Characterisation of the Structure and Activity of an Artificial Peroxygenase expressed in Yeast and Bacteria

Wendy Xiao Qin Robinson

Doctor of Philosophy

University of York

Chemistry

March 2022

1. Abstract

The oxyfunctionalisation of inert bonds in synthetic chemistry is a very desirable yet challenging process. Methods for this activation, especially for asymmetric products, often require harsh conditions, use valuable metals or need complex catalyst synthesis. Biocatalysis may provide a more favourable approach because reactions can be carried out under mild conditions and can be highly selective.

A group of heme-thiolate proteins called unspecific peroxygenases (UPOs), can carry out a diverse range of different oxyfunctionalisations using only hydrogen peroxide as a cosubstrate. The reactions carried out by UPOs are regio-, chemo- and enantioselective and they have a large substrate scope. In this work two UPOs were chosen for further investigation: an artificial UPO (artUPO) based on the UPO found in *Marasmius rotula* (*Mro*UPO), and the UPO derived from *Coprinus cinerea* (*Cci*UPO). The genes for artUPO and *Cci*UPO were cloned for *Pichia pastoris* expression but only artUPO resulted in successful purified recombinant protein. The gene for artUPO was also successfully expressed in *Escherichia coli*. Both forms of artUPO were evaluated by UV/visible activity assays which displayed a small reduction in activity compared to *Mro*UPO. The structure of both artUPO forms were determined by x-ray crystallography which developed our understanding of the mutations present in artUPO and enhanced understanding of general UPO structural biology. Substrate screens identified artUPO as effective biocatalyst for the asymmetric oxygenation of alkyl benzenes, styrenes and thioethers. With the latter, artUPO gives the complementary series of (S)-enantiomers to those obtained with the UPO from *Agrocybe aegerita* (*Aae*UPO), broadening the scope for application of the enzymes.

2. Contents

3. List of tables

4. List of Figures

5. Acknowledgements

First of all, I would like to thank Professor Gideon Grogan, Dr William Unsworth and Dr Alison Parkin for giving me the opportunity to work on this PhD project and for their guidance. I would also like to thank Professor Anne Duhme Klair for all of her support and advice not only throughout my undergraduate years, but also through my PhD as an IPM. I would like to thank Dr Jared Cartwright and Mick Miller in TF for their time and effort to teach me about yeast.

I would like to thank Louise, Simon, Jules and Tim Kirk from YSBL and David Neale from M2 for all of their technical support and kind faces. The labs would be nothing without them and they are nothing but legendary.

I would like to give a huge thanks to the Grogan Groupies past and present: Mark Petchey, Amina Frese, Muhi Omar, Mahima Sharma, Anibal Cuetos-Fernandez, Tamara Mielke/ Hughes, Ginevra Camboni, Megan Bennett, Verity Barber, Amelia Gilio, Zoe Ingold and Lorna (Qingyun) Tang. You have made me feel so welcome, taught a chemist to become a “biochemist”, and it has been wonderful to work and drink alongside you, but not at the same time of course. Thank you to my fellow revolutionary spirits Rhianna Rowland and Alex Snow for being incredible friends throughout this time, sharing the best and worst times of our PhD experiences. I also have to mention the whole of the YSBL family for making the department a great place to be and to boost morale with regular coffee. Thank you Jenny and Jo from Cookies for such a warm smile every day.

Thank you to my friends, parents, dogs (Winnie and Elsie) and family for all the love and support for all these years. You've made me smile, given me a shoulder to cry on and helped me through the tough times when I couldn't see myself finishing at all. Thank you Beth, Callum, Nick, Lydia, Oli and Dougie the dog for being here for my whole time at York, Molly, Harriet, Becca and Linda for all the love, and my sister Heather for being my best friend.

And finally thank you to Benjamin Rowlinson. I genuinely could not have made this journey without you in all aspects. Thank you for believing in me when I couldn't, and for all of your love and support.

I would not like to thank COVID-19 or my weak bones.

6. Author's Declaration

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

x-ray data were collected by Johan Turkenburg and Sam Hart either in house or at the Diamond Light Source. James Rossi-Ashton or William Unsworth collected and processed the ^{1}H-NMR spectra from synthesised compounds in the Chemistry department. Mass spectrometry experiments were carried out by Adam Dowle at the Technology Facility in the Department of Biology. Allison Dann performed the fermentation of rAaeUPO on industrial scale at GSK Worthing and Cyril Boudet assayed and concentrated the protein on site.

7. Introduction

7.1. Synthetic oxidations

Oxygenation reactions are important for numerous applications in synthesis, for example, the synthesis of human drug metabolites for drug metabolism and pharmokinetic (DMPK) studies, natural product synthesis and other synthetic intermediates or fine chemicals. However direct oxygenation, such as sulfoxidation or hydroxylation, can be challenging to due selectivity issues in oxidation of specific sites or achieving the correct stereoselectivity. Typical methods of oxidation include the use of *meta*-chloroperoxybenzoic acid (*m*-CPBA), hydroboration, ozonolysis and palladium but these methods suffer from lack of selectivity or control of the ratio of the primary and secondary oxidation products (Figure 7.1).

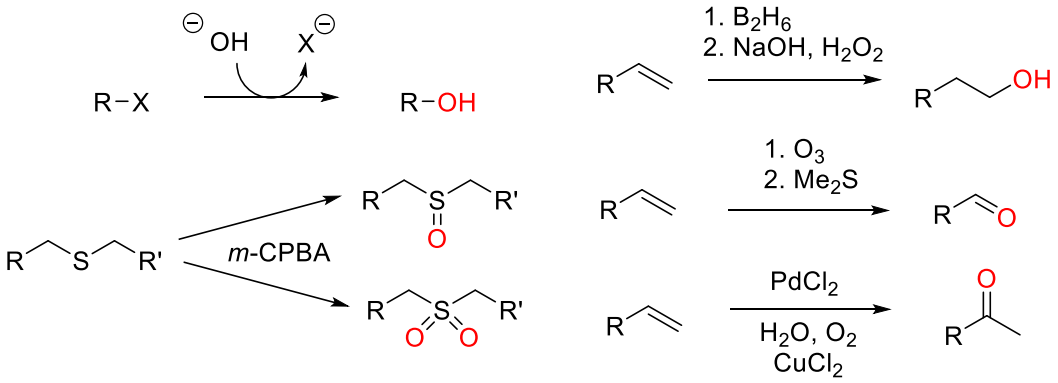


Figure 7.1 Typical methods of oxidation used in synthetic chemistry

Selective oxidations have been studied extensively to solve specificity problems which range in ease and complexity[1, 2] Examples include asymmetric dihydroxylation of unsaturated potassium trifluoroborates using osmium oxide and asymmetric

epoxidation using Jacobsen's catalyst, **4** as seen with the epoxidation of indene, **1** into indene oxide, **2** with use of ligand, **3** (Figure 7.2). [3, 4]

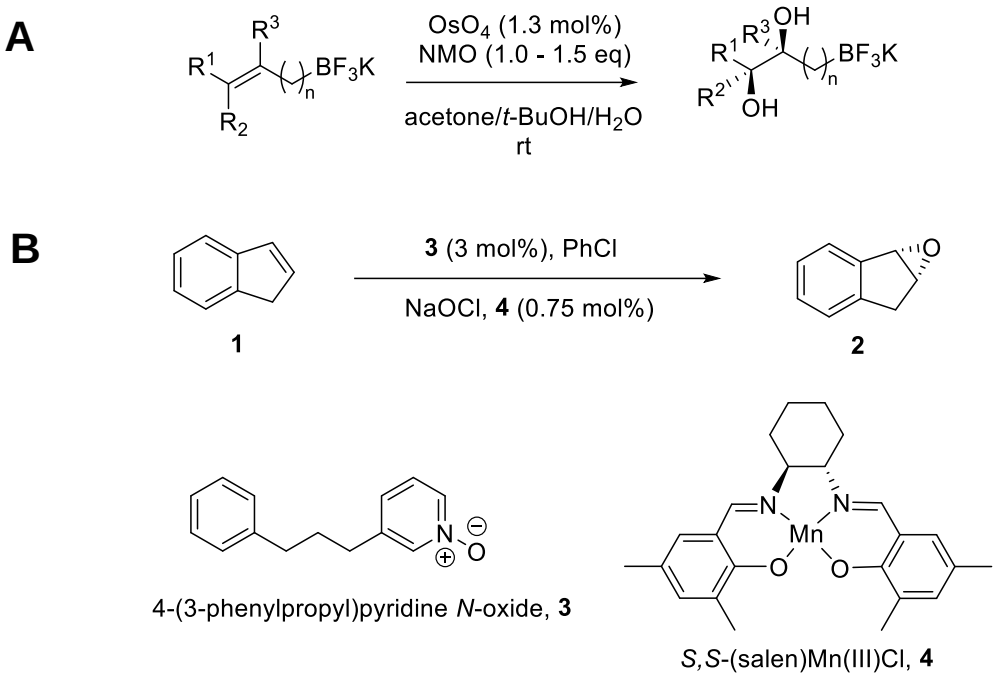


Figure 7.2 Examples of asymmetric oxygenations **A**: *Cis*-dihydroxylation of unsaturated potassium trifluoroborates. [3] **B**: Epoxidation of indene, **1** using Jacobsen's catalyst, **4**. [4]

Selective oxidations have been applied to the synthesis of human drug metabolites where starting materials are relatively complex. For example, the synthesis of the 10-hydroxy metabolite, **6** from the anti-malarial drug artemisinin, **5** required iron catalyst **7** and hydrogen peroxide to oxidise an inert C-H bond with retention of stereochemistry (Figure 7.3A). [5] The activation of an inert C-H bond is also seen in the hydroxylation of a cholesterol derivative, **8** (Figure 7.3B)**.** [6] The reaction occurs through an iridium-catalysed intramolecular silylation to activate a C-H bond γ from a hydroxyl group to form the 1,3-diol, **9**.

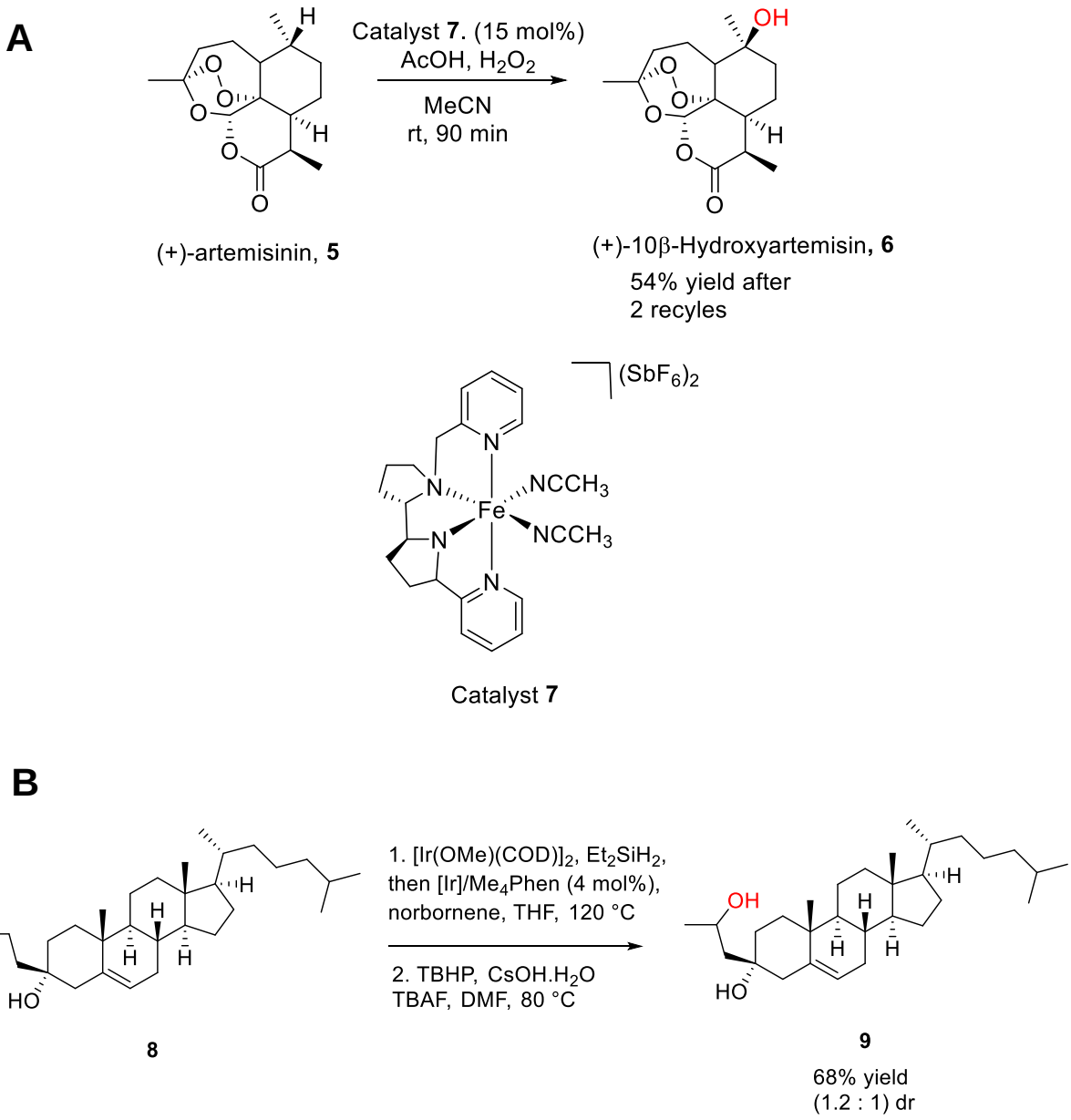


Figure 7.3 Asymmetric oxidation of biologically relevant compounds **A**: selective hydroxylation of artemisinin, **5** to form **6** using iron catalyst **7**.[5] **B**: iridium-catalysed C-H silylation to selectively form 1,3-diol, **9.**[6]

However despite efforts, all of these methods of asymmetric oxygenations in Figures 7.2 and 7.3 require either expensive metals, synthetically challenging catalysts, environmentally damaging solvents, or high temperatures to operate.

As well as oxygenations, oxidative halogenations are highly desired. Halogenations are of particular interest because the halogenated compounds can be used as synthetic intermediates for reactions such as Suzuki-Miyaura coupling or may have interesting medicinal properties. [7, 8] Traditional methods of halogenation include Sandmeyer reactions, Friedel-Crafts reactions, use of *N*-halosuccinimides and transition metal catalysts to activate aromatics, [9, 10] but these methods usually produce halogenated side products that are toxic to the environment and are difficult to dispose of. [11] These reactions also tend to lack selectivity where other regioisomers are formed or the compound is halogenated multiple times. [12] Examples of unspecific halogenations and polyhalogenation are seen in Figure 7.4, where chlorination of thymol, **10** with *N*-chlorosuccinimide and catalyst **13** results in two regioisomers **11** and **12**, [12] and bromination of phenol, **14** can result in subsequent brominations to yield **15**.

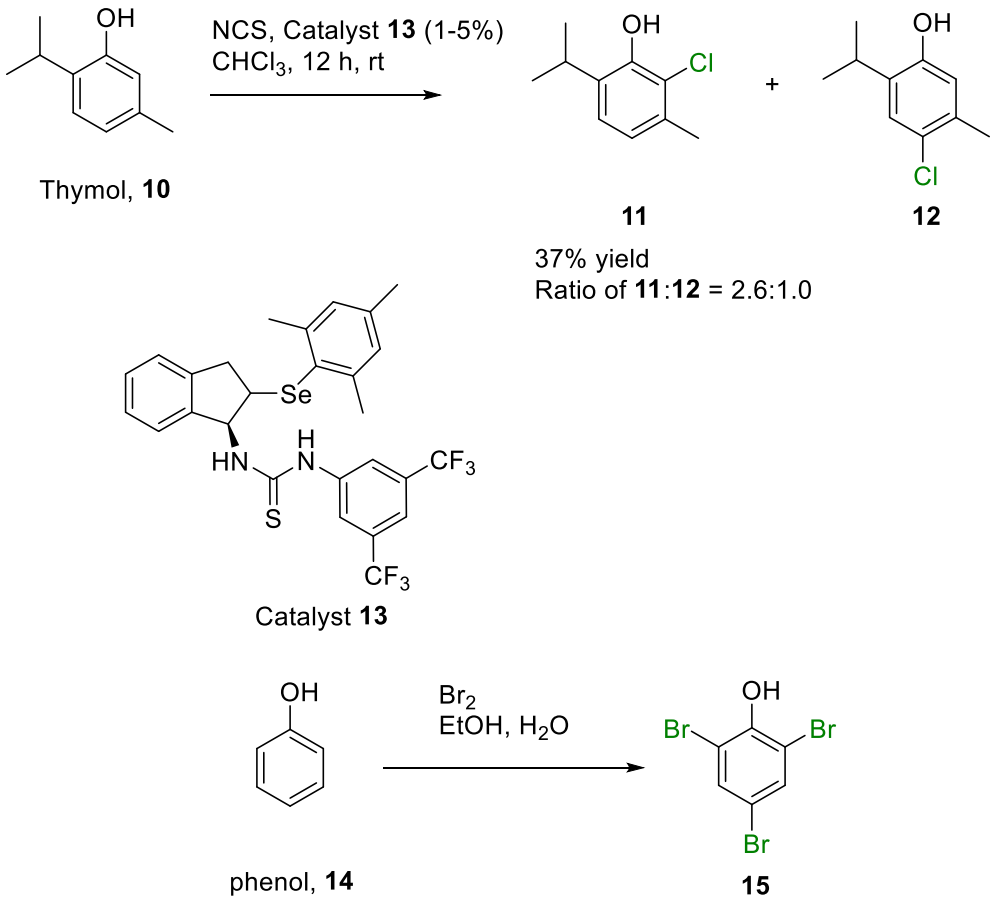


Figure 7.4 **Top**: Chlorination of thymol, **10** with catalyst **13** leads to a mixture of the *ortho* **11** and *para* **12** substituted chlorinated products. [12] **Bottom**: Bromination of phenol, **14** forms the tribrominated product **15** which may be undesirable

7.2. The role of oxidoreductases in C-H bond activation in biocatalysis

As seen in previous examples, synthetic chemistry can be used to access a variety of compounds using a variety of oxidation reactions. However, the catalysts and reaction conditions required for these transformations can be expensive, toxic or harmful to the environment. The use of biocatalysts has become increasingly popular as scientists look towards nature for green and sustainable regiospecific and stereospecific transformations. [13] Biocatalysts can reduce the need for harmful organic solvents, expensive precious metals and potentially dangerous chemicals often associated with difficult oxidation reactions.

Oxidoreductases, particularly oxygenases, have significant interest for industrial biotransformations due to their ability to activate inert C-H bonds in nature under mild conditions. [14] Heme-dependent oxidoreductases have been extensively studied due to both their relevance in drug metabolism and their broad reaction scope. [15] The protoporphyrin environment allows the heme-iron to reach unusually high oxidation states, which in turn allows for challenging oxidation reactions such as C-H activation. [16] The substrate binding site of these enzymes also allows for higher specificities than those seen with classic chemical oxidants, allowing for stereo- and regio- specific product formation without the need of protecting groups. [17]

The work described in this thesis is focused on investigating a group of heme-thiolate proteins (HTPs) for industrially relevant biotransformations. Here presented is a brief

overview of the development of HTPs for industrial application, highlighting a role that could be filled by unspecific peroxygenases (UPOs).

HTPs contain a protoporphyrin IX prosthetic group which contains an iron atom coordinated in the centre (Figure 7.5). [14] The heme within the enzyme has two sides; the proximal side has a cysteine residue coordinated to the iron atom and is involved in catalysis, the distal side allows for coordination and oxidation of the substrate.

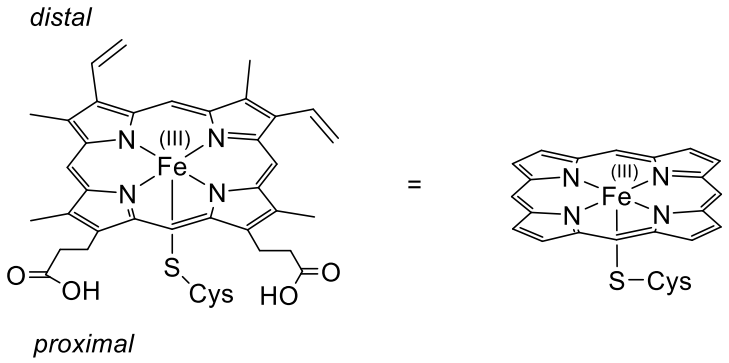


Figure 7.5 Protoporphyrin IX group with an iron (III) coordinated in the centre, these together form the heme group. The proximal side has a cysteine residue coordinated in HTPs and the distal side allows for catalysis. The right side is a simplified version of the heme group

7.2.1. Chloroperoxidases

Chloroperoxidase (CPO) was discovered in the 1960s from the fungus *Caldariomyces fumago* (*Cfu*CPO) (now known as *Leptoxyphium fumago*) and has since been extensively studied due to its potential applications in biotechnology. [18] In nature, *Cfu*CPO is involved in the synthesis of the antibiotic caldariomycin, **18** where *Cfu*CPO is involved with chlorinating 1,3-cylcopentanedione, **16** twice (Figure 7.6), however it

is capable of other reactions such as brominations, iodinations, sulfoxidations and epoxidations using only hydrogen peroxide as a cosubstrate. [19, 20]

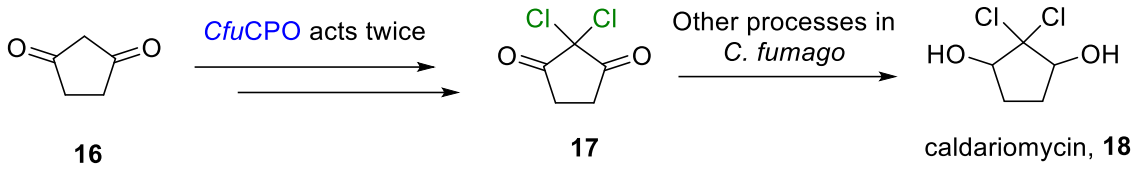


Figure 7.6 *Cfu*CPO di chlorinates **16** in its native host in part of the biosynthesis of the caldariomycin, **18**

7.2.1.1. Structure and mechanism of the heme-dependent CfuCPO

*Cfu*CPO is a monomeric protein with eight helices, surrounding a protoporhyrin prosthetic group (Figure 7.7). It is unique compared to other heme peroxidases because it contains a cysteine ligand that is coordinated through the sulfur atom to the iron centre, rather than a histidine residue which is observed in other heme peroxidases.

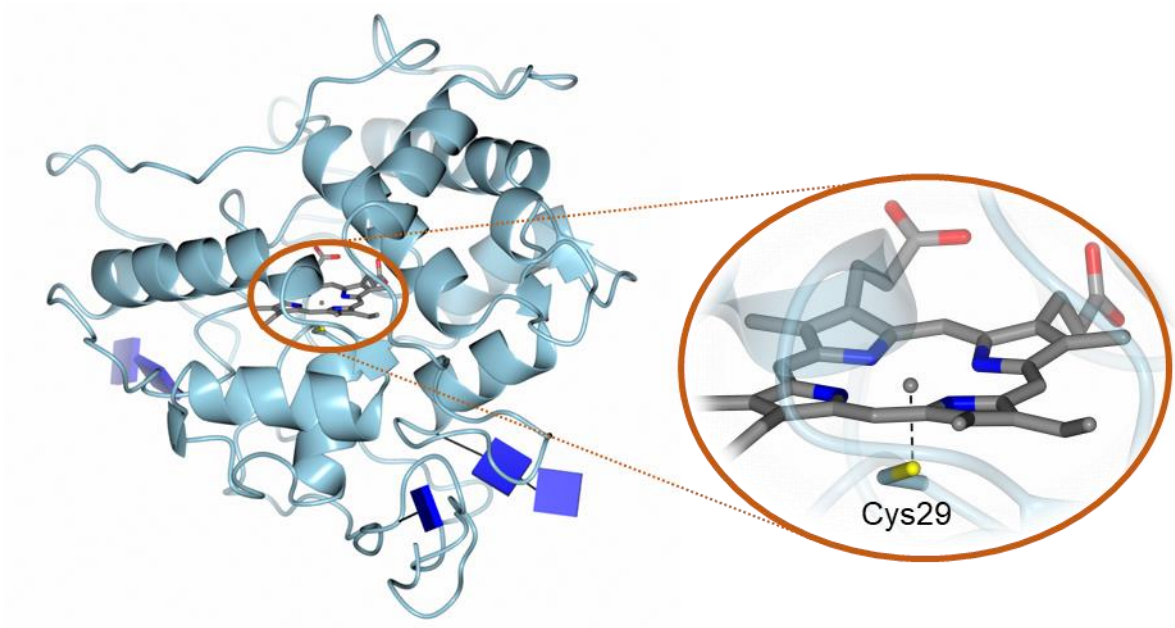


Figure 7.7 Ribbon structure of choloperoxidase (2CPO) coloured in pale blue with a close up of Cys29 coordinated to the heme (2.28 Å). The dark blue squares represent

glucosamide and the heme ligand is highlighted using a cylindrical representation where carbon is grey, oxygen is red and nitrogen is blue.

The proposed mechanisms for halogenation and hydroxylation are given in Figure 7.8, where hydrogen peroxide is used to form compound I in both cases. This step is catalysed by a glutamate residue, which is further involved in a distal charge stabilising pair with histidine. [21] In the halogenation reaction, a halide is introduced to form an iron (III) hypohalite intermediate which rapidly degrades to form a hypohalous acid and the enzyme returns to the resting state. [22] The hypohalous acid then diffuses out and halogenates an activated compound, however this halogenation is not inherently selective due to the free nature of the hypohalous acid and so the heme environment of *Cfu*CPO would not be expected to influence selectivity. [23] In the absence of a halide, *Cfu*CPO is able to oxygenate compounds through a peroxygenase mechanism (Figure 7.8). [24-26] Compound I is able to extract a proton from the substrate which forms a radical. The radical can then interact with Compound II and then form the hydroxylated compound. Both reactions avoid the need for expensive cofactors or partner proteins which has made CPOs very attractive for synthetic applications.

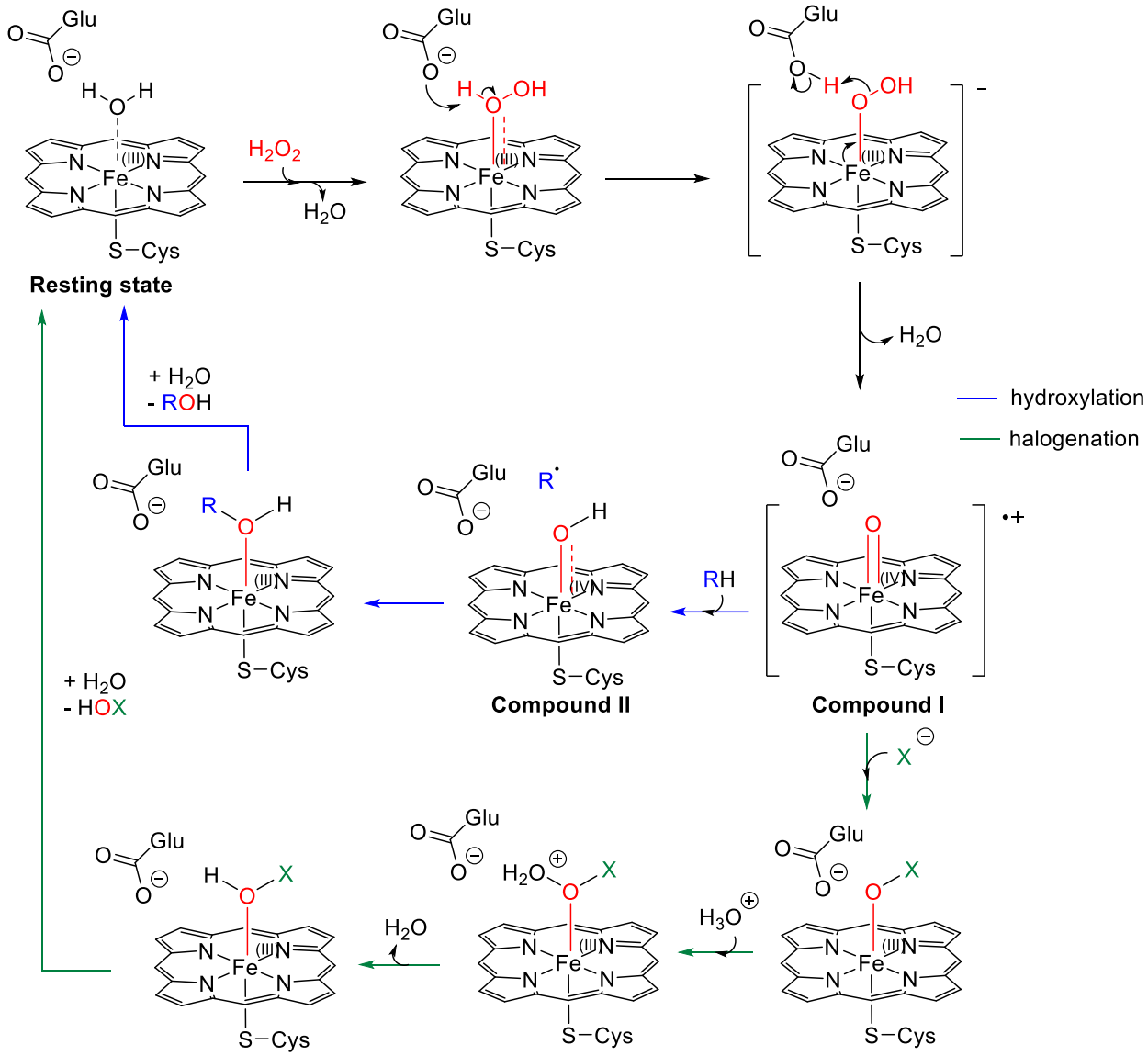


Figure 7.8 The mechanism of halogenation and hydroxylation in *Cfu*CPO. In both mechanisms, a glutamate residue catalyses the formation of Compound I through hydrogen peroxide. With the formation of Compound I, *Cfu*CPO can either halogenate with the presence of a halide, or hydroxylate *via* a peroxygenative mechanism

7.2.1.2. The reaction scope of *Cfu*CPO

*Cfu*CPO has the ability to chlorinate electronically activated compounds. In the native reaction, *Cfu*CPO chlorinates 1,3-cyclopentadione **16** to 2,2-dichloro-1,3-cyclopentadione, **17** in two steps for the synthesis of caldariomycin, **18** (Figure 7.6). [18] *Cfu*CPO was also able to chlorinate the terpene thymol **10**, where the *ortho* substituted

compound **11** was the major product whereas the *para* substituted compound, **12** was the minor product (Figure 7.9). [27] It has shown other halogenating capabilities such as bromination of *trans* cinnamic acid, **19** to yield **20** and **21**, and iodination of tyrosine, **21** to form **22** (Figure 7.9) but interestingly fluoride ions act as inhibitors of *Cfu*CPO because these can compete for the hydrogen peroxide or the anion-binding sites on *Cfu*CPO. [28, 19]

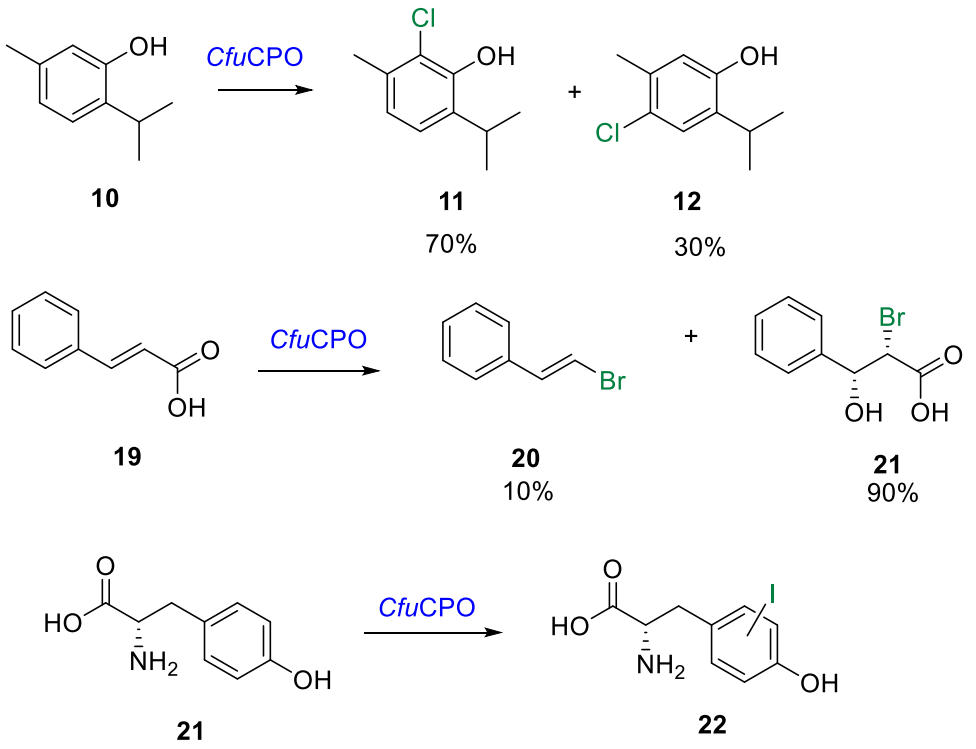


Figure 7.9 Halogenation reactions carried out by *Cfu*CPO, substrates include thymol **10**, [27] *trans* cinnamic acid, **19**, [28] and tyrosine, **21**[19]

Through its peroxygenative activity, *Cfu*CPO has been able to carry out a variety of oxygenation reactions including chiral sulfoxidations, [29-32] epoxidations[25,33,34] and hydroxylations[25,35-38] as shown in Figure 7.10. In all these reactions, where applicable, high enantioselectivity was observed with >90% *ee* being typical.

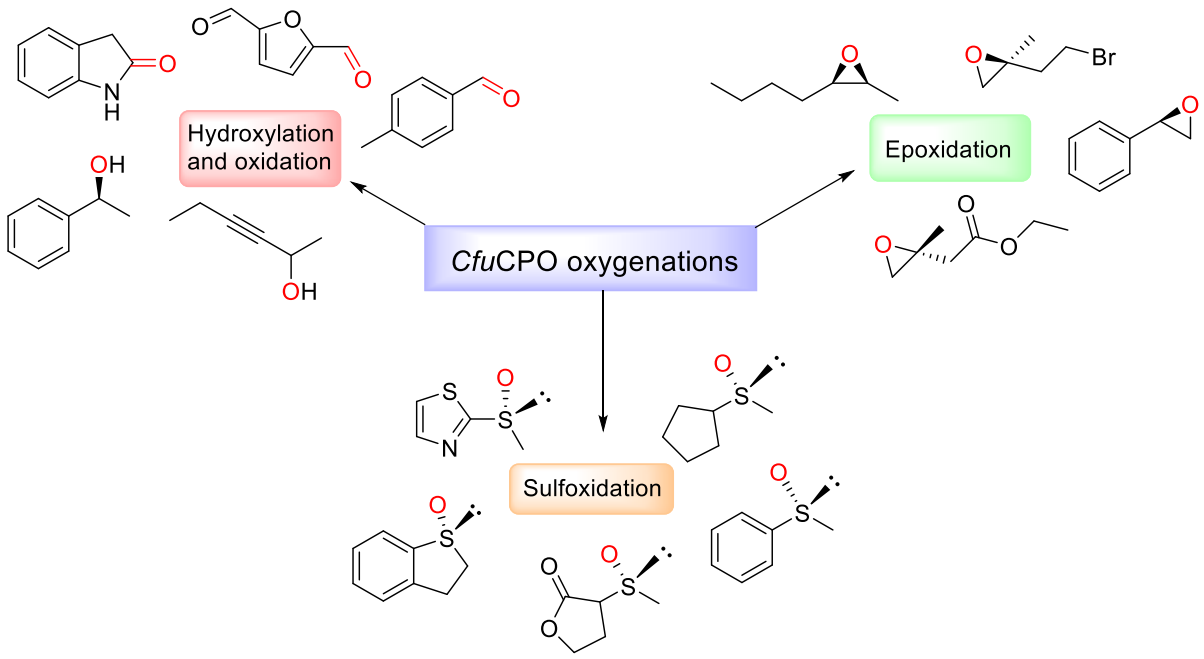


Figure 7.10 *Cfu*CPO oxygenations such as sulfoxidations, epoxidations, and hydroxylations

*Cfu*CPO has also been used in a cascade with other enzymes to carry out Achmatowicz-type rearrangement of furylcarbinols such as **25** using 10% *tert* butanol as a cosolvent (Figure 7.11). [39] This study suggests that CPO can oxygenate the furan core, which goes on to rearrange to give the pyranone products.

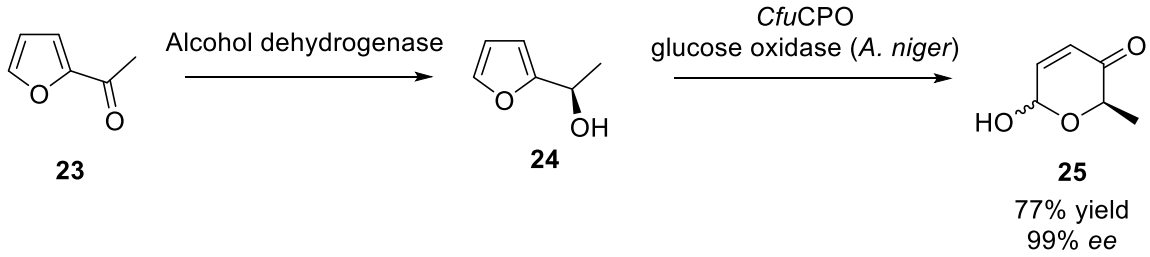


Figure 7.11 A biocatalytic Achmatowicz rearrangement through a multi-enzyme cascade

Despite the promising reaction scope, *Cfu*CPO is limited because it cannot react with non-activated carbons found in aromatic systems, alicyclic rings or *n*-alkanes. Other biocatalysts must be considered if these reactions are to be carried out enzymatically.

7.2.2. Cytochrome P450 monooxygenase enzymes

Cytochrome P450 monooxygenases (P450s) are a large superfamily with over 300,000 P450 sequences found in all kingdoms of life. [40] They have been given this name because they have an intense spectrophotometric absorbance band at 450 nm when reduced with carbon monoxide. [41] Like chloroperoxidases, they have a globin fold with a heme prosthetic group coordinated by an axial cysteine. They are able to oxyfunctionalise sp^3 and sp^2 inert C-H bonds in a regio-, chemo- and stereospecific manner under mild conditions, which has made them popular biocatalysts for the synthesis of a variety of compounds, such as pharmaceuticals, human drug metabolites, natural products and fine chemicals. [42, 43, 44]

7.2.2.1. P450 classes and mechanism

All P450s require a source of electrons and they can be found in the cofactors nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADPH). P450s also need a redox partner protein to transfer the electrons from NAD(P)H to the iron in the heme centre. There are multiple different classes of P450 systems where the organisation of the P450 and the redox partner varies. These can be broadly divided into three groups: P450s that interact with redox partners or cofactors, P450s that are tethered to redox proteins, and P450s that are tethered to proteins not involved with P450 reduction. Some examples of these are shown in Figure 7.12. [43,45,46]

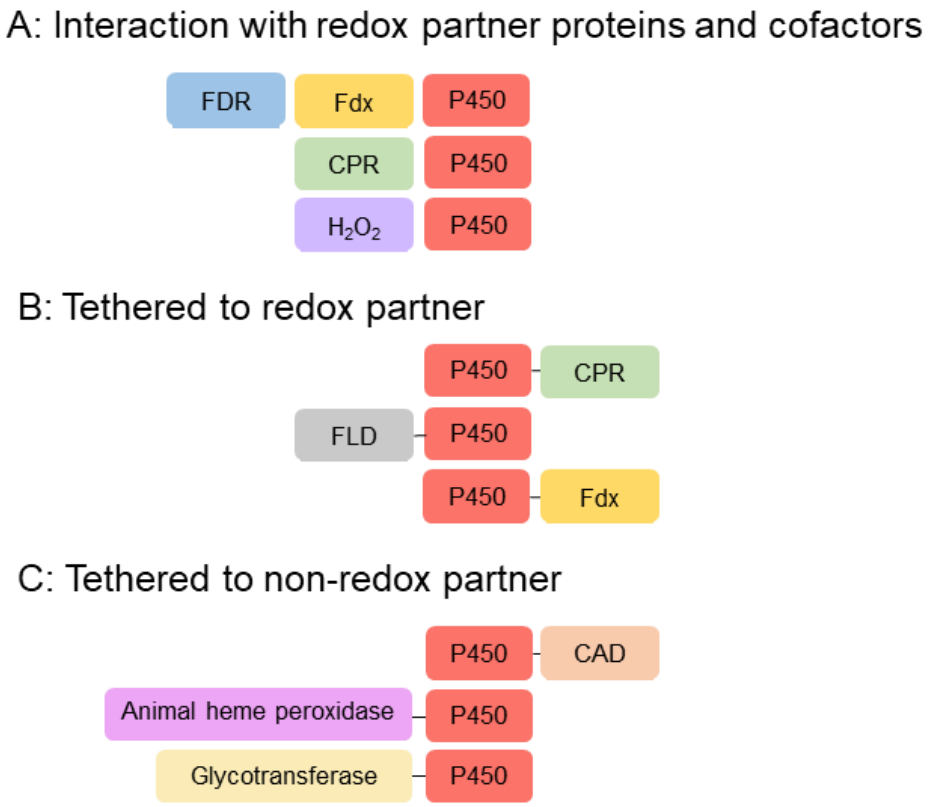


Figure 7.12 The different classes of P450s divided into three groups. **A**: P450s which interact with partner proteins and cofactors including a FAD/NAD(P)H-binding reductase (FDR) and a ferredoxin (Fdx) (Class I), a cytochrome P450 reductase (CPR) (Class II) and H_2O_2 P450 peroxygenases. **B**: P450s which are tethered to redox partners such as a CPR (P450 BM3), FMN-binding flavodoxins (FLD) or Fdx. **C**: P450s tethered to a non-redox partner such as cinammyl alcohol dehydrogenase-type module, animal-like heme peroxidase and glycotransferase-like module

Like CPOs, P450 catalysis occurs on the distal side of the heme and an example of hydroxylation is shown in Figure 7.13. In Stage 1, the substrate displaces a distal water molecule in the resting state and the ferric iron changes spin state from low spin to high spin. [47] This change in spin increases the redox potential of the heme to allow for the reduction in Stage 3, where an electron is provided by the cofactors NADH or NADPH. [48, 49] In Stage 4, the complex binds to molecular iron to form a ferric-oxo

species. [43, 50] Further reduction leads to the formation of the ferric peroxo state, which is quickly protonated to form the transient ferric hydroperoxo state, Compound 0. Protonation occurs again in Stage 5, where the O-O bond is heterolytically broken to form water and porphyrin radical cation called Compound I. [51, 52] Compound I is then able to oxygenate the substrate and the product is released, allowing water to bind to the iron and the resting low spin iron is restored, ready for another cycle. [43, 44] Non-productive routes can also occur alongside the favoured mechanism, for example in the autooxidation route the ferric-oxo species in Stage 4 can return to the beginning of the cycle and superoxide is formed. Compound 0 in Stage 5 can degrade to form hydrogen peroxide in the peroxide shunt route, but this route can be reversible, and therefore productive, if an excess of peroxide is added to the system. [53] For most naturally occurring P450s, this step is very inefficient. The third alternative route involves the degradation of Compound I in Stage 6 to form water in the oxidase shunt. [44] These alternative shunts can occur when the delivery of electrons or protons are inefficient and so the productive cycle is disrupted.

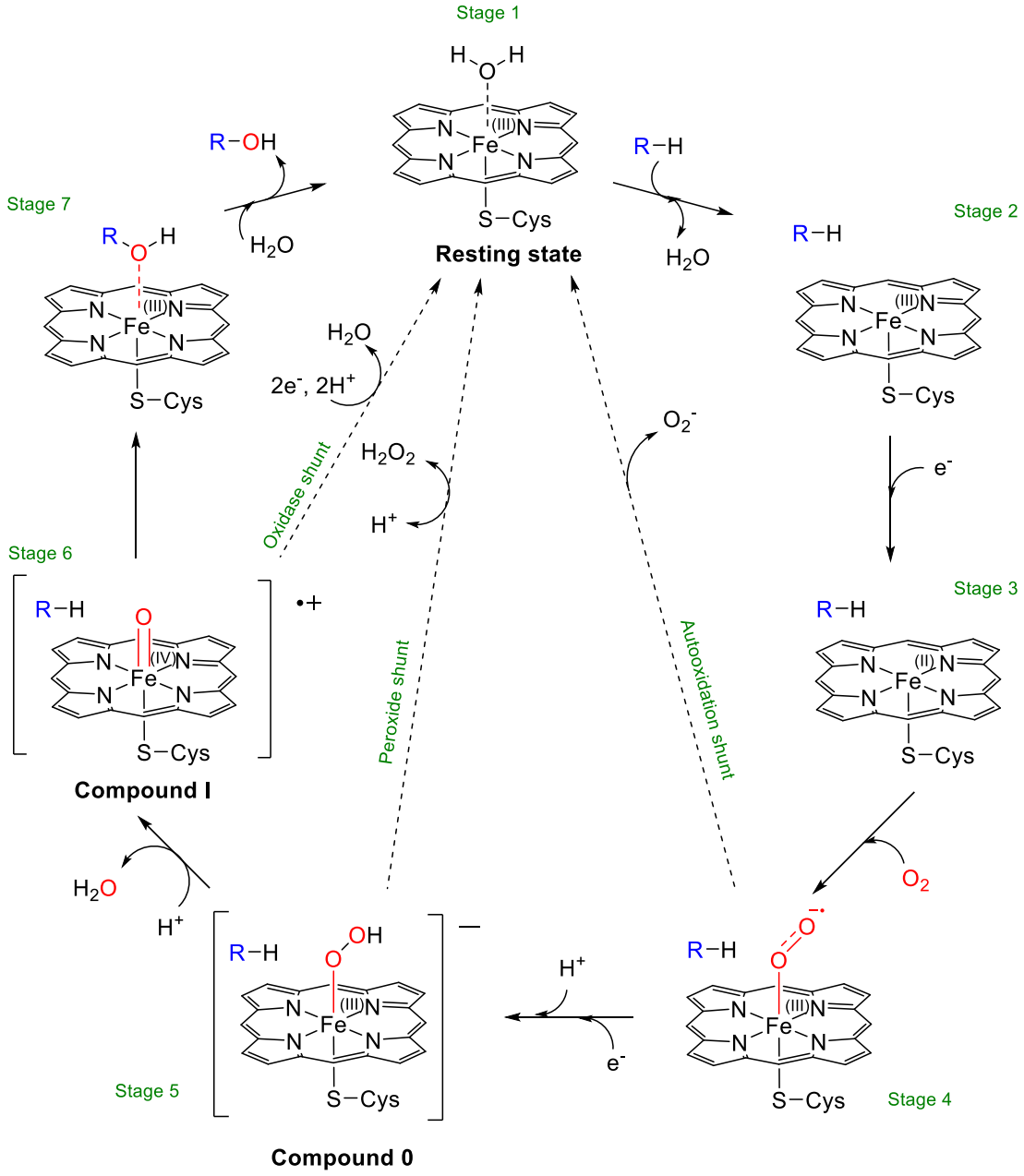


Figure 7.13 Catalytic cycle for the hydroxylation of a substrate carried out by P450s. Stage 1: A substrate displaces water on the distal site and the iron changes from low spin to high spin. Stage 2: an electron provided by a redox partner is delivered to reduce the iron to the ferrous form. Stage 3: molecular oxygen is able to bind to form the ferric-oxo species. Stage 4: a second electron from a redox partner is delivered. This is protonated to form Compound 0. Stage 5: Compound 0 receives another proton which forms water and reactive species Compound I. Stage 6: Compound I abstracts a hydrogen atom from the substrate and rebounds to form the hydroxylated product coordinated to the heme. Stage 7: the product is released which allows water to bind and the system returns to resting state

7.2.2.2. P450s for use in industry

P450s have great potential for use in industry due to their reaction and substrate scope. One of the most successful applications of P450s is for the production of pravastatin, **27** developed by Daiichi-Sanko and Bristol Myers Squibb (Figure 7.14). [54] In this process, the precursor compactin, **26** is selectively hydroxylated by CYP105A3 in a fermentation with *S. carbophilus*. Reactions carried out by P450s in whole cells are also used for the production of pharmaceutical hydrocortisone, **29**, scent and flavour of grapefruit nootkatone, **31** and the intermediate artemisinic acid, **33** for the synthesis of artemisinin, **5** (Figure 7.14). [55-58]

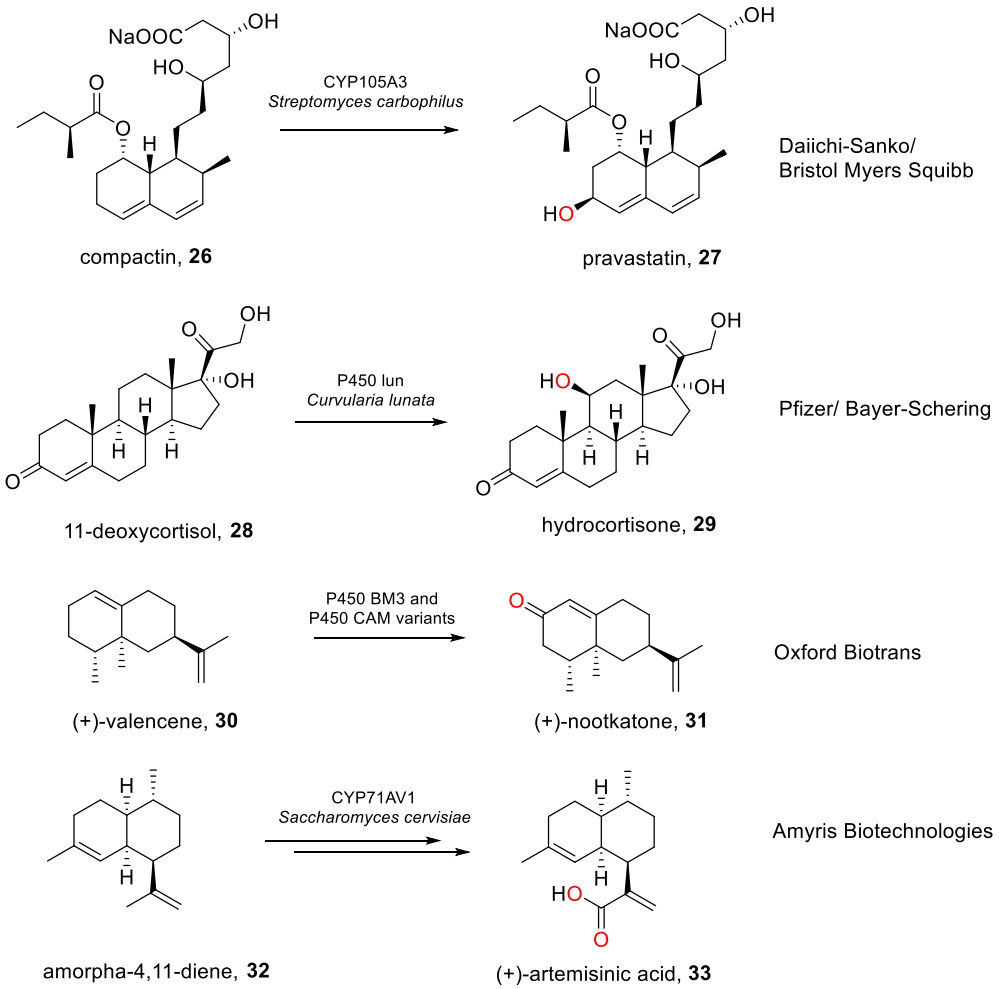


Figure 7.14 Industrial applications of P450s to produce the pharmeceuticals pravastatin, **27**, hydrocortisone, **29**, the fragrance nootkatone, **31** and synthetic intermediates such as (+)-artemisinic acid, **33**

7.2.2.3. Limitations of P450s

Although P450s have great potential and have proven utility in several scenarios, there are still issues that limit their application synthetically and industrially on a large scale. Firstly, P450s need the cofactor NAD(P)H as an electron source which is very costly. To address this issue, the cheap NADH biomimetic compound *N*-benzyl-1,4-dihydronicotinamide, used alongside an evolved P450, has been shown to be an effective alternative cofactor to NADH, potentially reducing financial costs. [59] Alternatively *in situ* cofactor recycling systems have been implemented to regenerate NAD(P)H through enzymatic, chemical, or electrochemical means. [60, 61]

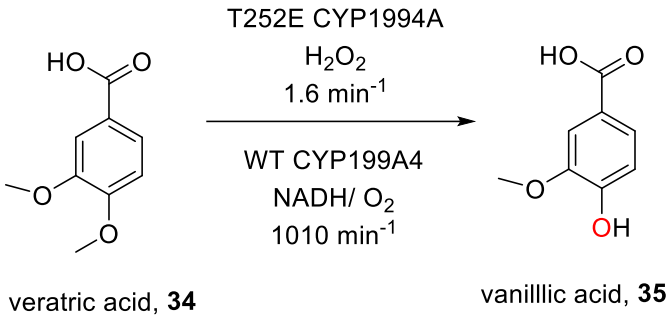
The rate and efficiency of electron transfer in P450s is paramount to its function and so uncoupling reactions (as mentioned in Section 7.2.2.1) tend to decrease the catalytic efficiency which leads to poor catalytic control of the enzyme. Self-sufficient redox systems, such as P450 BM3 isolated from *B. megaterium* and P450 Rhf from *Rhodococcus sp.* strain NCIMB 9784, have proven to be very effective catalysts because of their high catalytic turnovers and the simplicity of using a fused system with improved electron transfer efficiency. [62, 63] Further studies on these systems, including protein engineering and fusion of the reductase domain to other P450s, has vastly improved their application where reactivity, substrate specificity, and selectivity were enhanced. [43, 45, 15]

Another problem is that P450s are generally unstable towards both heat and solvents, especially because many P450s are membrane bound or have multiple domains that are essential for their function. Different approaches have been taken to solve the

problem with stability, including protein engineering, immobilisation and using whole cell catalysis. [45, 64]

To circumvent many of the problems found in P450s, some research has turned to exploiting the peroxide shunt of P450s, where the reverse of this reaction can avoid complexities such as redox-partner proteins, molecular oxygen and cofactors. Instead hydrogen peroxide can be used as a cosubstrate and is the oxygen and electron donor. Cirino *et al.* 2003 were able to evolve the heme domain of P450 BM3 for the hydroxylation of medium chain fatty acids and the epoxidation of styrene *via* peroxygenase activity. [65] This work has demonstrated the powerful use of protein engineering to tailor enzymes for desirable activities. Other examples of engineering of P450 BM3 for peroxide driven biotransformations include hydroxylation of small and medium alkanes, production of atorvastatin metabolites, and regioselective aromatic *O*-demethylation. [66-69] A different approach to encourage peroxygenase activity in P450 BM3 was carried out by Ma *et al*. 2018 using “dual-functional small molecules” (DFSMs). [70] In this study, different N-(ω-imidazolyl fatty acyl)-L-amino acids were anchored in the P450 active site and the imidazole group acted as an acid-base catalyst to activate hydrogen peroxide, which is similarly seen in unspecific peroxygenases (UPOs), to carry out hydroxylations, sulfoxidations and epoxidations. Inspiration from the mechanism of UPOs has also been demonstrated in a different evolved P450, CYP199A4 derived from *Rhodopseudomonas palustris* HaA2, where a highly conserved threonine in P450s was mutated to a glutamate residue. [71] The glutamate residue also acted as an acid-base catalyst and was able to activate the peroxide heme complex to form compound I. The CYP199A4 T252E mutant, like the wild type, was able to regioselectively demethylate veratric acid (3,4-

dimethoxybenzoic acid), **34** in the para position to form exclusively vanillic acid, **35**, presumably *via* an initial methyl oxidation (Figure 7.15). However the rate of product formation was much lower by a factor of 650 compared the wild-type using cofactors.



veratric acid, **34** vanilllic acid, **35**

Figure 7.15 *O*-demethylation of veratric acid, **34** with the mutant CYP199A4 using peroxygenase activity and the wild type CYP199A4 using monooxygenase activity. The production formation (micromoles per micromole of P450 per minute) is vastly different between the two forms. [71]

In some cases, P450 naturally use hydrogen peroxide as an oxidant and these are called P450 peroxygenases such as P450 SP_{α} from *Sphingomonas paucimobilis*, P450 BS_{β} from *Bacillus subtilis* and P450 CLA from *Clostridium acetobutylicum*, all of which were able to selectively hydroxylate fatty acids. [72, 73, 74] The P450 $OleT_{JE}$ from *Jeotgalicoccus* sp. ATCC 8456 is another P450 peroxygenase and was able to decarboxylate fatty acids to produce α-alkenes. [75] In all of the natural P450 peroxygenases, the acid-alcohol amino acid pair usually seen in P450s is replaced by arginine and proline, suggesting that these enzymes have different evolutionary routes to other P450s. [76] Despite efforts to discover new P450 peroxygenases or optimise current P450s for peroxygenase activity, unspecific peroxygenases provide an alternative as they are stable and are reactive to a variety of substrates as discussed in the next section.

7.3. Unspecific peroxygenases (UPOs) as biocatalysts

In 2004, Ulrich and Hofrichter were the first to report an unspecific peroxygenase (*Aae*UPO) excreted from the basidiomycetous fungus *Agrocybe aegerita* (now *Cyclocybe aegerita*), also known as the black poplar mushroom, and was originally described as a novel haloperoxidase (AaP). [77] The enzyme was later renamed as an unspecific peroxygenase (EC 1.11.2.1) and was the first member of a new oxidoreductase sub-class. [78] Initial studies of *Aae*UPO carried out by the Hofrichter group showed that this enzyme could catalyse aromatic hydroxylations, aryl alcohol and aldehyde oxidations and halogenations using only H_2O_2 as a cosubstrate at neutral pH. [79] *Aae*UPO has since been demonstrated to catalyse a large variety of oxidations which involve the activation of inert C-H bonds. The large reaction scope of these proteins and the lack of cofactor dependency make UPOs viable candidates for the use in industrial biocatalysis. Since this work, a large number of UPOs have been found in the fungal kingdom but only a select few have been studied to investigate their reaction capabilities. Further investigation of the reaction and substrate scope of less studied UPOs will help build a UPO toolbox to allow for a wider range of targets to be produced. The native function of UPOs is currently unknown however they are secreted from the fungus and are believed to be involved in detoxification or the degradation of large biomolecules such as lignin and humus. [80, 14]

7.3.1. The proposed mechanism of UPO catalysis

Both UPOs and P450s can carry out similar oxidation reactions however, there are key differences between their reaction mechanisms. P450s incorporate an oxygen

atom from molecular oxygen into substrates and are reliant on electron transfer systems to carry out several one electron transfers. [44] UPOs possess both peroxygenative (single two-electron transfer) and peroxidative activity (one-electron transfer) using only H_2O_2 as a cosubstrate but it is the peroxygenative activity that is of the greatest interest because of its ability to activate C-H bonds.

In the proposed mechanism (Figure 7.16), hydrogen peroxide replaces water as the distal ligand from its resting state and the ferric iron centre changes spin state from low spin to high spin (Stage 1). [81] The glutamate residue then promotes heterolytic cleavage of the O-O bond in Stages 2 and 3 to form an oxoiron (IV) porphyrin radical cation, Compound I. [82] In Stage 4 Compound I abstracts a proton from the substrate to form Compound II. The cysteine residue in compound II is ligated to the heme iron which causes the newly formed hydroxyl group to have a basic pKa of 10.0. [83] This basic character is believed to be essential as it provides a driving force for the hydrogen atom abstraction of compound I and is necessary for difficult C-H bond scission. The newly formed substrate radical combines with the hydroxyl group in Stage 5 to form the hydroxylated product and a new catalytic cycle begins. [81]

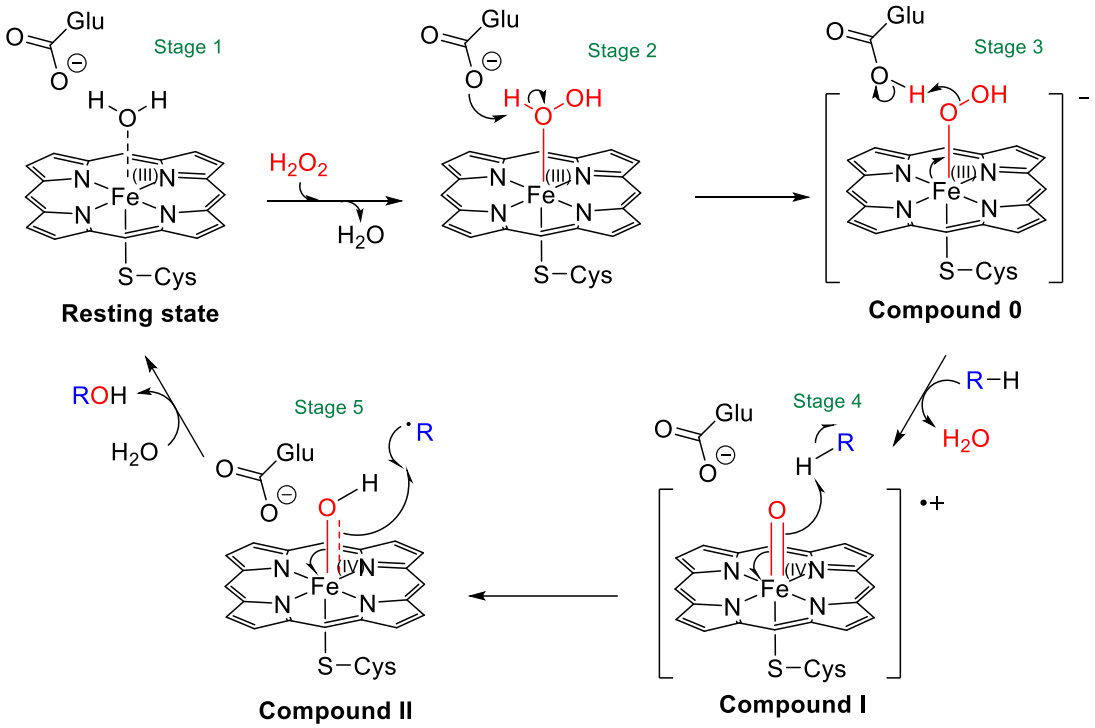


Figure 7.16 Schematic of the proposed catalytic hydroxylation mechanism for UPOs. Stage 1: hydrogen peroxide displaces water on the distal site and the iron changes from low to high spin. Stage 2: a glutamate residue abstracts a proton to form Compound 0. Stage 3: protonated glutamine facilitates the cleavage of the O-O bond to form Compound I. Stage 4: Compound I abstracts a proton from the incoming substrate to form Compound II. Stage 5: the cysteine residue assists in radical rebound to form the hydroxylated product and the cycle restarts with an incoming molecule of water.

7.3.2. Structural insights into UPOs

UPOs share many structural similarities with both P450s and CPOs with all of these proteins adopting a globin type fold. They are heme-thiolate proteins (HTPs) where a cysteine residue acts as a proximal ligand to the heme group. UPOs are classified into two groups based on their size, “short” UPOs (Group I) or “long” UPOs (Group II). [84] Short UPOs typically have a molecular weight of roughly 30 kDa and are found

throughout the fungal kingdom, whereas long UPOs usually have a molecular weight of around 45 kDa and are only found in ascomyota and basidiomycota. [85] This cysteine residue is found in a “proline-cysteine-proline” motif which is found in all UPOs. [86] UPOs are typically *N*-glycosylated in their native form and this glycosylation has been reported to as high as 42% in the UPO from *Coprinus verticillatus*. [87]

Oxidations in the active site of UPOs occur on the distal side of the heme group and are facilitated by a distal charge stabilising pair. In long UPOs, this pair involves a glutamate and an arginine residue but in in short UPOs, like in CPO, the basic residue is a histidine instead (Figure 7.17A). [88] Close to the heme active site, UPOs possesses a cation binding site which contains a structural Mg^{2+} ion which is coordinated in an octahedral fashion with one of the proprionate groups of the heme, the three protein residues and two molecules of water (Figure 7.17B).

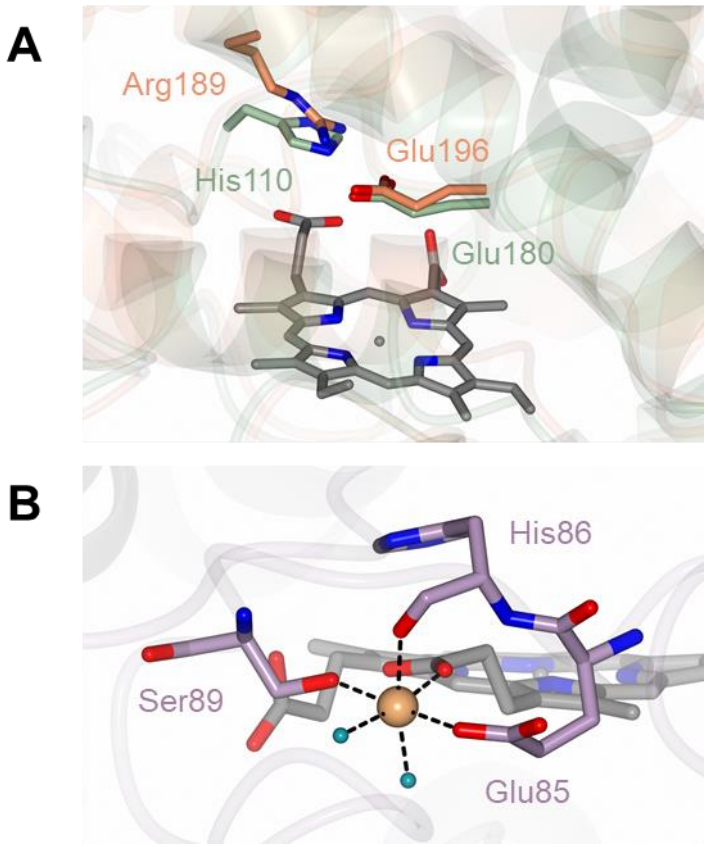


Figure 7.17 **A** Overlay of acid base pair in a long UPO *Aae*UPO (2YOR orange) and a short UPO *Hsp*UPO (7O1R green) **B** A structural Mg^{2+} ion (brown) sits near the heme group in *Mro*UPO and is coordinated in an octahedral geometry with two water molecules (blue), a heme proprionate group and three residues

Currently in the protein databank a small range of UPO structures have been solved, including *Aae*UPO and its variants, *Mro*UPO (from *Marasmius rotula*), and *Hsp*UPO (from *Hypoxylon* sp. EC38) with or without ligands. [89-91] All of these structures were solved by protein x-ray crystallography. Short UPOs, such as *Mro*UPO and *Hsp*UPO, have eight α-helices, four located on the proximal side and four located on the distal side of the heme group and also have a small antiparallel β-sheet pair (Figure 7.18). *Aae*UPO, a long UPO, has two additional helices at the *C*-terminal end of the protein and five short β-sheets (Figure 7.18).

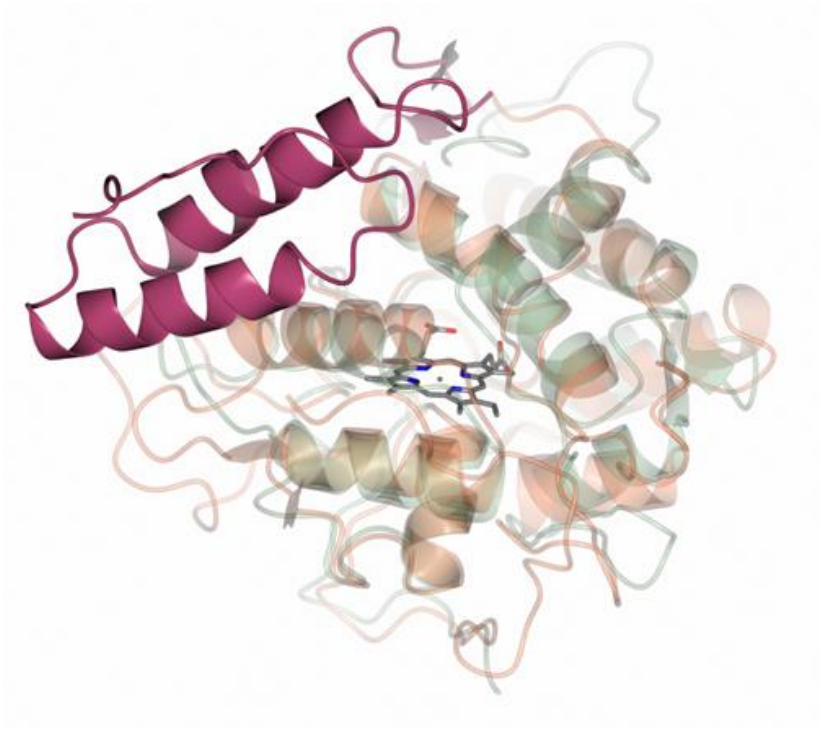


Figure 7.18 Ribbon structure an overlay of a short UPO *Hsp*UPO (7O1R in green) and a long UPO AaeUPO (2YOR in orange), with the two additional helices in AaeUPO coloured in maroon. The heme ligand is highlighted using a cylindrical representation where carbon is grey, oxygen is red and nitrogen is blue.

All known UPOs have mannose *N*-glycosylation sites at asparagine residues on the surface of the proteins, but the extent of glycosylation varies depending on the host organism and the UPO itself. *Mro*UPO exists as a dimer and the two units are joined together by a disulfide bond. *Aae*UPO and *Hsp*UPO on the other hand exist as monomers. *Aae*UPO also has an internal disulfide bond The potential effects of dimerization on stability and activity of UPOs would be an interesting area for further study.

The substrate channel and active site residues for each UPO differ and are believed to give rise to the differences in the selectivity of substrates. *Aae*UPO has a cone-shaped substrate channel leading to the active site with a diameter of 8.5 Å which is

mostly non-polar and aromatic in character. [89] Five of the phenylalanines in this tunnel are considered to be key in the selectivity in *Aae*UPO; three phenylalanine residues are found close to the active site and are believed to orientate the substrate. The other two phenylalanine residues are located near the entrance of the channel and potentially restrict the degrees of freedom of the substrate. [89, 92] In the case of *Mro*UPO, the substrate channel consists of mostly leucine and isoleucine residues which gives rise to a wider substrate channel (12.9 Å diameter) than in *Aae*UPO with a more aliphatic character (Figure 7.19A and 7.19B). [93] The substrate channel in *Mro*UPO is also shorter than in *Aae*UPO. The substrate tunnel in *Hsp*UPO differs from both *Aae*UPO and *Mro*UPO due to the presence of some polar residues such as serine, asparagine and lysine at the top of the tunnel (Figure 7.19C). [90] The differences in the residues of the substrate channel affect which compounds can be accepted by different UPOs. The implications of these effects can be seen in the protein engineering work carried out on UPOs discussed later in this chapter.

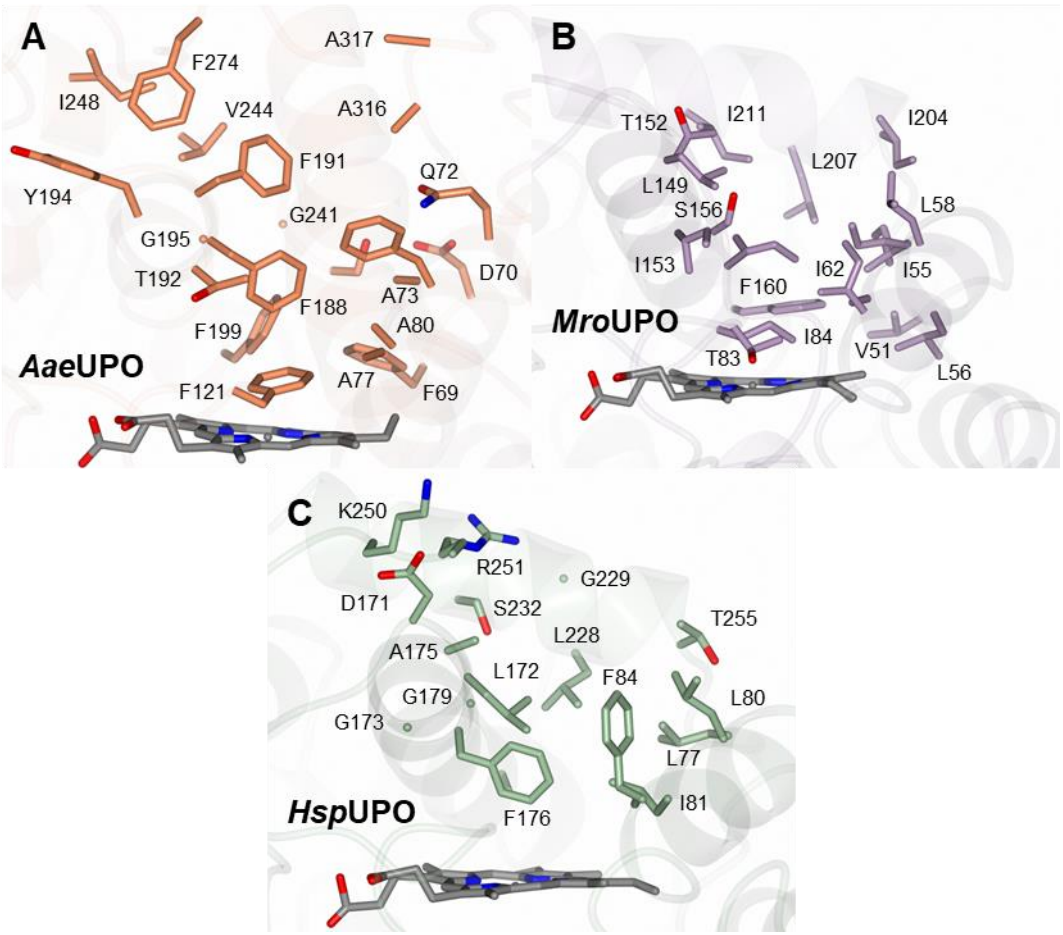


Figure 7.19 Comparison of the heme substrate channels in **A** *Aae*UPO (2YOR in orange) **B** *Mro*UPO (5FUJ in purple) and **C** *Hsp*UPO (7O1R in green)

The existing structures of UPOs have revealed important details into the mechanism and substrate specificity of these enzymes. Important differences between the structures reveal major differences between long and short UPOs, as well as differences between the short UPOs. The information provided by these structures is extremely useful to guide engineering studies by identifying important sites and residues. The few currently published UPO structures have already enhanced our understanding of these enzymes. Additional structures of UPOs would likely reveal more key sites of interest and would further our knowledge and understanding of UPOs.

7.3.3. The expression of UPOs is a potential limiting factor for industrial use

In 2004, *Aae*UPO was expressed homologously with a yield of 0.007 mg L^{-1} in a stirred-tank bioreactor in a 10 day fermentation. [77] For industrial applications, higher expression levels would be needed and so further studies looked towards using a heterologous expression. *Escherichia coli* and *Saccharomyces cerevisiae* are usually popular hosts for heterologous expression due to their ease of use, they are well studied and evolution can be relatively easily carried out. [94] In previous studies, *E. coli* has been attempted to be used as a host organism to express the long UPOs *Aae*UPO and *Abi*UPO (from the fungus *Agaricus bisporus*) but the expressed enzymes were not active. [95]

In 2014, Molina-Espeja *et al*. first reported the use of directed evolution to express soluble, stable and active *Aae*UPO in *S. cerevisiae*. [96] Nine point mutations were introduced to the native protein using their evolution technique Mutagenic Organised Recombination Process by Homologous *in vivo* Grouping (MORPHING) in *S. cerevisiae*, four of the mutations were located in the signalling peptide. [97] This resulted in an increase to 8 mg L^{-1} for the new mutant (PaDa-I) with an activity of 1300 U L^{-1} towards the compound 1,2-(methylenedioxy-4-nitrobenzene) (NBD). The mutations in the signalling peptide (F12Y, A14V, R15G, A21D) were found in the hydrophobic core of the leader and increased secretion levels by 27-fold. The five mutations in the enzyme itself (V57A, L67F, V75I, I248V, F311L) preserved the charge and polarity of the original residues in the wild type but the size of the residues were changed. This resulted in improved hydrophobic interactions within the UPO leading to tighter folding

and increased stability as well as increased secretion levels. *S. cerevisiae* tends to hyperglycosylate proteins and so the level of glycosylation increased from 22% in the wild type to 30% in PaDa-I, but this did not affect the stability of the enzyme. The kinetic properties and spectrophotometric features were very similar for both native and mutant UPOs. The same group later established a tandem yeast expression system for UPOs, where potential directed evolution could be carried out in *S. cerevisiae* as described previously and large scale production would use a different yeast host, *Pichia pastoris* (now known as *Komagataella phaffii*). [98] Expression of PaDa-I in *P. pastoris* was controlled by an *alcohol oxidase* I promoter and was induced by the addition of methanol. In a 7 L fermenter, PaDa-I was secreted with levels as high as 217 mg L^{-1} whilst also retaining the characteristics of the recombinant enzyme expressed in *S. cerevisiae* such as catalytic constants, thermal stability and pH profiles.

Since this work, most studies with *Aae*UPO have used the tandem-yeast expression system as evolution can be carried out easily and the system reliably generates high yield of protein due to the ability of *P. pastoris* to grow with high cell densities in fermentations. This system has been applied to modify different UPOs to improve activity and selectivity for different reactions, and simplify the purification of the protein. [99, 100, 101, 102] *Aae*UPO was also recently expressed in the filamentous fungus *Magnaporthe oryzae* for potential application in industry. [103] This work attempted to demonstrate the ability of *M. oryzae* to carry out post-translational modifications such as *N*-glycosylation and disulfide bond formation with eukaryotic glycoproteins. *M. oryzae* may be a preferable system to yeast expression as it avoids hyperglycosylation that is common in yeast systems. However as hyperglycosylation did not appear to

affect the activity or stability of UPOs, this is unlikely to be of much consequence in this case. [103]

To access other UPOs, most studies expressed UPOs using either the tandem-yeast expression system or use of their native hosts such as *Coprinellus radians* (*Cra*UPO), [104] *Marasmius wettsteinii* (*Mwe*UPO), [105] and *Chaetomium globosum* (*Cgl*UPO). [106] Another heterologous system using a eurkaryotic host was implemented to express the UPO derived from *Coprinius cinerea* (*Cci*UPO) using a Novozymes *Aspergillus oryzae*. [107] No expression in prokaryotic hosts such as *E. coli* was observed until 2019 when the first UPO and its variants were expressed in *E. coli* from *Marasmius rotula* (r*Mro*UPO). [108, 109] These UPOs were expressed in a BL21 strain with expression controlled by a T7*lac* promoter. The mutations that were carried out in the study changed whether the UPO would carry out higher or lower ratios of epoxidation to hydroxylation reactions of oleic acid. Shortly after two other short UPOs from *Collariella virescens* (*Cvi*UPO) and *Daldinia caldariorum* (*Dca*UPO) were successfully expressed and purified from *E. coli* with yields of 1.0 mg L^{-1} and 0.4 mg L^{-1} respectively. [110] Both UPOs were shown to be active following the oxidation of 2,2-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), **113** and veratryl alcohol (VA), **120**.

7.3.4. UPOs are amenable to protein engineering

As mentioned previously, UPOs have two main activities, peroxidative activity and peroxygenative activity but latter tends to be of more interest because it can carry out the oxidation of inert C-H bonds. The peroxidative activity is usually a competing reaction with the peroxygenative activity and so ideally the ratio of these reactions

should be changed to favour desired peroxygenase activity. Evidence for optimising peroxygenase over peroxidase activity through evolution was first seen in 2016 by Molina-Espeja *et al.* [101] Using PaDa-I as a starting point, mutant libraries were created through random mutagenesis, staggered extension process (StEP) recombination and *in vivo* shuffling to optimise the regiospecific hydroxylation of naphthalene to 1-naphthol. UV/vis assays with Fast Red reagent and 2,6-dimethoxyphenol (DMP) were used to detect and measure peroxygenase and peroxidase activity respectively. The variant that resulted from these studies, JaWa (G241D-R257K), was able to produce 156% more 1-naphthol than PaDa-I, showed twice the amount of peroxygenase activity and half the peroxidase activity compared to PaDa-I. Computational analysis using protein energy landscape exploration (PELE) and quantum mechanics/molecular mechanics (QM/MM) calculations indicated the two mutations affected the position and stability of one of the loops and the α-helix that contained the arginine-glutamine pair which in turn affected the orientations naphthalene could take, and affected the accessibility to naphthalene to the heme.

The same group then targeted the flexible loops of UPOs for mutagenesis and implemented MORPHING in order to improve the peroxygenase to peroxidase activity ratio. [111] Initially mutants with higher ratios suffered poorer thermal stability but the most thermostable mutants were then chosen for further evolution using combinatorial saturation mutagenesis (CSM). Two variants were generated from this study, QuiGon and WinDu which both showed significantly lower peroxidative activity than their parent PaDa-I by as much as 74% through ABTS and NBD assays. This is potentially due to the reduction of long-ranged electron transfer (LRET) reactions as potential

residues, such as tryptophan or tyrosine, that may have been involved in LRET pathways to the heme were mutated and affected the peroxidase activity.

MORPHING and CSM were used together in another study involving the regioselective hydroxylation of propranolol, **36** to its 5-hydroxy derivative, **37** (Figure 7.20). [102] The variant SoLo was produced and was able to generate higher turnover numbers (TTNs) than native *Aae*UPO, PaDa-I or JaWa for the hydroxylation with and without the radical scavenger ascorbic acid. The *Aae*UPO mutants were also tested on their abilities to form human drug metabolites of the drugs dextromethorphan, **38**, naproxen, **40** and tolbutamide, **42** , demonstrating the different capabilities of the mutants PaDa-I, JaWa and SoLo (Figure 7.21). [112]

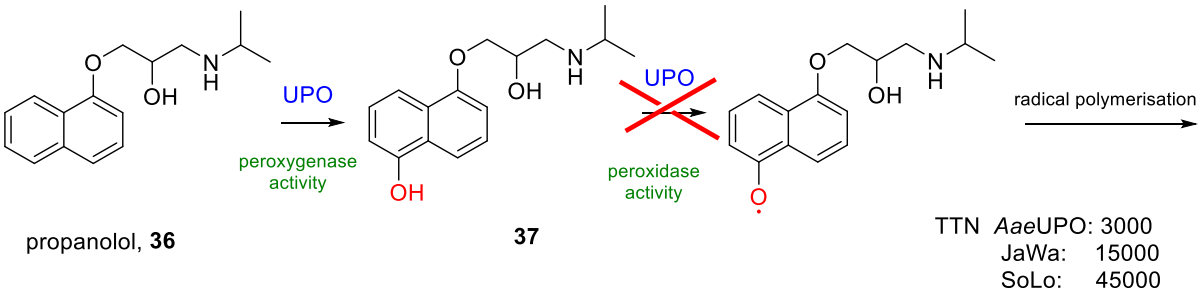


Figure 7.20 Regioselective hydroxylation of propranolol into 5'-hydroxypropanolol by *Aae*UPO and evolved mutants that reduce the peroxidase activity and hence reduce unwanted radical polmerisation[102]

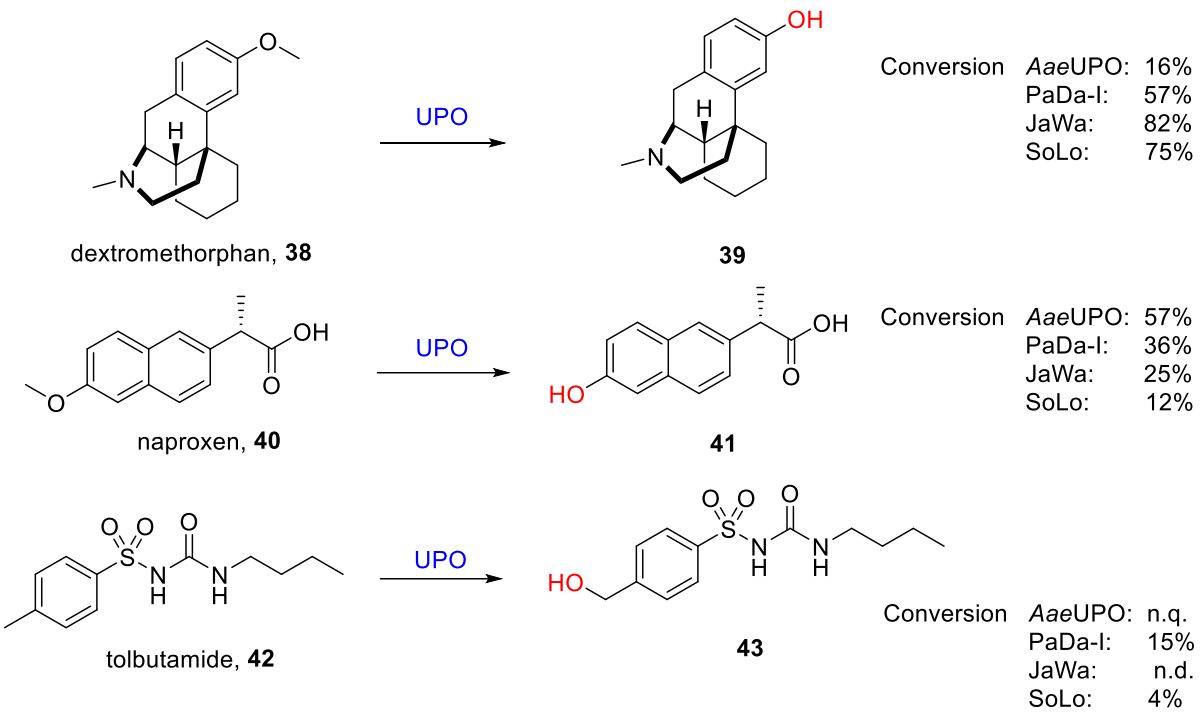


Figure 7.21 Formation of human drug metabolites through various oxidations by *Aae*UPO and evolved variants[112]

The demonstrated ability to engineer UPOs towards different substrates and reaction types is key to their use as industrial biocatalysts. Additional structural characterisation may aid in enzyme engineering efforts by allowing for identification of key sites and residues for mutation, as well as providing additional models for computational analysis.

7.3.5. Reaction capabilites of UPOs

UPOs are able to carry out a variety of oxidation reactions which will be discussed in this section. Some of these reactions are analogous to the oxidations carried out by P450s which require the use of cofactors. UPOs offer a cofactor independent alternative to P450s. making them potentially cheaper and more efficient industrial biocatalysts.

7.3.5.1. Halogenations

The first study of *Aae*UPO by Ullrich and Hofrichter showed that *Aae*UPO was capable of brominating and chlorinating monochlorodimedone (MCD), **44** where it demonstrated high brominating activity (Figure 7.22). [77] *Aae*UPO can also brominate **14** in the presence of potassium bromide to yield 2-bromo, **46** and 4-bromophenol, **47** in a ratio of 1:4 (Figure 7.22). [113] However, the ability to chlorinate phenol in the presence of potassium chloride was much lower and only traces of 2-chlorophenol were detected. In the halogenation mechanism, compound I is able to oxidise halide ions into their respective hypohalous acid, which can then oxidise the desired substrate. [114]

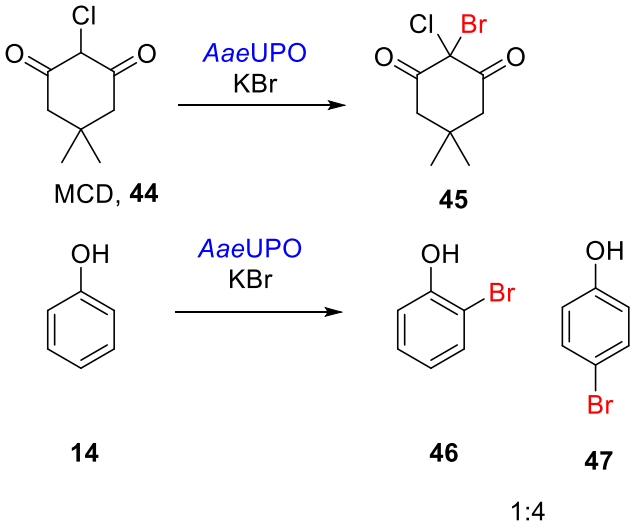


Figure 7.22 Bromination of MCD, **44** and phenol, **14** using *Aae*UPO and potassium bromide

The UPO derived from *Coprinus radians* (*Cra*UPO) is also able to brominate phenol in the same ratio as *Aae*UPO and only traces of the chlorinated product were found. [115] No other UPOs as of yet have shown any halogenating activity.

7.3.5.2. Aromatic hydroxylations

One of the first reactions that were demonstrated for UPOs was aromatic hydroxylations. In 2004, napthalene was converted by *Aae*UPO into 1-naphthol and 2-naphthol in a ratio of 36:1. This ratio demonstrates *Aae*UPO's ability to hydroxylate regioselectively. [79] Toluene was also hydroxylated, but gave rise to multiple products including benzyl alcohol, benzaldehyde, benzoic acid, *o*- and *p*-cresol, and methyl-*p*-benzoquinone. The mechanism of aromatic hydroxylations catalysed by UPOs involve the formation of an epoxide intermediate which spontaneously rearranges to form the final product (Figure 7.23). [116]

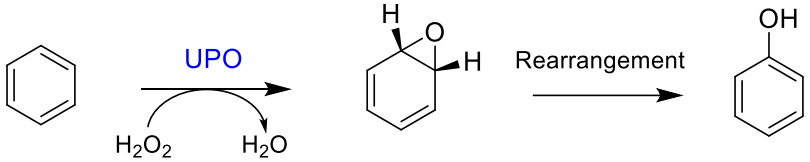


Figure 7.23 Aromatic hydroxylation catalysed by a UPO through an epoxide intermediate

UPOs can regioselectively hydroxylate more complex molecules such as flavonoids, where *Aae*UPO can predominantly hydroxylate the C6 position (Figure 7.24). [117] The use of industrially focused substrates for the synthesis of agrochemicals and pharmaceuticals have also been reported, such as (*R*)-2-phenoxypropionic acid, **50**, diclofenac, **52**, antiplatelet metabolites of clopidogrel, **54** and acetanilide, **56** to form paracetamol, **57** (Figure 7.24). [117-121] In all these examples, hydroxylation was regioselective. In the case of Figure 7.24B, a racemic mixture of **50** was added and both enantiomers of the product were formed. However, (*R*)-**50** appeared to be the preferred substrate, and a 60 % *ee* of (*R*)-**51** was observed.

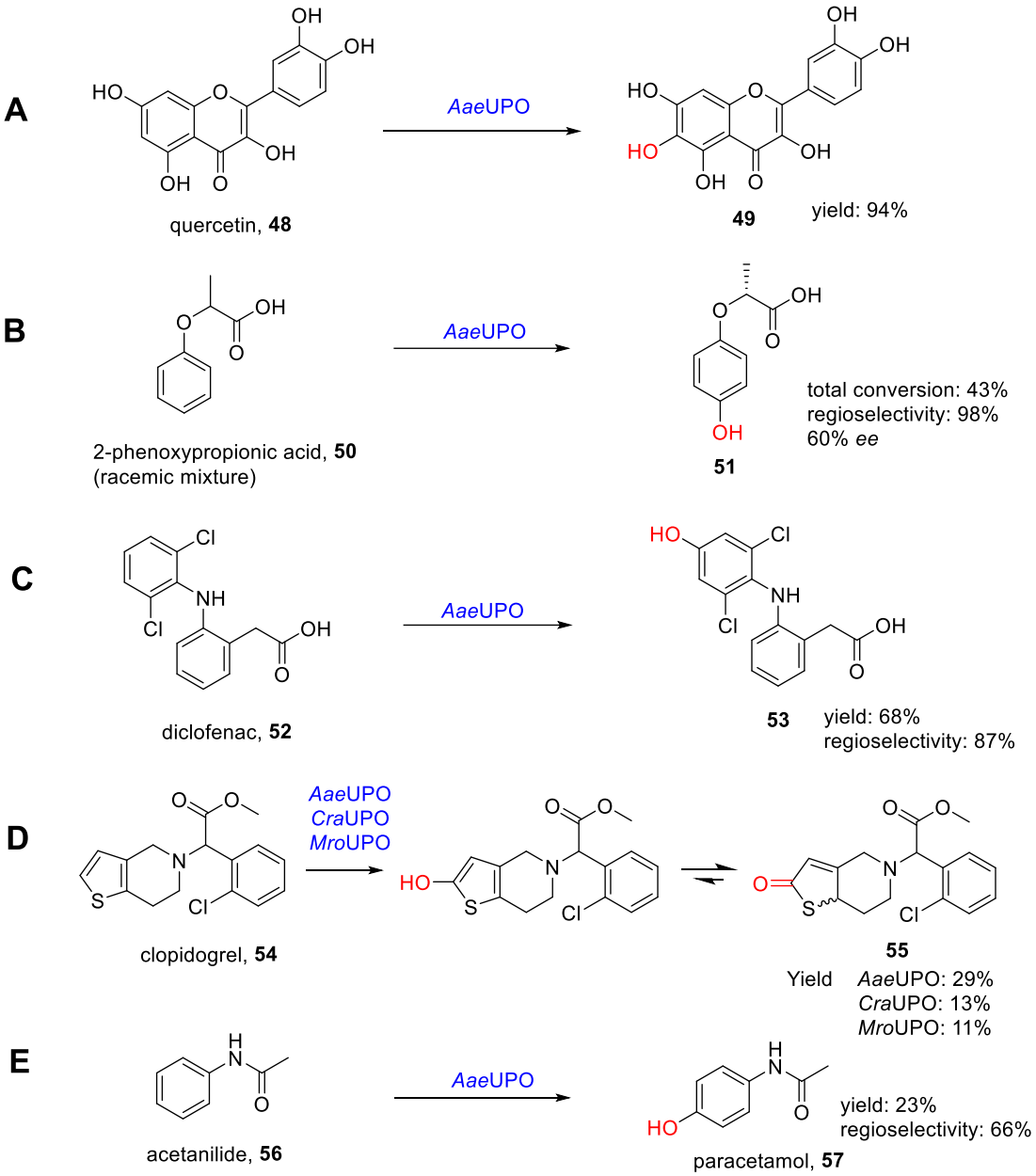


Figure 7.24 Examples of aromatic hydroxylation by UPOs to synthesise pharmaceuticals and agrochemicals: **A** hydroxylation of quercetin, **48** at the C6 position. [117] **B** para position is hydroxylated to form a precursor, **51** to aryloxyphenoxypropionic acid-type herbicides. [118] **C** hydroxylation to form human drug metabolite 4'-hydroxydiclofenac, **53**. [119,120] **D** formation of antiplatelet metabolite 2-oxo clopogrel, **55**. [121] **E** Hydroxylation of acetanilide, **56** to form paracetamol, **57**. [120]

Aromatic hydroxylations by UPOs have also been used to oxidise a variety of EPA (Environment Protection Agency) organic pollutants. [122] A comparison was run between *Aae*UPO and *Mro*UPO to test their abilities to oxidise compounds such as chlorinated benzenes, halogenated biphenyl ethers, nitroaromatics, polycyclic aromatic hydrocarbons (PAHs) and phthalic acid derivatives. *Aae*UPO was able to oxidise most of the pollutant screen but in some instances, *Mro*UPO was able to hydroxylate some compounds to a higher yield. The difference in activity between the two UPOs highlight the need to study the reaction capabilities of more UPOs in order to build an effective enzymatic toolbox. Some examples of this study is shown in Figure 7.25 which include the conversion of chlorinated benzene **58**, and PAHs **59** and **60**.

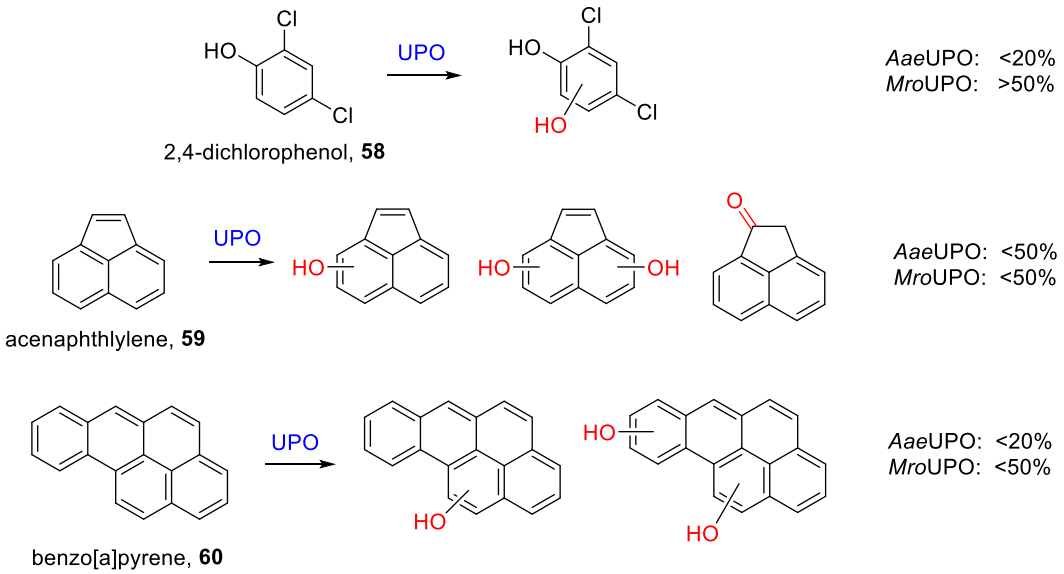


Figure 7.25 Oxidation of EPA pollutants by *Aae*UPO and *Mro*UPO. Relative conversions are shown in terms of either poor conversion, decrease of substrate by <20%, moderate conversion, decrease of substrate by <50% or good conversion, decrease of substrate by >50% [122]

Aromatic hydroxylations can be challenging because the phenolic products generated can become substrates for the peroxidative activity of the enzyme, leading to the

formation of phenoxy radicals which either couple or polymerise. To prevent the production of the peroxidation products, ascorbic acid can be added as a radical scavenger. [123]

7.3.5.3. Aliphatic hydroxylations

UPOs have been shown to catalyse the hydroxylation of linear, branched and cyclic aliphatic compounds. The regioselectivity is dependent on the substrate, however high enantioselectivities have been reported. The ability of UPOs to enantioselectively functionalise aliphatic compounds, which can be very chemically challenging and often requires the use of extreme conditions, makes them potentially very desirable for the synthesis challenging previously inaccessible targets.

*Aae*UPO has demonstrated its ability to hydroxylate linear alkanes up to 16 carbon atoms long, yielding in the respective alcohols in either the 2- or 3-position. [124] Interestingly, *n*-hexane yielded both (*R*)- and (*S*)-enantiomers of the alcohol products, but *n*-butane and *n*-pentane generated the (*S*)-form of the alcohol, and *n*-heptane and *n*-octane gave the (*R*)-enantiomers with *ee*s as high as 99%. This indicates the importance of chain length in selectively hydroxylating linear compounds. In the same study, branched and cyclic alkanes were monohydroxylated by *Aae*UPO but branched compounds were regioselectively hydroxylated at tertiary carbon atoms (Figure 7.26).

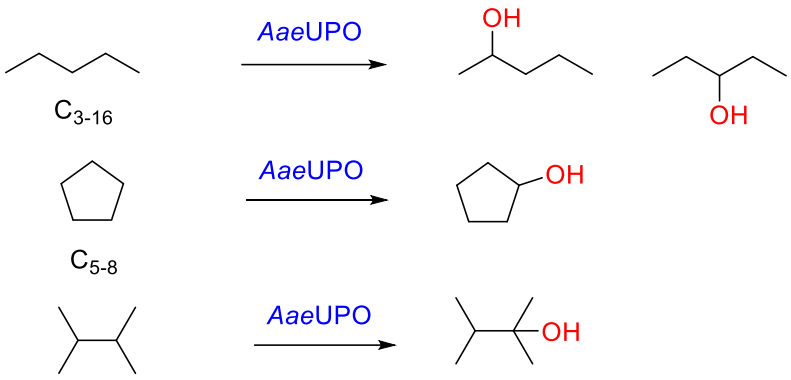


Figure 7.26 Examples of aliphatic hydroxylation of linear, cyclic and branched alkanes carried out by *Aae*UPO[124]

Four different UPOs, *Aae*UPO, *Mro*UPO, r*Cci*UPO and rNOVO were investigated for the conversion of cyclohexane into the products cyclohexanol and cyclohexanone. Catalytic efficiencies for the first oxidation step for all of the UPOs were very similar but differences arose in the second oxidation to form the ketone. The catalytic efficiencies of *Mro*UPO and rNOVO for the second oxidation were 1-2 orders of magnitude higher than the values observed for the other two UPOs. [125] This demonstrates that different UPOs can be used to selectively generate primary or secondary oxidation products.

Both *Aae*UPO and r*Cci*UPO are able to monohydroxylate at the 2 and 3 positions or dihydroxylate at the 2 or 3 position from both ends of the linear alkane tetradecane **61** (C14), as well as form the keto derivatives (Figure 7.27). [107, 126] The conversion of *n*-tetradecane **61** was 51% using r*Cci*UPO which was higher than the conversion for *Aae*UPO with a conversion of 24%. Interestingly, for both UPOs as the percentage of acetone (*v*/*v*) in the reaction was increased from 20% up to 60%, the conversion of **61** increased and the monohydroxylated compounds became the major product (only product found in 60% acetone reaction). This increase in conversion and

monohydroxylated product is believed to be due to increased solubility of *n*-tetradecane with an increasing percentage of acetone and a decreased solubility of the product. Small amounts of other oxidation compounds, including hydroxy/keto-, carboxy- and dicarboxy-acid derivatives were also detected. *Mro*UPO is also able to convert **61** with 45% conversion to give rise to the primary products which were the 2- and 3- monohydroxylated derivative. [93] Only *Mro*UPO was able to oxygenate the terminal ends of *n*-dodecane, **62** resulting in the formation of 1-dodecanol, **63**, dodecanoic acid, **64**, ω-hydroxydodecanoic acid, **65** and 1,12-didodecandioic acid, **66**.

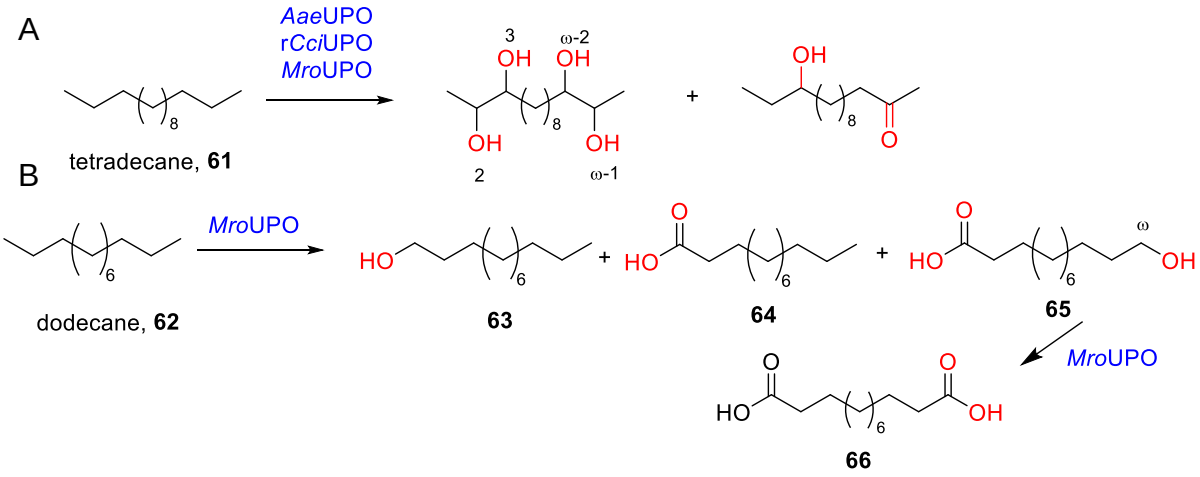


Figure 7.27 Oxidation of alkanes by UPOs **A** *Aae*UPO, r*Cci*UPO and *Mro*UPO oxidation of *n*-tetradecane **61** generated a mixture of mono- and di-dihydroxy/keto products **B** *Mro*UPO was able to oxygenate in the terminal positions of dodecane **62**[93, 107, 126]

In alkyl- and cycloalkylbenzenes, AaeUPO was able to hydroxylate the benzylic carbon to form secondary alcohols with very high stereospecificities of over 99 % *ee* as shown with the conversion of ethylbenzene **67a** and 1,2,3,4-tetrahydronapthalene **68** in Figure 7.28. [127] Again, the chain length affected the reactivity as longer alkyl chains decreased the enantiomeric purity and the number of side products increased.

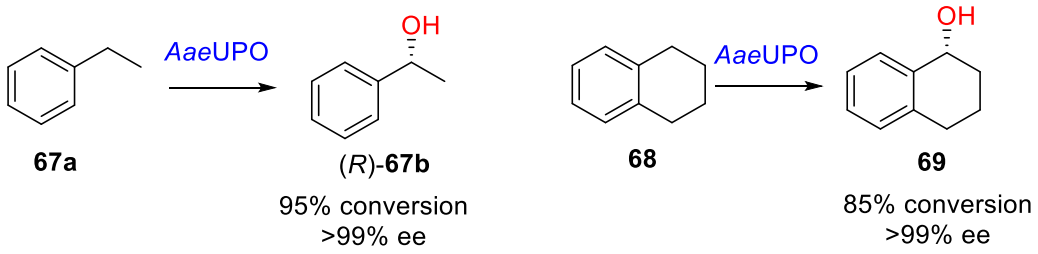


Figure 7.28 selective hydroxylation of the benzylic carbon by *Aae*UPO

Fatty acids have also been reported as a suitable substrate for UPOs to carry out hydroxylations. *Aae*UPO has been shown to monohydroxylate a range of fatty acids in both the ω-1 and ω-2 positions, as well as form some of the keto derivatives, however, conversions were poor to moderate. [126] The reactivity of the fatty acids towards *Aae*UPO decreased as the length of the chain increased, and monounsaturated fatty acids were more reactive than their respective saturated analogues. r*Cci*UPO, like *Aae*UPO, can carry out regioselective monohydroxylations and form some of the keto derivatives predominantly in the ω-1 and ω-2 positions of saturated/unsaturated fatty acids and fatty acid esters. [107] The ratio of these products depended on the length of the chain where the longer the alkyl chain, the more ω-2 hydroxylation product was formed. r*Cci*UPO consistently gave rise to higher conversions of oxidation products compared to *Aae*UPO, indicating a preference of r*Cci*UPO for long aliphatic chains.

*Mro*UPO is also able to hydroxylate fatty acids when ten times the amount of H_2O_2 was added but the regioselectivity depended on the fatty acid itself. [128] With tetradecanoic acid, *Mro*UPO catalysed the oxygenation to form hydroxy or keto derivatives at the terminal or subterminal (ω-2) position. However when tetradecandoic acid was used as a substrate, hydroxylation occurs twice at the α position which causes a chain shortening reaction. This mechanism is shown in Figure

7.29. This reaction was also observed with decanoic acid but no reaction with any of the substrates was observed with *Aae*UPO.

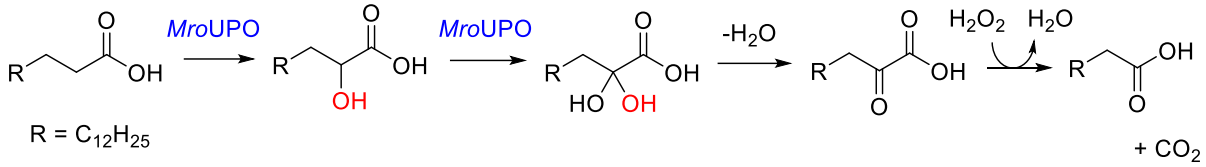


Figure 7.29 Mechanism for fatty acid chain shortening using *Mro*UPO as a catalyst *via* two α-hydroxylations

As well as chain shortening of fatty acids, *Mro*UPO is also capable of another unusual reaction. Both *Mro*UPO and *Mwe*UPO are able to selectively remove the hydroxyacetyl side chain from C17 of corticosteroids such as cortisone, **70** *via* two oxygenation steps (Figure 7.30). [105] *Aae*UPO was unable to convert these substrates.

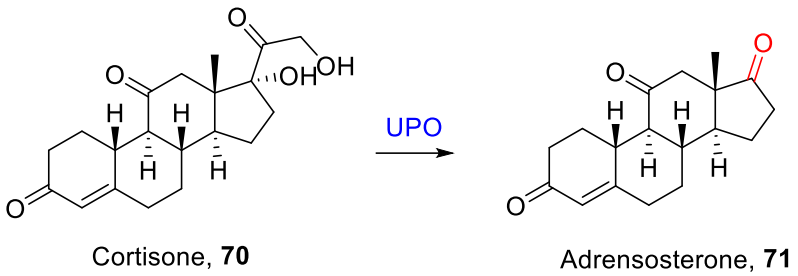


Figure 7.30 Side chain removal from cortisone, **70** catalysed by UPOs

Through analysis of the crystal structure of *Mro*UPO and docking studies, Kiebst *et al*. propose that the difference in substrate acceptance between *Aae*UPO and the *Marasmius* UPOs is due to two reasons. Firstly the dimensions of the substrate channel found in *Mro*UPO was 4 Å shorter and 5 Å wider than *Aae*UPO's channel, which would be able to accommodate larger, bulkier molecules. Secondly, the nature of the channel in *Mro*UPO was more likely to accept bulky costeroids due to the flexible, hydrophobic residues found in the tunnel compared to the more rigid aromatic

residues found in the channel of *Aae*UPO. As *Mwe*UPO has 90% sequence identity to *Mro*UPO, the proposed conclusions can be applied to this enzyme as well.

Other biological molecules such as steroid derivatives cholesterol, **72**, cholesta-3,5-dien-7-one, **74** and vitamin D **76** are also suitable substrates for aliphatic hydroxylation, where the C25 hydroxy-derivative in these cases are preferentially formed by *Aae*UPO, r*Cci*UPO and *Mro*UPO (Figure 7.31). [129,130, 131]

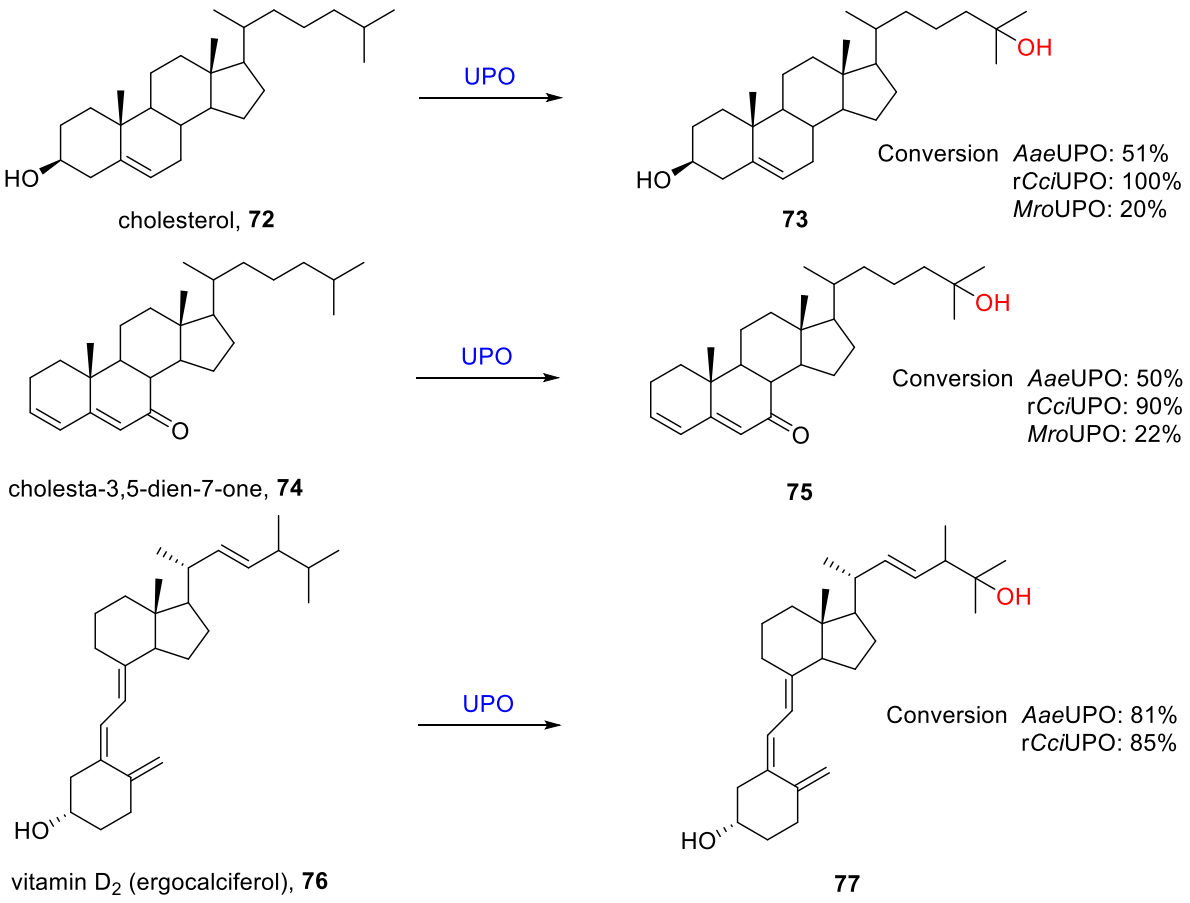


Figure 7.31 Regioselective hydroxylation of various steroid-like molecules at the C25 position carried out by UPOs

7.3.5.4. Epoxidations

UPOs are able to catalyse the epoxidation of linear, branched and cyclic alkenes, however selectivity issues tend to arise depending on the type of alkene (Figure 7.32). In the case of *Aae*UPO, long and linear allylic alkenes with a terminal double bond often produce both epoxidised and hydroxylated products where the hydroxylation occurs on the methylene group next to the double bond. [132] However, branched and benzylic alkenes are catalysed with excellent conversion and regioselectivity. [127,132]

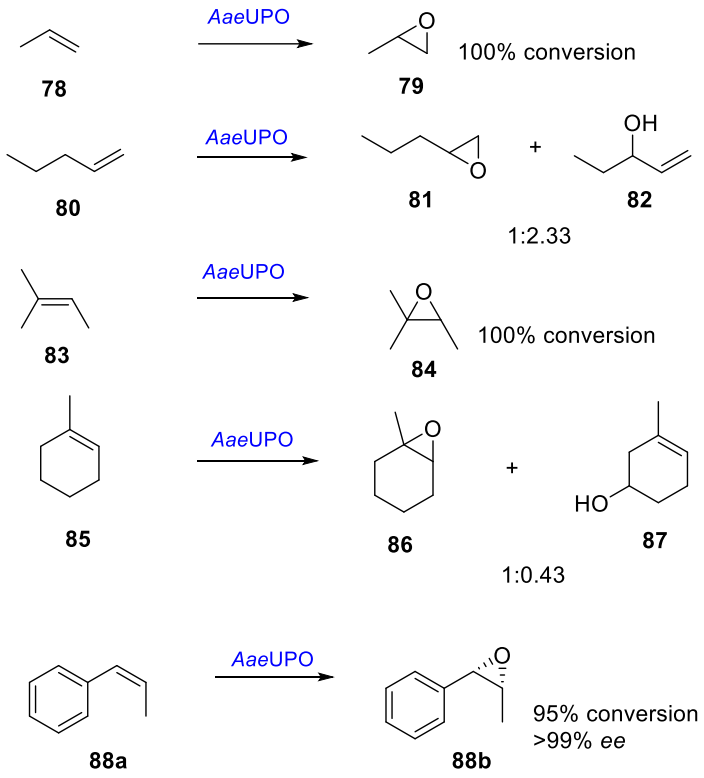
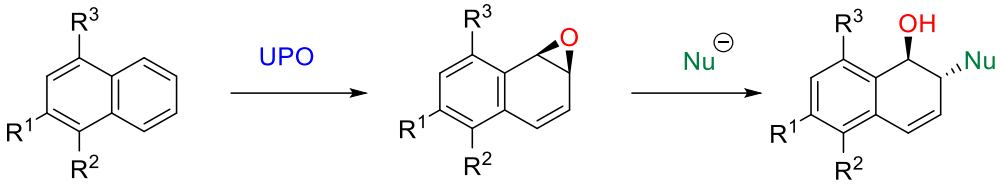


Figure 7.32 Examples of epoxidation of double bonds carried out by *Aae*UPO[127,132]

Aromatic hydroxylations catalysed by UPOs occur through an epoxide intermediate. [116] Zhang *et al*. 2021 exploited this intermediate to synthesise chiral, trans-disubstituted cyclodienes using naphthalene derivatives as starting materials and a nucleophile to attack the naphthalene epoxide. [133] Two evolved variants of *Aae*UPO, Pada-I and SoLo, were used in this study (Figure 7.33).



R^1	R^2	R^3	Nucleophile	Conversion/ %	
				PaDa-I	SoLo
H	H	H	N_3	60	n.d
H	Br	H	N_3	73	n.d.
H	CF_3	H	N_3	45	n.d.
H	CH_3	H	N_3	19	37
CH_3	H	H	N_3	51	n.d.
H	Cl	Br	N_3	31	n.d.
H	Cl	H	N_3	49	n.d.
H	F	H	N_3	26	n.d.
Br	H	H	N_3	45	n.d.
H	Cl	H	COOH	15	n.d.
H	Br	H	COOH	6	n.d.

Figure 7.33 Formation of trans-disubstituted cyclodienes from naphthalene epoxides *via* UPO catalysed epoxidation. [133] n.d. = not determined

A variety of conversions were observed using azide as a nucleophile. The use of formate as a nucleophile resulted in poorer yield, likely due to the fact that formate is a poorer nucleophile than azide.

*Mro*UPO and *Cgl*UPO were able to epoxidise monounsaturated fatty acids, **89** with the epoxide compounds as the main product (Figure 7.34). [134] *Cgl*UPO was highly

chemoselective as it could convert over 90% of the fatty acids tested into epoxides, whereas the level of epoxidation for *Mro*UPO varied. Other side products included the oxygenated (hydroxy, keto or carboxy) derivatives of the epoxide product or the fatty acid at the terminal, subterminal positions or at the allylic carbons. Diunsaturated fatty acids, **91** were also tested. Both *Cgl*UPO and *Mro*UPO were capable of diepoxidations, generating both anti and syn-enatiomers, **92**, but lower enzyme concentrations resulted in favouring the monoepoxidation reaction. In a later study, *Cvi*UPO was expressed in *E. coli* and could also carry out fatty acid epoxidation but only of monounsaturated compounds. [135] Because of the ease of expressing in *E. coli* rather than the native host, mutagenesis could be carried with the aim of widening the substrate channel. To achieve this, the residues in the substrate channel (identified by computational modelling based on the *Aae*UPO structure) were changed to smaller, less bulky residues using the residues in the *Cgl*UPO substrate channel as a guide. A mutant was found where a phenylalanine residue was mutated to a lysine (F88L) to open the channel which decreased the level of side reactions whilst maintaining high epoxidation activity. This indicates the potential of structure guided enzyme engineering to aid improving UPOs towards industrial applications. Another recombinant UPO from NovoZymes (r*Hin*UPO) was capable of forming the triepoxide product, **94** of α-linolenic acid, **93** and its fatty ester derivative with yields as high as 52% (Figure 7.34). [136] Interestingly, *Aae*UPO and r*Cci*UPO were unable to carry out epoxidations on any of the fatty acids and only formed hydroxylated derivatives at the subterminal positions.

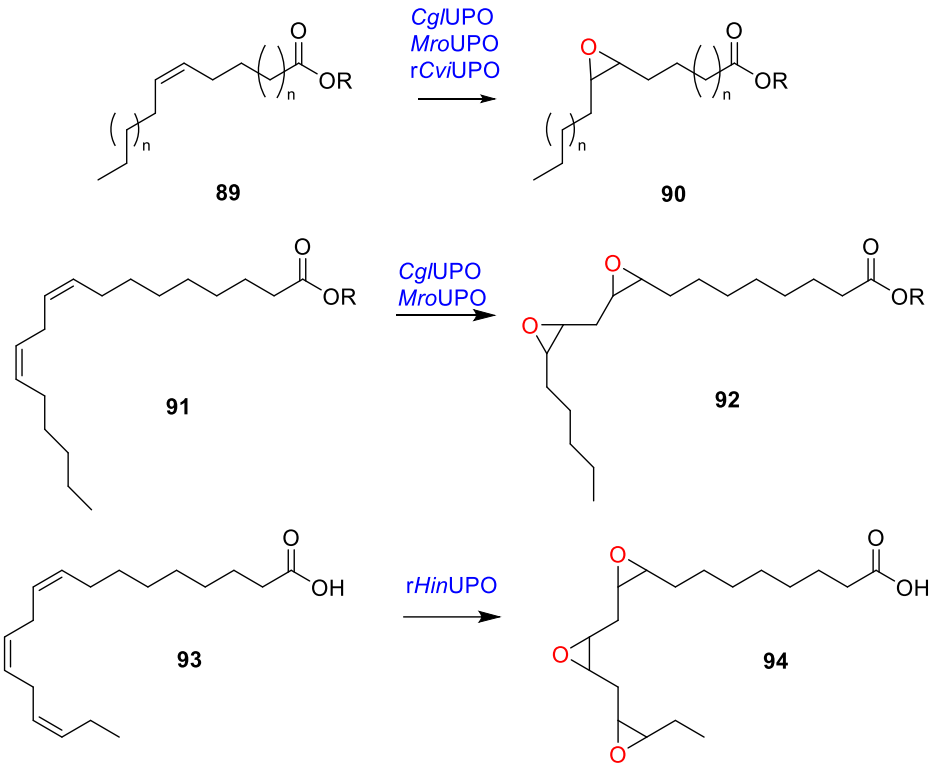


Figure 7.34 Epoxidation of fatty acids and their epoxide derivatives carried out by *Cgl*UPO *Mro*UPO, r*Cvi*UPO and r*Hin*UPO[134, 135, 136]

7.3.5.5. Sulfoxidations

Aranda *et al.* 2009 investigated the ability of *Aae*UPO and *Cra*UPO to convert dibenzothiophene into its sulfoxidation products. [137] In fungal culture experiments, *Aae*UPO was able to completely convert dibenzothiophene, **95** within 10 days, yielding in the sulfone, **97** as the major metabolite (42%), the sulfoxide, **96** (15%) and traces of hydroxylated products, **98** (Figure 7.35). *Cra*UPO was only able to convert 59% of **95** after 16 days but only sulfoxidation products were formed. Howevever, *Aae*UPO *in vitro* produced a variety of polyhydroxylated metabolites and only traces of the sulfoxidation products were observed in the presence of ascorbic acid. In contrast, *Cra*UPO produced mostly sulfoxidation products where the sulfoxide **96** was the major product with or with ascorbic acid, only traces of hydroxylated products **98** were found.

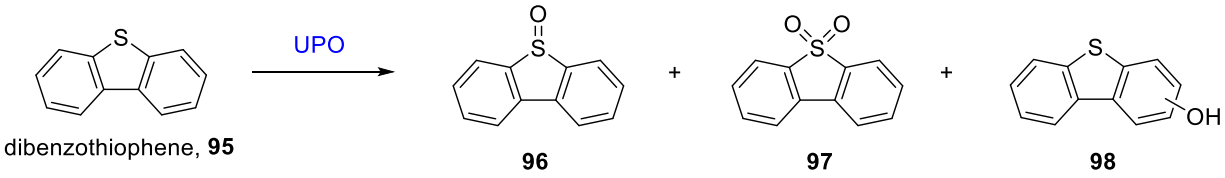
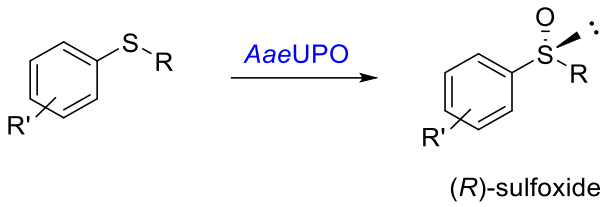


Figure 7.35 Conversion of dibenzothiophene **95** by *Aae*UPO into the respective sulfoxide **96**, sulfone **97** and hydroxylated products **98**

A more recent study focused on the enantioselective sulfoxidation of thioanisole derivates using *Aae*UPO (Figure 7.36). [138] In all of the substrates tested, the (*R*)-sulfoxide was the major product with no traces of the over-oxidation sulfone product or other by-products. High conversions and enantioselectivities were reported for compounds with electron donating substituents on the para position of the ring, and with small and relatively inflexible groups on the alkyl side of the compound.



R	R'	Conversion / %	*ee* / %
CH_3	H	>99	80
CH_3	*o*-Cl	35	74
CH_3	*p*-Cl	>99	90
CH_3	*p*-CH_3	83	90
CH_3	*p*-OCH_3	87	70
$CH{=}CH_2$	H	85	>99
C_2H_5	H	95	>99

Figure 7.36 Enantioselective sulfoxidation of thioanisole derivatives carried out by AaeUPO[138]

7.3.5.6. Heteroatom dealkylations

In 2009 *Aae*UPO showed its ability to cleave ethers in a variety of compounds to produce carbonyl and alcohol derivatives with limiting H_2O_2 *via* a hemiacetal intermediate. [139] When a non-limiting amount of H_2O_2 was added, the alcohol derivatives were further oxidised by *Aae*UPO to form the respective secondary oxidation product. Examples of ether cleavage reactions carried out by *Aae*UPO are shown in Figure 7.37. Interestingly in the example of 3,4-dimethoxybenzyl ether, the benzyl ether linkage was selectively cleaved whilst the other ethers within the compound were not cleaved by *Aae*UPO.

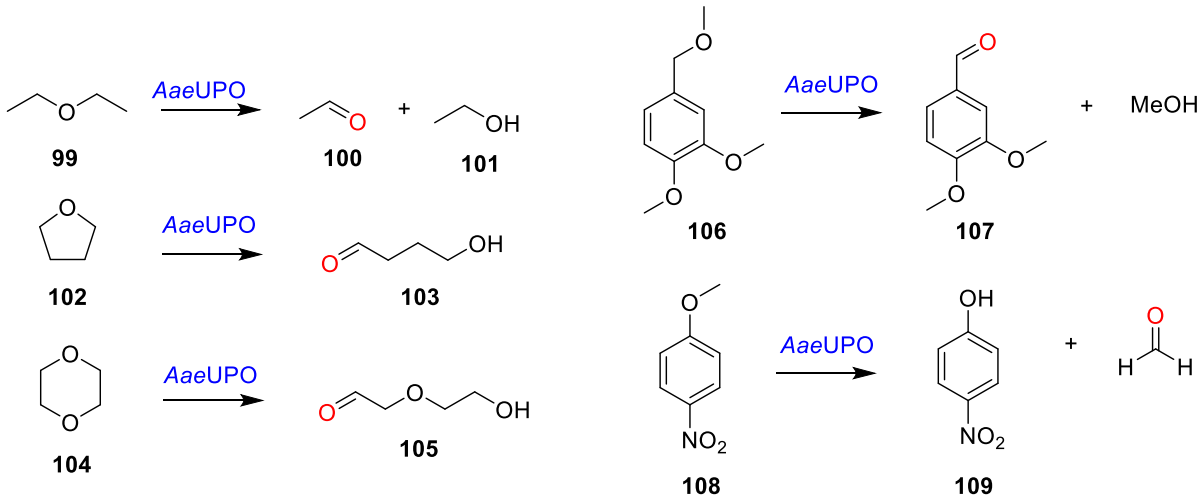


Figure 7.37 Oxidative ether cleavages carried out by *Aae*UPO

Pada-I, an evolved variant of *Aae*UPO, was also tested for ether cleavage activity with a range of aromatic ethers. [140] Aromatic ethers with hydrophobic, non-polar substituents were accepted by PaDa-I whereas more polar substituents were not accepted. An excess of H_2O_2 lead to inactivation of the UPO and so low concentrations were added slowly to the reactions to reduce heme oxidation.

In human metabolism, dealkylations are common to metabolise and excrete drugs. It is therefore ideal to be able to mimic this process to create human drug metabolites

for medicinal research. In a study screening various drugs against *Aae*UPO, *O*-demethylation and *O*-deethylation were observed in metoprolol, naproxen, **40** and phenacetin with conversions of 17 %, 57 % and 23 % respectively. [120] *Aae*UPO was also able to catalyse *N*-dealklyations of compounds with tertiary amines including sildenafil, **110** and lidocaine, **112** in the same study (Figure 7.38). The precursor to SAR548304, **114** a bileacid reabsorption inhibitor, has a tertiary amine that can be dealkylated by *Mro*UPO but *Aae*UPO was unable to accept this compound as a substrate (Figure 7.38). [125]

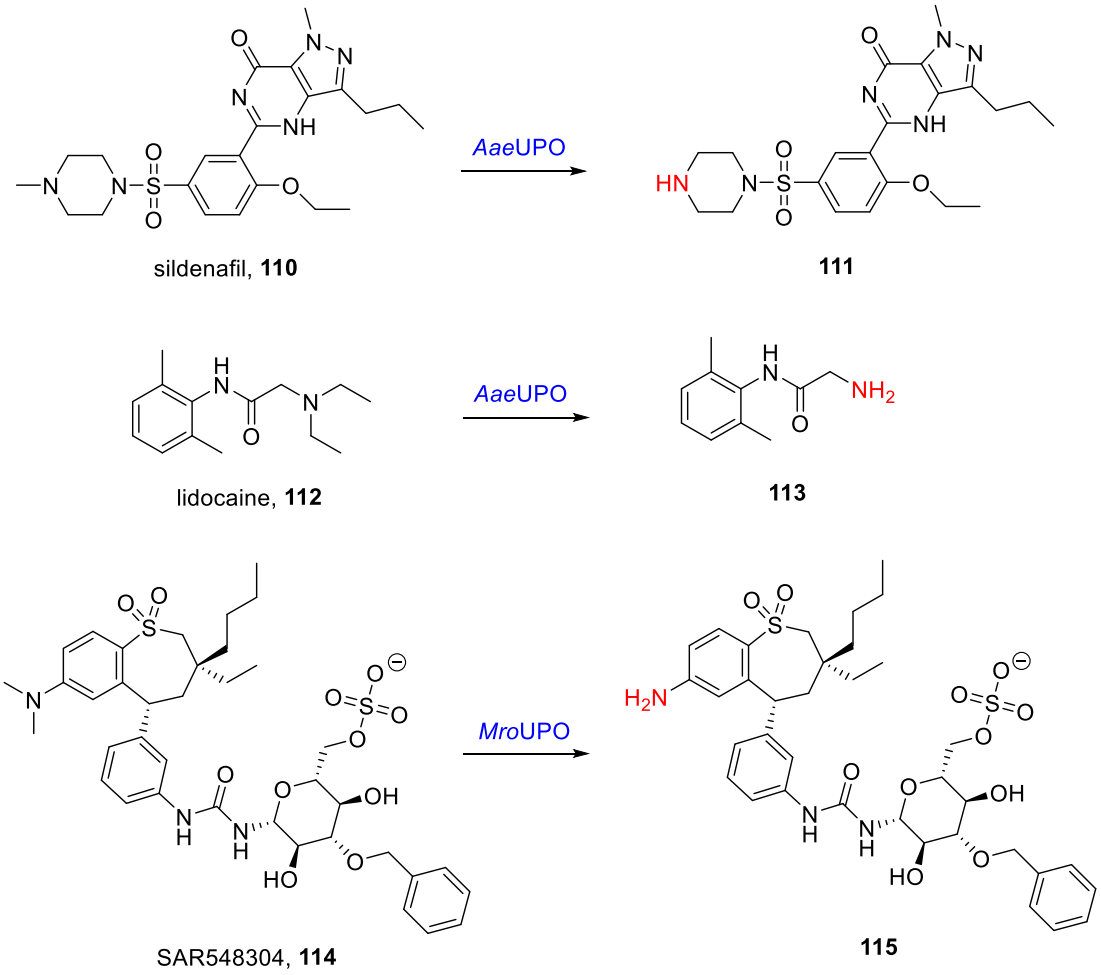


Figure 7.38 N-dealkylations to form human drug metabolites using UPOs[120, 125]

7.3.5.7. *In situ* production of hydrogen peroxide

One of the major advantages of UPOs over other oxidoreductases is the the simplicity of only needing hydrogen peroxide as a cosubstrate. However, the presence of H_2O_2 at high concentrations can also become a problem for UPOs due to the catalase activity of UPOs. The first issue is that hydrogen peroxide dismutation is catalysed by UPOs to produce molecular oxygen and two molecules of water. [80] This reaction competes with the desired peroxygenation reaction for H_2O_2 and so the peroxide becomes a limiting factor, reducing productivity of the enzyme. The second and main issue arises when an excess of H_2O_2 is added to the reaction and heme destruction occurs. This reaction has been observed in other heme-proteins such as chloroperoxidases, mangansese peroxidases and P450s. [141] In UPOs a hydroxy radical can be formed from unwanted catalase activity which in turn converts heme into biliverdin, carbon monoxide and Fe^{2+} *via* a compound III intermediate. [142] To resolve this issue, many studies involving UPOs add hydrogen peroxide slowly to the reaction to avoid heme destruction. However, the issue may be resolved by slowly producing H_2O_2 *in situ* so that the UPO can consume the peroxide without the effects of heme bleaching. This has been achieved in a variety of methods including enzyme cascades involving alcohol oxidase, formaldehyde dismutase, and choline oxidase, gold-titanium dioxide photocatalysts that can reduce either methanol or water to form hydrogen peroxide, and contact with non-thermal plasma. [143, 144, 145, 146, 147] A self-sufficient peroxygenase was recently created through the fusion of an evolved aryl-alcohol oxidase derived from *Pleurotus eryngii* and *Aae*UPO (SoLo variant) joined together with a peptide linker (Figure 7.39). [148, 149] This system was able to selectively demethylate the antitussive drug dextromethorphan, **38** to form the metabolite dextrophan, **39** with *in situ* hydrogen peroxide production through the oxidation of 4-

fluorobenzyl alcohol, **116** by aryl-alcohol oxide, proving its capabilities in commercial use.

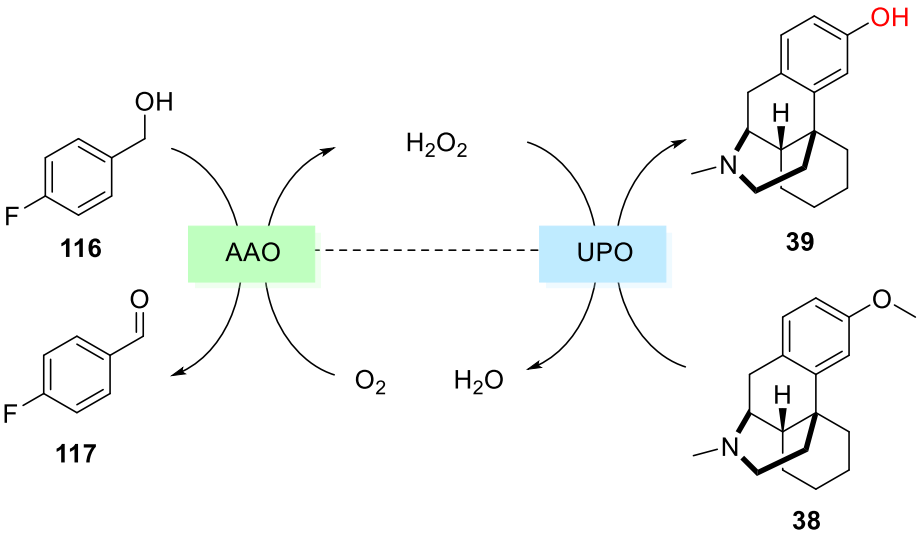


Figure 7.39 A self-sufficient UPO where aryl-alcohol oxidase and *Aae*UPO are joined together and the system was able to demethylate to form the metabolite dextrophan, **39**

7.3.5.8. UPO catalysis in solvents

For the practical use of UPOs in industry, they must be compatible with organic solvents as many organic starting materials of interest are only soluble in organic solvents and not in 100% aqueous solution. Organic solvents can have a detrimental effect to the activity of enzymes, as the lack of water can affect the rigidity and the conformation of the protein, leading to denaturation. [150, 151] Promisingly UPOs have been shown to be active in cosolvent systems as shown in some of the examples already described which most often includes the use of acetone and acetonitrile. [107, 126, 128, 133, 138, 125] In one case, immobilised r*Aae*UPO was able to carry out enantioselective hydroxylation of ethylbenzene in neat ethylbenzene starting material on gram scale, showing potential industrial application. [152]

In order to further improve the compatibility of UPOs with organic solvents, the Alcade group have used their MORPHING strategies to develop a new mutant based upon PaDa-I, WamPa, that was 23 fold more active in 30% acetonitrile than its parent and was also stable in acetone, methanol and dimethyl sulfoxide. [153] The mutations introduced were located on the surface of the protein and hypothetically introduced more rigidity and stability. The ability of UPOs to work in cosolvent systems and to be mutated towards improved stability in these systems is very promising for industrial application of these enzymes.

7.3.6. Identification of an artificial UPO

In 2016, a patent was published by NovoZymes which describes a new sequence of an artificial UPO (artUPO). [154] This sequence is derived from *Mro*UPO and has 75% sequence similarity. However, it is not clear what effect these mutations have on the reactivity, stability and substrate acceptance of artUPO compared to *Mro*UPO. Investigation of the effect of these mutations on structure and activity may provide valuable information for the evolution of this class of enzymes towards industrial biotransformations.

Carro *et al.* 2019 first demonstrated the ability of *Mro*UPO, a short UPO, to be expressed in *E. coli*. [108] This posed the question of how the structure and activity could change for a given UPO expressed in either yeast or bacteria, especially given the differences in glycosylation capabilities between the two systems. Given that artUPO was derived from *Mro*UPO it was predicted that expression of artUPO could be achieved in bacteria. A comparison of artUPO expressed in different hosts could then

be made in order to investigate the applicability of each expression system for industrial use.

7.4. Aims of the project

It is clear that extensive studies into the structure, characterisation and the reactivity of UPOs are essential to further improve their use for industrial application. Further investigation of the biocatalytic use of artUPO would be valuable in discovering the effects of the mutagenic changes carried out by NovoZymes on its catalytic capability. artUPO also serves as a useful target to investigate the practicalities of expressing a recombinant UPO in different hosts.

As discussed previously, the enzyme *Cci*UPO has only been successfully expressed in *S. cerevisiae* and has not been further explored due to this limitation despite its potential in complementary reaction scope to *Aae*UPO. Potential expression of *Cci*UPO in a different host may aid in exploration of *Cci*UPO's reaction scope. In addition, in the protein data bank there are only structures of one long UPO (*Aae*UPO) and its variants. Another structure of a long UPO such as *Cci*UPO would greatly contribute to the understanding of UPOs and would allow for specific targeting of residues for structure-guided engineering.

The aims of this project were to:

- Clone the gene of artUPO and express in *P. pastoris* and *E. coli*
- Clone the gene of *Cci*UPO and express in *P. pastoris*
- Establish purification protocols for C*ci*UPO and artUPO from *P. pastoris* and /or *E. coli*
- Determine the structure of artUPO from both *P. pastoris* and *E. coli* by x-ray crystallography to allow for comparison of structure between expression hosts

- Biochemical and kinetic characterisation of artUPO from both *P. pastoris* and *E. coli* to further compare differences in expression hosts
- Determine the structure of *Cci*UPO by x-ray crystallography in order to inform future structure guided engineering work
- Investigate the substrate and reaction scope of both enzymes to better assess their industrial potential and to help build an enzymatic toolbox of UPOs

8. Expression and Purification of artUPO in yeast and bacteria and attempted expression of *Cci*UPO

An emerging interest in the use of UPOs for industrial biotransformations has been observed due to their ability to carry out a large range of oxidation reactions using only hydrogen peroxide as a cosubstrate. [78, 88] One of the factors limiting their use in industry is the difficulty in producing the active enzymes in high enough yields. Production in native hosts are often too costly and have too low yields to be industrially applicable. [94] Industrial application of enzymes tend to rely on heterologous expression in standard host organisms to produce high yields of enzyme in a quick and efficient manner. [96] Heterologous expression systems also allow for quick generation of mutant proteins for enzyme engineering. [85] The production of UPOs in heterologous hosts is made more challenging by the need for post translational modifications, such as glycosylation and disulfide bond formation which been observed in some UPOs. To date, a few examples of recombinant production of UPOs exist, including expression in *S. cerevisiae*, *P. pastoris*, *M. oryzae*, and *E. coli*. [96,98,107,108,110]

P. pastoris has proven to be a desirable host for expression of UPOs because it has been shown to produce *Aae*UPO in high yields (217 mg L^{-1}) and was able to carry out glycosylation successfully using an optimised signalling peptide. [96, 98] For these reasons, *P. pastoris* was chosen for the heterologous expression of artUPO and *Cci*UPO in this project. It has been recently demonstrated that some UPOs can be successfully produced using *E. coli* expression systems. [108,110] The use of *E. coli* for a recombinant expression system is highly desirable due to its ability to produce high yields of protein at low costs and is easily manipulated for mutagenesis studies. [94] One

of the UPOs successfully expressed in *E. coli* was *Mro*UPO. As artUPO is derived from *Mro*UPO, it was anticipated that bacterial expression of artUPO would also be possible. Given the inability of *E. coli* to glycosylate target proteins and the differences in chaperones compared to eukaryotic systems, questions remain as to whether UPOs expressed in bacteria are just as active and stable as those expressed in eukaryotic systems. A comparison of artUPO expressed in both *E. coli* and *P. pastoris* may provide valuable information to aid in the selection of expression systems for UPOs. In this chapter, the genes for artUPO and *Cci*UPO were cloned into their appropriate expression vectors, and expressed in *P. pastoris*. The gene for artUPO was also expressed in *E. coli* and both forms of artUPO, artUPO$_{yeast}$ and artUPO$_{bact}$, were purified for further characterisation later in the project.

8.1. Materials and Methods

8.1.1. Materials

Chemicals and other molecular biology materials were purchased from Alfa Aesar, Fisher Scientific, Fluorochem, Generon, Merck Chemicals Ltd, Millipore, New England Biolabs, Qiagen, Scientific Laboratory Supplies Ltd, Takara Bio, Tokyo Chemical Industry and VWR International. Primers were synthesised and purchased from Eurofins Scientific.

8.1.2. Expression of artUPO and CciUPO in *P. pastoris*

8.1.2.1. Gene amplification

The genes for r*Cci*UPO and artUPO were ordered from Invitrogen in a 17ADGGHP plasmid and the sequences of the genes are given as given in Table 8.1.

Table 8.1 Gene sequences of the target UPOs. **Top**: artUPO. **Bottom**: *Cci*UPO

artUPO DNA sequence
TCCCAGGATATTGTGGACTTCTCCCAGCATCCGTGGAAGGCACCTGGACCCA
ACGATCTCCGCTCCCCCTGTCCCGGACTCAACACGCTCGCGAACCACGGCTT
CCTCCCTCGCAACGGTCGAAACATCACGATCCCGATGATCGTCCAGGCAGGC
TTCGACGGCTACAACGTCCAGCCGGATATCCTCATTTTGGCAGCCAAGGTGG
GATTGCTCACCTCGCCTGAACCCGATACCTTCACCCTCGACGATTTGAAGTTG
CATGGAACCATCGAGCATGATGCATCGTTGTCGAGGGAGGACTTCGCGCTCG
GCGATAACCTCCACTTCAACGAGGCGATCTTCAACACCCTCGCAAACTCGAA
CCCTGGTTCCGATGTCTACAACATCACCTCGGCAGGACAGGTGCTCAAGGAC
CGCCTCGCCGACTCGCTCGCCAGGAACCCGAACGTCACGAACACCGGCAAA
GAGTTCACTATCAGGACTTTGGAGTCCGCCTTCTATCTCTCCGTCATGGGCAA
CGCAACAACAGGTGAAGCACCCAAAAACTTCGTGCAGATTTTCTTCCGAGAG
GAGCGGCTCCCCATCGAGGAGGGCTGGAAGAGGTCCACAACCCCTATCACG
TCGGACACGTTGAACCCCATTGCAGGCCAGATTTCGGAGGCGTCGAACTGGA
AGCCCAACCCCGACCAGTGTCCCTGGATCGTCCTCTCGCCCAACTTGTAG

r*Cci*UPO DNA sequence
TTCCCGGCGTACGCGTCCCTTGGAGGTTTAACCGAGCGTCAAGTCGAAGAGT
ACACGTCCAAGCTCCCTATCGTCTTTCCACCACCGCCGCCTGAACCTATCAA
GGACCCGTGGCTCAAGTTGGTCAATGACAGGGCTCATCCATGGAGACCCCTT
CGGAGAGGAGATGTCAGAGGACCCTGCCCGGGGTTGAATACGTTGGCATCC
CATGGGTATCTTCCTCGAGATGGTGTGGCGACTCCAGCTCAAATCATCACTG
CCGTCCAAGAAGGCTTCAACATGGAGTACGGGATCGCGACATTCGTCACCTA
CGCTGCCCACCTCGTCGATGGAAACCCACTCACCAATCTCATCAGCATTGGT
GGGAAGACGCGCAAAACTGGCCCCGATCCACCACCTCCCGCCATCGTTGGT
GGGTTGAACACTCACGCTGTTTTCGAAGGTGATGCGAGTATGACCCGAGGCG
ACTTTCACTTGGGGGATAACTTCAACTTCAACCAGACGCTTTGGGAGCAGTTC
AAGGACTACAGTAACCGCTATGGAGGTGGACGATACAACCTAACTGCGGCTG
CTGAGCTTCGCTGGGCACGTATCCAGCAATCCATGGCCACGAACGGTCAATT
CGACTTCACCTCCCCTCGGTACTTCACAGCCTACGCCGAATCCGTTTTCCCA
ATCAACTTCTTCACGGACGGACGGCTCTTCACTTCGAACACTACCGCACCAG
GCCCCGACATGGACTCGGCGCTCTCCTTCTTCCGGGACCACAGGTACCCCA

AAGACTTCCATCGCGCACCCGTTCCAAGTGGTGCTCGTGGACTCGATGTAGT
CGCCGCTGCCTACCCTATCCAGCCGGGCTACAATGCAGATGGGAAGGTGAA
CAACTACGTCCTCGACCCGACTTCCGCGGATTTCACAAAGTTCTGTCTGCTGT
ACGAGAACTTTGTGTTGAAGACTGTGAAGGGGCTCTATCCAAATCCGAGGGG
CTTCTTGAGGAAGGCACTGGAGACAAACTTGGAATACTTTTACCAGTCGTTCC
CTGGGTCGGGAGGCTGCCCGCAGGTCTTCCCCTGGGGCAAGAGTGATTAG

Two polymerase chain reactions (PCR) reactions, one for each gene plus duplicates, were carried out to amplify the r*Cci*UPO and artUPO genes. The primers used to amplify the genes are given in Table 8.2 and the PCR mix for each reaction and the PCR program used was carried out as described in Table 8.3.

Table 8.2 Primers for UPO gene amplification

H-artSTOP-For	CTGTTCCAGGGACAATCCCAGGATATTGTGGACTTC
H-artSTOP-Rev	ATGATGGTCGACGGCTCACTACAAGTTGGGCGAGAGG
H-CciSTOP-For	CTGTTCCAGGGACAATTCCCGGCGTACGCGTTCC
H-CciSTOP-Rev	ATGATGGTCGACGGCTCACTAATCACTCTTGCCCCAGGGG

Table 8.3 **Top**: components for gene amplification. **Bottom**: PCR program

21 µL	deionised water
10 µL	Q5 reaction buffer (5X)
10 µL	Reaction enhancer (5X)
5 µL	dNTPs (2 mM)
1 µL	template DNA
1.25 µL	forward primer (20 uM)
1.25 µL	reverse primer (20 uM)
0.5 µL	Q5 DNA polymerase

98 ºC	30 s	Initial denature	1 cycle
98 ºC	10 s	Denature	
60 ºC	30 s	Anneal	× 35 cycles
72 ºC	30 s	Extend	
70 ºC	3 min	Final extension	1 cycle
10 ºC	∞	Cool	

8.1.2.2. Agarose gel electrophoresis

For analysis of DNA samples, agarose gels (1%) were prepared by heating agarose in TAE buffer (40 mM tris base, 20 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA)). When cooled, SybrSafe dye was added to a final concentration of 1% and the gels were cast. The gels were transferred to a gel running chamber and the chamber was filled with TAE buffer. The DNA samples were prepared by adding a 6X DNA dye and a 10 kb ladder was used as a reference. The gels were run at 100 V for 1 h. The desired bands were then excised and purified using a gel extraction kit (Qiagen).

8.1.2.3. Vector amplification

PCR was used to amplify the pPCIZ vector. The vector contained a signalling peptide with the four mutations used to express *Aae*UPO in yeast, [98] a linker, a HRV 3C

protease site, an N-terminal His-tag and the gene for *Aae*UPO. This variant of the pPICZ plasmid was designed by Mielke previously in the group.[95] Gradient PCR was run, ranging from 55 – 65 ºC (3 vials). The primers used to amplify pPICZ are given in Table 8.4 and the PCR mix for each vial and the PCR program was run as described in Table 8.5.

Table 8.4 Primers for vector amplification

pPICZ 3C His For	GCCGTCGACCATCATCATCATC
UPO sig His3C Rev	TTGTCCCTGGAACAGAACCTCG

Table 8.5 **Top**: components for vector amplification. **Bottom**: PCR program

21 µL	deionised water
10 µL	Q5 reaction buffer (5X)
10 µL	Reaction enhancer (5X)
5 µL	dNTPs (2 mM)
1 µL	template DNA
1.25 µL	forward primer (20 uM)
1.25 µL	reverse primer (20 uM)
0.5 µL	Q5 DNA polymerase

98 ºC	30 s	Initial denature	1 cycle
98 ºC	10 s	Denature	
55-65 ºC	30 s	Anneal	× 35 cycles
72 ºC	3 min	Extend	
72 ºC	10 min	Final extension	1 cycle
10 ºC	∞	Cool	

A sample of the vectors was taken and analysed using gel electrophoresis. The template DNA was digested using *DpnI* (1 µL) in a cutsmart 10X buffer (5 µL) and incubated at 37 ºC for 1 h. The concentrations of the vectors and genes were measured using UV spectroscopy. In-Fusion cloning was used to integrate the genes

into the vector to create the desired recombinant plasmids by mixing 75 ng vector with 55-60 ng of the desired gene along with the In-fusion mix following the protocol given in the In-Fusion cloning kit user manual (Clonetech).

8.1.2.4. Bacterial media

For the growth of stellar cells, low-salt lysogeny broth (LB) was prepared (5 g L^{-1} yeast, 10 g L^{-1} tryptone, 5 g L^{-1} NaCl). To prepare low-salt LB plates, 20 g L^{-1} agar was added to the low-salt LB recipe. Both solutions were made using ultra pure water and were autoclaved before use.

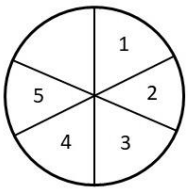
8.1.2.5. Transformation

For each construct following the In-Fusion cloning, 2.5 µL recombinant plasmid was added to 50 µL stellar cells and was left on ice for 20 min in an Eppendorf tube. The tube was then heat shocked to 42 ºC for 45 s and was returned to ice for 2 min. 1 mL low-salt LB was added to the tube and was incubated at 32 ºC at 180 rpm for 1 h. 200 µL cells were then plated out on a low-salt LB plate containing 25 µg mL^{-1} zeocin and was grown overnight at 37 ºC. For the new r*Cci*UPO plasmid, the DNA was isolated using a mini prep kit and the plasmid was incubated with *Xho*I at 37 ºC for 1 h. The digestion was analysed using electrophoresis and sent to Eurofins for sequencing to check for successful cloning.

8.1.2.6. Colony PCR

From the low-salt LB plates prepared in Section 8.1.2.5, five colonies for each homolog were selected. Each colony was transferred using a pipette tip to a section of a new low-salt LB plate (as shown in Figure 8.1), then immediately the tip was washed in a

colony PCR mix. This mix was prepared and the PCR program was run as described in Table 8.6.



1 segment = 1 colony

Figure 8.1 Preparation of separate colonies on a low-salt LB plate prior to colony PCR

Table 8.6 **Top**: components for colony PCR amplification mix. **Bottom**: PCR program

18.65 µL	deionised water
2.5 µL	Taq buffer (10X)
2.5 µL	dNTPs (2 mM)
0.75 µL	$MgCl_2$ (50 mM)
0.25 µL	forward Aox primer (20 uM)
0.25 µL	reverse Aox primer (20 uM)
0.1 µL	Platinum-Taq polymerase

94 ºC	2 min	Initial denature	1 cycle
94 ºC	30 s	Denature	
60 ºC	30 s	Anneal	× 25 cycles
72 ºC	70 s	Extend	
72 ºC	2 min	Final extension	1 cycle
10 ºC	∞	Cool	

The PCR products were isolated on 1% agarose gel and excised as described previously. The concentrations of the DNA determined using UV spectroscopy. 20 µL of the extracted DNA (30-50 ng μL^{-1}) and 20 µL of the primers (10 uM) were sent to GATC for sequencing after colony PCR to ensure that no changes to the desired sequence had occurred.

8.1.2.7. Preparation of linearised vectors

A 5 mL starter culture of stellar cells for each homolog was grown in low-salt LB for 8 h, and then transferred into 45 mL low-salt LB to grow overnight. The plasmids were harvested using a Qiagen midiprep kit and the concentrations were determined using UV spectroscopy with water as a reference. 20 µg of each plasmid was linearised using *SacI* and was prepared as described in Table 8.7.

Table 8.7 Components required to linearise both types of plasmids using *Sac*I

artUPO	**Components**	**r*Cci*UPO**
30 µL	Nuclease free water	45 µL
280 µL	Plasmid	400 µL
35 µL	10X cutsmart buffer	50 µL
5 µL	*Sac*I	5 µL

For the linearisation of r*Cci*UPO in pPICZ, 2 µg of DNA was incubated with two units of *Sac*I for 15 min at 37 °C. The temperature was then raised to 65 °C to deactivate *Sac*I.

Following the digest, the DNA were analysed using electrophoresis on an agarose gel to determine the success of the digestion. If successful, the DNA reaction mixtures were purified using a PCR clean-up kit (Qiagen), or the DNA was purified from the agarose gel using the *SacI* digests liquid protocol in the Qiagen DNA gel purification kit.

8.1.2.8. Yeast growth

For the growth of *P. pastoris* X-33, yeast peptone dextrose (YPD) plates (10 g L^{-1} yeast extract, 20 g L^{-1} peptone, 20 g L^{-1} agar) yeast peptone dextrose sorbitol (YPDS) plates

(10 g L^{-1} yeast extract, 20 g L^{-1} peptone, 2 % (w/v) glucose, 182.2 g L^{-1} sorbitol) and a YPD medium (10 g L^{-1} yeast extract, 20 g L^{-1} peptone, 2% (w/v) glucose) were prepared.

A glycerol stock was used to inoculate a YPD plate and was incubated at 30 ºC for 48 h to obtain single colonies. A 5 mL YPD culture was inoculated with a single colony from the plate and was grown at 30 ºC for 24 h.

For the standard electroporation, 50 mL fresh YPD media was inoculated with 2 mL of the 24 h culture and was incubated at 30 ºC until the OD_{600} reached between 1.3-1.5. The cells were centrifuged at 1500 x g for 5 min at 4 ºC and the cell pellet was resuspended in 50 mL ice cold sterile water. The cells were centrifuged as described and the pellet was resuspended in 25 mL ice cold sterile water. The cells were centrifuged again and the pellet was resuspended in 2 mL ice cold sterile sorbitol (1 M). A final centrifugation was carried out and the pellet was resuspended in 400 µL ice cold sterile sorbitol (1 M).

For growing cells in LiOAc, 50 mL fresh YPD media was inoculated with 5 mL of the 24 h culture in a baffled flask and was incubated at 30 ºC until the OD_{600} reached between 1.5-2.0. The cells were centrifuged at 2000 x g for 5 min at room temperature and the cell pellet was resuspended in 50 mL LiOAc solution with dithiothreitol (DTT) (10 mM) for 30 min. This was spun at 2000 x g for 5 min at 4 ºC, the supernatant was discarded and the cells were kept on ice. The cells were resuspended in 30 mL ice cold sorbitol (1 M), spun and this was repeated, discarding the supernatant each time. The cells were finally resuspended with 0.5 mL ice cold sorbitol (1 M).

8.1.2.9. Electroporation

Either 5-10 µg of linearised vector (1 µg μL^{-1} in water) or 200 ng of digested r*Cci*UPO in pPICZ, and 80 µL electro-competent *P. pastoris* cells were transferred into an ice cold 2 mm electroporation cuvette and was incubated on ice for 5 min. The cells were pulsed using a gene pulser II with a Capacitance Extender Plus and a Pulse Controller Plus (BioRad) with the capacitance set to 25 µF, resistance to 200 Ω and voltage at 2.0 kV. Immediately following the electroporation pulse, 1 mL ice cold sorbitol (1 M) was added to the cuvette and the contents were transferred to a sterile 1.5 mL tube and incubated on ice without shaking for 1.5 h. 100 µL of cells were spread onto a YPDS plate each containing 100 µg mL^{-1} zeocin and were incubated for 48 h. Four colonies from each plate were streaked onto fresh YPD plates containing zeocin (100 µg mL^{-1}) to grow single colonies and were incubated at 30 ºC for 48 h.

8.1.2.10. Small scale expression

Media buffered glycerol complex medium (BMGY) and buffered methanol complex medium (BMMY) were prepared as described in Table 8.8. One colony from each plate, after electroporation, was used to inoculate 5 mL of BMGY and was incubated at 30 ºC at 190 rpm for 24 h. A glycerol stock was made from the starter culture after the 24 h incubation. 250 µL of the starter culture was then added to 5 mL of BMMY and a 250 µL sample was taken. A 10% MeOH feed solution was made from an extra supply of BMMY and was stored at 4 ºC. Every 24 h, a 250 µL sample of the culture was taken and 250 µL MeOH (10 %) feed solution was added to the culture. The samples were centrifuged at 10000 × g for 20 min and the supernatant was stored at −80 ºC.

Table 8.8 Components to make 50 mL of BMGY and BMMY

BMGY		BMMY	
35 mL	yeast extract peptone (YEP) (2X)	35 mL	YEP (2X)
4 mL	Potassium phosphate buffer pH 6.0 (1 M)	5 mL	Potassium phosphate buffer pH 6.0 (1 M)
4 mL	Yeast nitrogen bases (YNB) (10X)	5 mL	YNB (10X)
4 mL	glycerol (10 %)	5 mL	MeOH (5 %)
80 µL	biotin (500X)	100 µL	biotin (500X)
40 µL	zeocin		

8.1.2.11. SDS Page analysis

The following buffers described in Table 8.9 were prepared for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE).

Table 8.9 Recipes to make buffers required for SDS PAGE analysis

Stacking buffer	
0.5 M	Tris-Cl
0.4%	SDS
pH	6.8

Resolving buffer	
1.5 M	Tris-Cl
0.4%	SDS
pH	8.8

4X SDS Running buffer	
767 mM	Glycine
100 mM	Tris base

Magic Dye	
60 g L^{-1}	G250 Blue dye
1%	HCl

5X SDS PAGE Sample Buffer	
250 mM	Tris-Cl pH 6.8
30%	Glycerol
10%	SDS
0.02%	Bromophenol blue
710 mM	β-mercaptoethanol

For the analysis, 12% acrylamide SDS gels were made using the recipe in Table 8.10.

Table 8.10 Recipes for the resolving and stacking 12% acrylamide SDS gels

Resolving gel		Stacking gel	
3.2 mL	Water	3.2 mL	Water
2.5 mL	Resolving buffer	1.3 mL	Stacking buffer
4.2 mL	Acrylamide stock	0.5 mL	Acrylamide stock
100 µL	10% APS	10 µL	Bromophenol blue
8 µL	TEMED	50 µL	10% APS
		8 µL	TEMED

Samples were prepared with 5X SDS PAGE Sample buffer to a total volume of 20 µL and were heated to 95 ºC for 5 min. They were then loaded onto 12% acrylamide SDS gels along with a low molecular weight marker (BioRAD). Running buffer was made by the addition of 0.1% SDS and added to the system. Gels were run at 200 V for 50 min and visualised using Magic Dye.

8.1.2.12. Western Blotting

The samples were prepared and run using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using a pre-cast gel. 40 µL of each sample was taken, was added to 10 µL 5X dye and was incubated at 100 ºC for 3 min to denature the protein. When cooled, 20 µL of each sample and a molecular weight marker (BioRAD) were loaded into a pre-cast gel and the gel was transferred into a running chamber, where it was covered by a running buffer (62.8 g L^{-1} MOPS, 72.6 g L^{-1} Tris, 6 g L^{-1} SDS, 6.5 g L^{-1} Na bisulfite) and was run at 150 V for 50 min. The protein in the gel was then transferred to blotting paper using a blotting machine for 8 min. The blotting paper was soaked in Ponceau S dye (0.1 % Ponceau S, 5 % acetic acid) and was then rinsed out with a small amount of milk/ TBST (tris-buffered saline solution with Tween-20). The paper was washed and incubated in 5 min intervals with milk/ TBST four times

before incubation overnight with milk/ TBST and monoclonal anti-polyhistidine peroxidase antibody. The paper was washed again with milk/ TBST before imaging.

8.1.2.13. Small-scale shake flask expression

From a glycerol stock made after electroporation, cells were resuscitated by streaking onto a YPD plate with zeocin (100 µg mL^{-1}) and grown for three days at 30 ºC. Twelve colonies were selected and were each grown in 5 mL BMGY starter cultures for 24 h at 220 rpm at 30 ºC. Six of the cultures were pooled together and were used to inoculate a 0.6 L BMMY culture in a 2.5 L baffled shake flask and were grown at 180 rpm at 30 ºC for 72 h. The other six cultures were pooled and added to BMMY in the same way as described previously, but were grown for 144 h. Every 24 h, a 250 µL sample was taken and 30 mL of a 10 % MeOH feed was added. At the time of harvest, the cultures were centrifuged at 5,000 × g and the supernatant was stored at −80 ºC for further analysis. The cell pellet was discarded.

8.1.2.14. 200 mL Fed-Batch Fermentation

The procedure for the fed-batch fermentation followed the instructions given in the Pichia Fermentation Process Guidelines (Invitrogen) and were adapted for use of a 200 mL fermentation.

8.1.2.15. Fermenter preparation

In a 500 mL glass vessel, 200 mL of the basal salts medium was prepared at pH 5 using the recipe given in Table 8.11. The pH probe was calibrated using a two-point system at pH 4 and pH 7 and the electrolyte in the dissolved oxygen (DO) probe was replenished. To the glass vessel, the antifoam, base and carbon-source addition

tubing was attached and the vessel was autoclaved. After cooling, the vessel was attached to a condenser, to control the temperature of the culture, and an O_2 source. The following parameters were set: temperature at 30 ºC, pH 5, stirrer limits set to 200 - 1750 rpm, airflow at 100 mL min^{-1}, DO at 30%, condenser with pump output of 30 - 75%, and antifoam with pump output of 15%. The pH was adjusted to 5.0 using ammonium hydroxide (undiluted), which also functioned as a nitrogen source, and the PTM_1 trace salts were added (4.35 mL L^{-1}) aseptically. Two vessels were prepared for two identical fermentations.

Table 8.11 Components to make Fermentation Basal Salts Medium and PTM_1 salts

Basal salts medium		**PTM_1 Trace salts**	
26.7 mL L^{-1}	H_3PO_4, 85%	6.0 g L^{-1}	$CuSO_4 \times 5\ H_2O$
1.17 g L^{-1}	$CaSO_4 \times 2\ H_2O$	0.08 g L^{-1}	NaI
18.2 g L^{-1}	K_2SO_4	3.0 g L^{-1}	$MnSO_4 \times H_2O$
14.9 g L^{-1}	$MgSO_4 \times 7\ H_2O$	0.2 g L^{-1}	$Na_2MoO_4 \times 2\ H_2O$
4.13 g L^{-1}	KOH	0.02 g L^{-1}	H_3BO_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}	$FeSO_4 \times 7\ H_2O$
		0.5 g L^{-1}	Biotin
		5.0 mL L^{-1}	H_2SO_4

8.1.2.16. Glycerol Batch Phase

From a glycerol stock, *P. pastoris* cells were streaked out onto a YPD plate with zeocin (100 µg mL^{-1}) and grown for three days at 30 ºC. Eight colonies were selected and each were grown in 5 mL BMGY starter cultures for 24 h at 220 rpm at 30 ºC. Four cultures were pooled together and added to each vessel (2 × 20 mL BMGY cultures). The batch was grown until all the glycerol had been consumed.

8.1.2.17. Methanol Fed-Batch Phase

Once the glycerol had been consumed, a 100% methanol feed with PTM_1 salts (12 mL L^{-1}) was added at 3.6 mL h^{-1} L^{-1} of initial fermentation volume for 4 h. When the culture had adapted to the methanol feed rate, had a steady DO % and then had a fast DO spike after stopping the methanol source, the methanol feed rate was doubled to 7.3 mL h^{-1} L^{-1} of initial fermentation volume. After the culture adapted to the increase in feed rate, the feed rate was increased again to 10.9 mL h^{-1} L^{-1} of initial fermentation volume. Samples were taken at least every 24 h, with the first sample taken at the beginning of the methanol fed-batch phase. The samples were centrifuged and the supernatant was frozen at −20 ºC for SDS-PAGE analysis and Western blotting. In addition, the wet cell weight was calculated. The phase lasted for either 56 h or 72 h and the cultures were harvested. The cultures were spun at 10,000 × g and the supernatant was decanted and stored at −80 ºC.

8.1.2.18. 2 L Fed-Batch Fermentation

8.1.2.18.1. 2 L Fermenter preparation

In a 5 L glass vessel, 2 L of the basal salts medium were prepared at pH 5. The pH probe was calibrated using a two-point system at pH 4 and pH 7 and the electrolyte in the dissolved oxygen (DO) probe was replenished. To the glass vessel, the antifoam, base and carbon-source addition tubing was attached and the vessel was autoclaved. After cooling, the vessel was attached to a condenser, a heating jacket, an air source and an O_2 source. The following parameters were set: temperature at 30 ºC, pH 5, stirrer limits set to 1000 rpm, airflow at 2 mL min^{-1}, O_2 flow at 0.2 mL min^{-1} and DO at 30%. The pH was adjusted to 5.0 using ammonium hydroxide (undiluted), which also functioned as a nitrogen source, and the PTM_1 trace salts were added (4.35 mL L^{-1}).

8.1.2.18.2. Glycerol Batch Phase

From a glycerol stock, *P. pastoris* cells were streaked out onto a YPD plate with zeocin (100 µg mL^{-1}) and grown for two days at 30 ºC. Two colonies were selected and each were grown in 5 mL BMGY starter cultures for 24 h at 220 rpm at 30 ºC. 1 mL of each culture was then inoculated into 2 × 100 mL BMGY and was shaken at 220 rpm at 30 ºC. A total of 200 mL of starter culture was then added to the fermenter and was left to grow until the glycerol had been consumed (roughly 20 h).

8.1.2.18.3. Methanol Fed-Batch Phase

Once the glycerol had been consumed, a 100% methanol feed with PTM_1 salts (12 mL L^{-1}) was added at 0.003 mL min^{-1} for 2.5 h. When the culture had adapted to the methanol feed rate, had a steady DO % and had a fast DO spike after stopping the methanol source, the methanol feed rate was increased to 0.030 mL min^{-1} with a 2 sec off pulse. This rate was slowly increased to 0.050 mL min^{-1} over the two days of the methanol fed-batch phase until harvest, where the culture was centrifuged at 15000 × g for 30 min and a total of 2.5 L of supernatant was either concentrated using a tangential flow diafiltration unit (30 kDa filter) or frozen at −80 °C until further use.

8.1.3. Expression of artUPO in *E. coli*

8.1.3.1. Gene preparation

The artUPO gene was purchased from Genscript in a pEt-28a(+) vector which included an N-terminal His tag and kanamycin resistance, and the sequence was codon optimised for *E. coli* as shown in Table 8.12.

Table 8.12 Codon optimised gene sequence of artUPO for the expression in *E. coli* in pET-28a(+)

artUPO DNA sequence codon optimised
CATATGAGCCAGGACATCGTGGATTTTAGCCAACATCCGTGGAAAGCGCCGG
GTCCGAACGACCTGCGTAGCCCGTGCCCGGGTCTGAACACCCTGGCGAACC
ACGGTTTCCTGCCGCGTAACGGCCGTAACATCACCATTCCGATGATCGTGCA
GGCGGGTTTTGACGGCTACAACGTTCAACCGGATATCCTGATTCTGGCGGCG
AAGGTTGGTCTGCTGACCAGCCCGGAGCCGGACACCTTCACCCTGGACGAT
CTGAAACTGCACGGTACCATTGAGCACGATGCGAGCCTGAGCCGTGAAGACT
TTGCGCTGGGCGATAACCTGCACTTCAACGAAGCGATCTTTAACACCCTGGC
GAACAGCAACCCGGGTAGCGATGTGTACAACATTACCAGCGCGGGCCAGGT
TCTGAAGGACCGTCTGGCGGATAGCCTGGCGCGTAACCCGAACGTGACCAA
CACCGGCAAAGAGTTCACCATTCGTACCCTGGAAAGCGCGTTTTATCTGAGC
GTGATGGGTAACGCGACCACCGGTGAAGCGCCGAAGAACTTTGTTCAAATCT
TCTTTCGTGAGGAACGTCTGCCGATTGAGGAAGGTTGGAAACGTAGCACCAC
CCCGATCACCAGCGACACCCTGAACCCGATTGCGGGTCAGATTAGCGAAGC
GAGCAACTGGAAACCGAACCCGGATCAATGCCCGTGGATTGTTCTGAGCCCG
AACCTGTAACTCGAG

8.1.3.2. Bacterial strains and media

For the growth of the *E. coli* expression strains, lysogeny broth (LB) was prepared (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 10 g L^{-1} NaCl). To prepare LB plates, agar was added to LB (15 g L^{-1}). Both solutions were made using ultra pure water and were autoclaved before use. Antibiotics were added to media according to commercial protocols for each strain as needed.

8.1.3.3. Expression tests of E. coli expression strains

50 µL of three different *E.coli* strains: BL21 (DE3), Rosetta PlysS (DE3) and B834 was transformed with 2 µL DNA. The strains were heat shocked at 42 °C for 10 s, 45 s, and 30 s respectively, placed on ice for 2 min, and incubated in 150 µL LB media at

37 °C for 1 h. After this period, the cultures were streaked onto a LB plate with kanamycin (30 μg mL^{-1}) or in the case of Rosetta, an additional antibiotic chloramphenicol was added to the plate (33 μg mL^{-1}). The plates were incubated at 37 °C overnight for colonies to form. A single colony from each plate was picked, added to 10 mL LB with kanamycin (30 μg mL^{-1}) and additional chloramphenicol for Rosetta (33 μg mL^{-1}), and incubated at 37 °C at 180 rpm until the OD_{600} reached 0.7. To induce gene expression, 5 μL IPTG (1 M), 10 μL 5-ALA (0.5 M) and 10 μL $FeSO_4$ (0.2 M) was added to each culture and was incubated at 180 rpm at either 37 °C for 2 h, 30 °C for 5 h or 16 °C overnight.

To harvest the protein, the cultures were spun at 10000 × g for 20 min , the pellets were collected and resuspended in 1 mL of nickel buffer A (Table 8.13). The cells were then lysed using sonication with pulsing of 10 s on, 10 s off, repeating three times. These mixtures were spun at 15000 × g for 40 min and the supernatants were collected (soluble fraction). The cell pellets was suspended in 700 μL 1X running buffer used for SDS PAGE (insoluble fraction). These samples were analysed using SDS PAGE.

8.1.3.4. 1 L Shake Flask Growth

From the expression tests described as above, Rosetta PLysS (DE3) was chosen to express artUPO. To grow a 1 L culture, Rosetta cells were transformed and spread onto a LB plate as described in Section 8.1.3.3. A single colony was picked and was added to 10 mL LB with 30 μg mL^{-1} kanamycin and 33 μg mL^{-1} chloramphenicol to grow overnight at 37 °C at 180 rpm. In the following morning, the 10 mL culture was added to 1 LB with kanamycin and chloramphenicol and was incubated at 37 °C at

180 rpm until the OD_{600} reached 0.7. To induce expression, 1 mL IPTG (1 M), 2 mL 5-ALA (0.5 M) and 2 mL $FeSO_4$ (0.2 M) was added and the culture was incubated overnight at 16 °C at 180 rpm.

To harvest the protein, the cultures were spun at 5000 × g for 20 min, the pellets were collected and resuspended in 30 mL of nickel buffer A (Table 8.13). The cells were then lysed using a cell disruptor with a pressure of 26 kPsi and were then centrifuged at 15000 × g for 40 min. The supernatant was then collected for purification.

8.1.4. Purification of artUPO$_{yeast}$ and r*Cci*UPO expressed in yeast

The pH of the supernatant was adjusted to pH 8.0 either directly with acid or base, dialysed prior to purification or was passed through a desalting column with the salt-free buffer. If the protein had an intact His-tag, it was purified using immobilised nickel affinity chromatography followed by size exclusion chromatography (SEC). If the His-tag was lost, the protein was purified using anion exchange chromatography (Q) followed by SEC. The buffers used for protein purification are given in Table 8.13.

Table 8.13 Recipes of buffers required for protein preparation and purification

Nickel buffer A	
20 mM	Tris
300 mM	NaCl
20 mM	Imidazole
10%	Glycerol
pH	8.0

Nickel buffer B	
50 mM	Tris
300 mM	NaCl
300 mM	Imidazole
10%	Glycerol
pH	8.0

Salt-free buffer	
50 mM	Tris
pH	8.0

Anion exchange buffer	
50 mM	Tris
2 M	NaCl
pH	8.0

Size exclusion buffer	
50 mM	Tris
300 mM	NaCl
10%	Glycerol
pH	8.0

8.1.4.1. Immobilised Nickel Affinity Chromatography

A 5 mL HisTrap FF crude column (GE Healthcare) was equilibrated with 5 column volumes (CV) of buffer A before loading the protein sample. The run was carried out using the FPLC system at 2.5 mL min^{-1} (ÄKTA Pure or ÄKTA Start, GE Healthcare). The column was washed with 5 CV of the nickel buffer A, followed by a gradient run from 20 mM to 300 mM imidazole over 20 CV. The flow throughs and fractions of interest were collected and analysed using SDS PAGE analysis.

8.1.4.2. Anion Exchange Chromatography

Protein samples for anion exchange chromatography were dialysed in a salt-free buffer prior to purification. A 5 mL HiTrap Q HP column (GE Healthcare) was equilibrated with 5 column volumes (CV) of salt-free buffer before loading the protein

sample. The run was carried out using the FPLC system at 5 mL min^{-1} (ÄKTA Pure or ÄKTA Start, GE Healthcare). The column was washed with 5 CV of the salt-free buffer, followed by a gradient run from 0 mM to 300 mM NaCl over 20 CV. The column was then washed with 5 CV of the anion exchange buffer. The flow throughs and fractions of interest were collected and analysed using SDS PAGE analysis.

8.1.4.3. Size Exclusion Chromatogrpahy (SEC)

The volume of the protein samples was concentrated to 2 mL and centrifuged at 2000 × g prior to purification using a HiLoad 16/600 Superdex 75 prep grade column (GE Healthcare). The column was equilibrated with SEC buffer before and the sample was loaded directly onto the column. The sample was eluted with 1.2 CV of size exclusion buffer at 0.8 mL min^{-1} using the FPLC system (ÄKTA Pure or ÄKTA Start, GE Healthcare). Fractions of interest were collected and analysed using SDS PAGE analysis.

8.1.5. Biochemical assays

8.1.5.1. UV/vis colour activity assays

Two different reactions were carried to determine if the UPO possessed peroxidative and peroxygenative activity. For a simple quantitative analysis, a spectrophotometer was not needed. The buffers used in these assays are given in Table 8.14. For the peroxidase activity assay, a 1 mL aqueous reaction was prepared containing 50 mM citrate phosphate, 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), **118a** (final concentration 0.2 mM), 10 µL UPO sample and the reaction was initiated with

hydrogen peroxide (final concentration (2 mM). For a positive result, the solution will change colour from colourless to blue (Figure 8.2).

For the peroxygenase activity assay, a 1mL reaction was prepared containing 50 mM potassium phosphate buffer, 1,2-(methylenedioxy-4-nitrobenzene) (NBD), 1**19a** (final concentration 1 mM) with 10% acetonitrile, 10 µL UPO sample and the reaction was initiated with hydrogen peroxide (final concentration (2 mM)). for a positive result, the solution will change colour from colourless to yellow (Figure 8.2).

Table 8.14 Recipe for buffers used in activity assays

Citrate phosphate buffer (50 mM)		**Potassium phosphate buffer (50 mM)**	
200 mM	Na_2HPO_4	61.5% mol	K_2HPO_4
100 mM	Citric acid	38.5% mol	KH_2PO_4
pH	4.4	pH	7.0

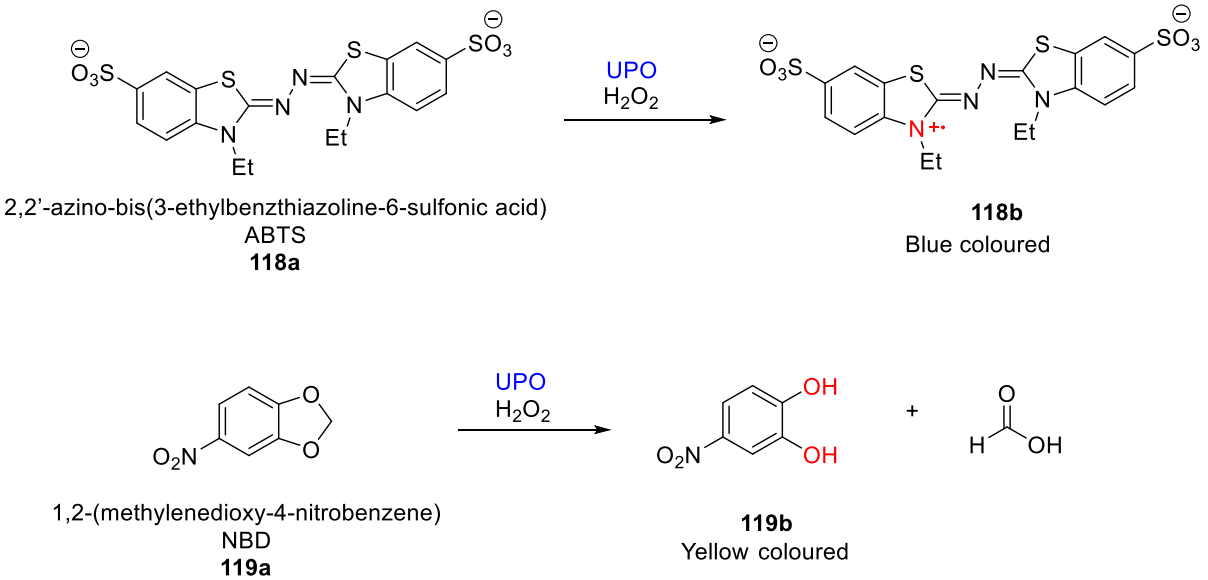


Figure 8.2 UV/vis colour activity assays for the quantitative determination of peroxidase and peroxygenase activity using compounds **1118a** or **119a** respectively

8.1.5.2. Bradford assay

To make a standard curve, a range of solutions with bovine serum albumin (BSA) were prepared with concentrations varying from 0 -1000 µg mL^{-1}. 20 µL of BSA solution was mixed with 1 mL of BioRAD Bradford assay reagent and left at room temperature for 20 min. The solution changed colour from red-brown to blue depending on the concentration of protein. After, the absorbance was measured for each concentration at 595 nm and a standard curve of absorbance against BSA concentration was plotted.

To determine the concentration of an unknown protein, 20 µL of protein sample was mixed with 1 mL of BioRAD Bradford assay reagent and left at room temperature for 20 min. The absorbance at 595 nm of the sample was then compared to the standard curve created with BSA.

8.1.5.3. Deglycosylation using PNGase F

To deglycosylate the UPO, 5-10 µg of UPO was added to 1 µL glycoprotein denaturing buffer (10X) and the mixture was adjusted to 10 µL H_2O. This mixture was incubated at 100 ºC for 10 min, then cooled and centrifuged for 10 s. To this mixture, 2 µL glycol-buffer 2 (10X) and 2 µL NP-40 (10%) were added, and the total volume was adjusted to 20 µL with H_2O. 1 µg of PNGase F was added and gently mixed before incubating at 37 ºC for 1 h. The resulting mixture was analysed using SDS-PAGE.

8.1.5.4. UV/vis spectroscopy

The UPO was scanned from 800 – 250 nm at 30 nm min^{-1}. The Reinheitszahl (Rz) value of the protein was determined by the equation given below. [96]

Equation 1. Calculation for Reinheitszahl value

$$Rz = \frac{A_{418\ nm}}{A_{280\ nm}}$$

Where Rz = Reinheitszahl value, A = absorbance

8.1.6. Mass spectrometry experiments

For analysis using mass spectrometry, artUPO$_{yeast}$ was purified from a fermentation and was sent to the Technology Facility, Metabolomics & Proteomics, University of York for ESI-MS intact mass spectrometry. The sample was diluted 1 in 10 (v/v) with aqueous 50% acetonitrile containing 1% formic acid before infusion at 3 µL/min into a maXis qTOF mass spectrometer. A spectrum was also acquired with the source ion funnel voltage raised from zero to 80 eV.

8.2. Results and Discussion

8.2.1. Expression and purification of artUPO$_{yeast}$ and *CciUPO* from *P. pastoris*

8.2.1.1. Gene and Plasmid Design

In 2014 Molina-Espeja *et al*. developed the first heterologous system for expressing UPOs in *S. cerevisiae* where directed evolution was used to optimise the secretion levels and catalytic activity of *Aae*UPO. A total of nine mutations were introduced to create the mutant Pada-I; four mutations were located in the signal peptide (F12Y-A14V-R15G-A21D), which originated from the native organism *A. aegerita*, and five were found in *AaeUPO* itself (V57A-L67F-V75I-I248V-F311L). The changes in the signalling peptide led to an increase in secretion by 41-fold in *S. cerevisiae* compared to using the wild-type in the same host. [96]

Soon after their work in *S. cerevisiae*, the same group translated this expression system into *P. pastoris* where the system could be scaled up to a large fed-batch fermentation for potential industrial use. Pada-I, which included the mutated signalling peptide, was expressed in high yields of 217 mg mL^{-1} and was very similar in terms of catalytic activity and stability. [98] In this project, we decided to use a similar plasmid and mutated signalling peptide in our expression system in *P. pastoris* X-33 due to its success in producing high yields of UPO. In the previous project in our group, several changes were made to the pPICZα B plasmid used in the Molina-Espeja study. This included deletion of the *N*-terminal α-factor, a native *S. cerevisiae* secretion signal which may aid secretion, and the *C*-terminal *c-myc* epitope tag, which allows detection

of the recombinant protein with the Anti-*myc*-antibody. [95, 155] Other features were introduced including an *N*-terminal 6 × His-tag, HRV-3C protease cleavage site and InFusion overhangs to allow the introduction of desired UPO genes. [95] This plasmid was called pPICZ and was used to express the genes involved in this study, artUPO and recombinant *Cci*UPO (will be now called r*Cci*UPO). The vector map of pPICZ is shown in Figure 8.3.

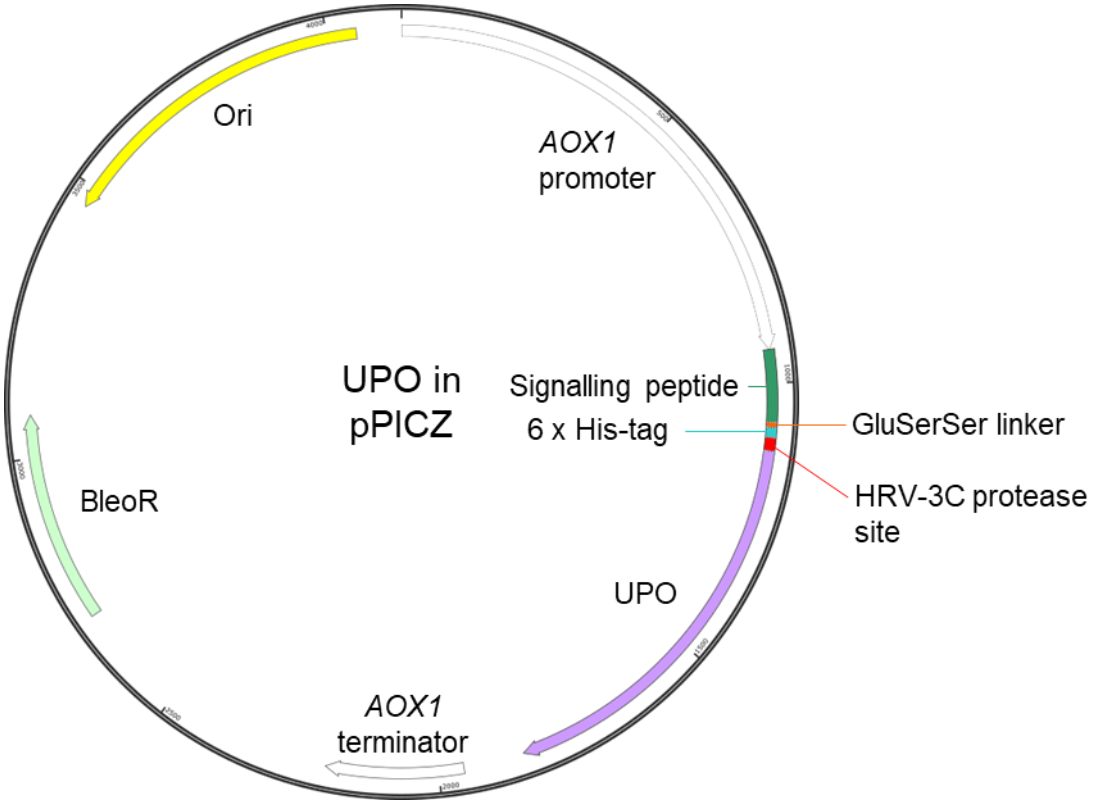


Figure 8.3 Map of pPICZ expression vector containing the UPO gene. UPO expression is controlled by the *AOX1* operon. The construct consists of a signalling peptide (dark green), a linker (orange), a 6x His-tag (blue) and the gene for the UPO (purple). Resistance for zeocin is conferred by the Bleomycin resistance protein (BleoR) (pale green). The origin of replication (ori) is shown in yellow, which is the site of DNA replication and enables the plasmid to reproduce itself.

The expression of the UPO in pPICZ is controlled by the *AOX1* operon. In this system, the introduction of methanol induces the expression of the *AOX1* gene to metabolise the methanol as its source of carbon. Adding an excess of methanol will lead to overexpression of the *AOX1* gene and hence the UPO gene, and the UPO is secreted into the media. Because *P. pastoris* secretes low levels of native proteins, heterologous secretion is desired as it simplifies the purification of the protein in downstream processes. A BleoR gene is also found in pPICZ for selection of cells containing the plasmid using the antibiotic zeocin.

The 6x His-tag at the N-terminal of the UPO allows for the purification of the protein *via* immobilised nickel affinity chromatography. The presence of the HRV-3C protease site gives the option to cleave the His tag from the protein which may be useful for structural purposes, such as crystallography.

8.2.1.2. Preparation of the gene insert

The artUPO and r*Cci*UPO genes plus duplicates were amplified using PCR and the gels were imaged as shown in Figure 8.4. The amplification was judged to be successful because the bands that appeared on the gel corresponded to the expected size of the gene, where the size of the r*Cci*UPO gene is ~1200 kbp and the artUPO gene is ~750 kbp.

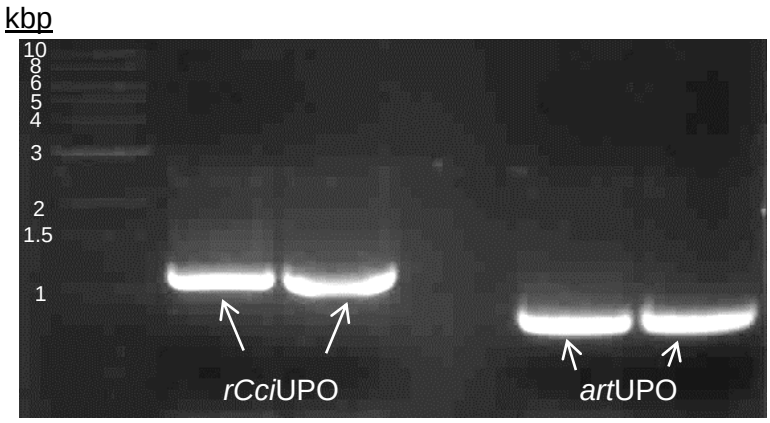


Figure 8.4 1 % agarose gel showing successful gene amplification of *Cci*UPO and artUPO

8.2.1.3. Preparation of the pPICZ vector

The vector pPICZ was amplified using PCR with three different annealing temperatures at 55 ºC, 60 ºC and 65 ºC. The three different vectors were named vector 1, 2 and 3 for each different annealing temperature respectively. The agarose gel in Figure 8.5 indicates that the amplification of vectors 2 and 3 was successful whereas vector 1 was not.

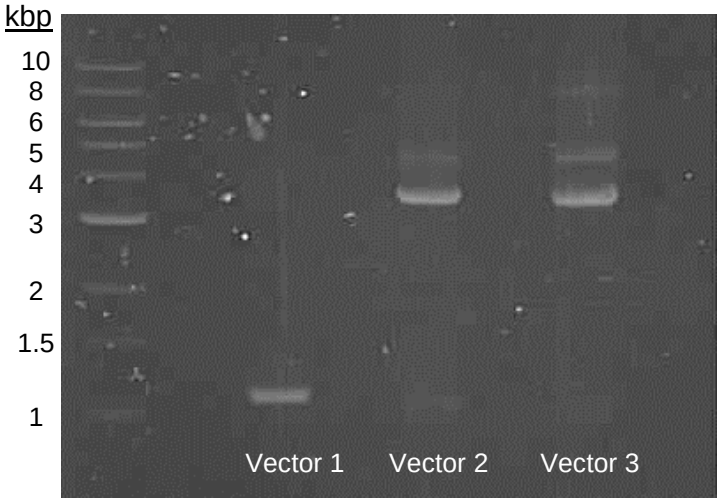


Figure 8.5 1 % agarose gel of the amplification of vector 1, vector 2 and vector 3 at annealing temperatures of 55 ºC, 60 ºC and 65 ºC respectively

The DNA containing the highest concentration of the successful vector and UPO gene were selected and used for InFusion cloning. The recombinant plasmids were incorporated into stellar cells using heat shock method and colony PCR was used to determine the success of the incorporation. The use of low-salt LB for bacterial cell growth was necessary to prevent deactivation of zeocin due to high salt concentrations. From Figure 8.5, Colony 5 with the r*Cci*UPO plasmid and colony 2 with the artUPO plasmid had the highest concentrations of DNA. This concentration was determined using a UV spectrometer. These colonies can be seen as the brightest bands in the gel (Figure 8.6). These colonies were grown, and their recombinant plasmids were extracted and linearised using *Sac*I for electroporation into *P. pastoris*.

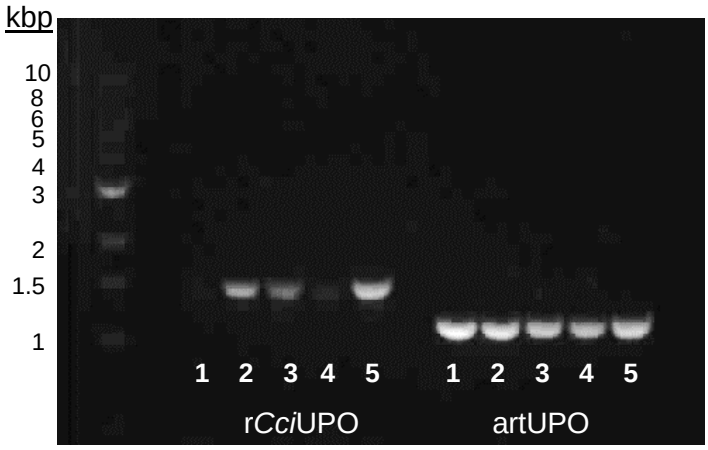


Figure 8.6 1 % agarose gel of colony PCR. Five colonies were selected for each UPO plasmid. Colony 5 from r*Cci*UPO and colony 2 from artUPO were selected to continue

8.2.1.4. Expression tests

P. pastoris colonies containing the UPO plasmids after electroporation were streaked onto a YPD plate with zeocin, and four colonies corresponding to each UPO were transferred to grow in 5 mL BMGY for 24 h. Each of the samples were then transferred to 5 mL BMMY starter cultures and samples were taken every 24 h to determine whether electroporation was successful, if the UPO could be expressed in *P. pastoris*,

and the optimum growth time to produce the maximum amount of protein. A western blot revealed that the expression of r*Cci*UPO was not successful as no protein was observed (Figure 8.7), even when the sample was concentrated, partially purified or repeated. Failure of expression could be due to a number of reasons, including problems with the signalling peptide, misfolding of the protein or loss of the His-tag.

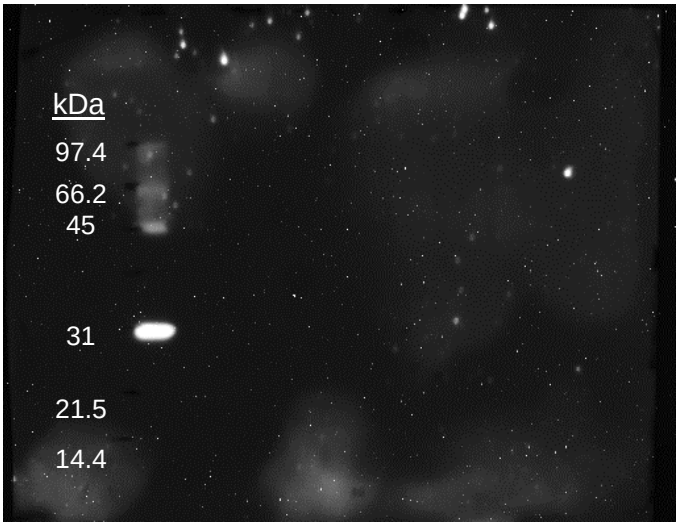


Figure 8.7 Western blot of the time course for r*Cci*UPO expression with four colonies. No expression of protein can be observed

The expression of artUPO$_{yeast}$ appeared to be very successful and the intensity of the band for colony A increases over time (Figure 8.8). The optimum time of growth after inoculation for colony A appeared to be 72 h. This is because the brightest band corresponding to the protein was observed using colony A and the band becomes brighter over time. The calculated molecular mass of artUPO$_{yeast}$ is 32 kDa but the mass appears to be larger in the western blot. This may be due to the glycosylation of the protein and Figure 8.8 also indicates that there are two glycoforms of artUPO$_{yeast}$ due to the presence of two bands. Further investigation of the two bands were carried out after protein purification later in this chapter. Because of these observations, colony A was chosen for a 0.6 L shake flask expression.

Colonies

Time /h

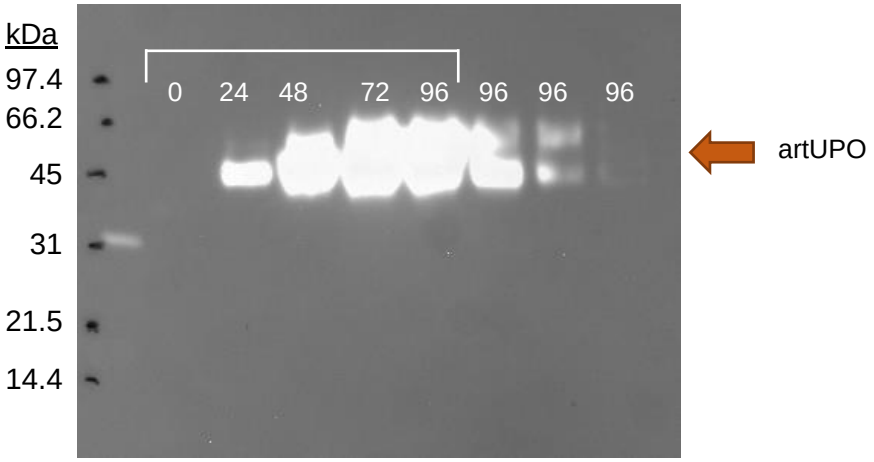


Figure 8.8 Western blot of the time course for artUPO$_{yeast}$ expression with four colonies A, B, C and D. Expected molecular weight of artUPO is 32 kDa not accounting for glycosylation and the corresponding bands are marked next to the gel

8.2.1.5. 0.6 L Shake flask expression

For the shake flask expression two identical flasks were prepared, however one was harvested 72 h after inoculation and the other was harvested 144 h after inoculation. This required resuscitation from glycerol stocks and the transfer of six colonies into 6 × 5 mL BMGY cultures. These cultures were then transferred to 600 mL BMMY and samples were taken from this point. A western blot with these samples (Figure 8.9) showed that a harvest after 72 h is optimal because the His-tag remains attached to the protein, which simplifies the method for protein purification. There could also be possible degradation of the protein after 72 h as the protein cannot be observed in the Western blot as strongly after 72 h.

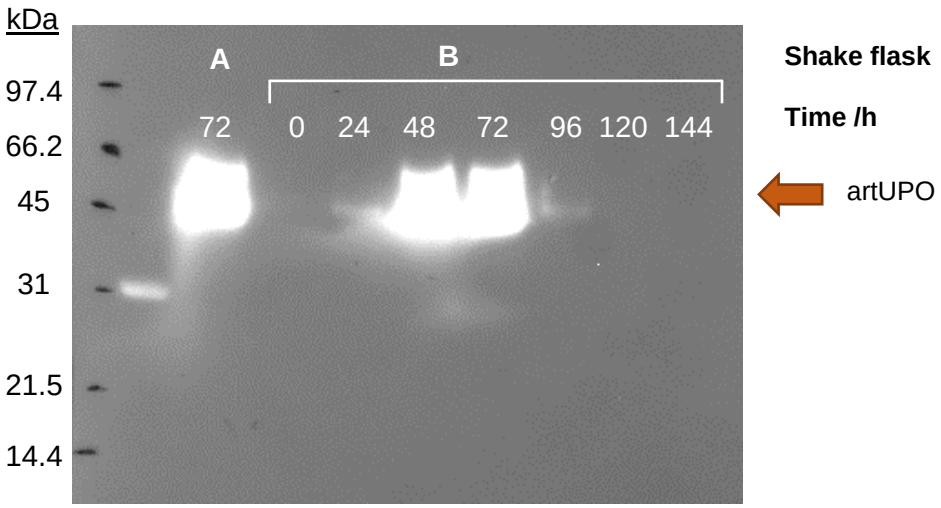


Figure 8.9 Western blot of the time course for artUPO$_{yeast}$ in shake flasks with shake flask A (harvest at 72 h) and shake flask B (harvest at 144 h). Expected molecular weight of artUPO is roughly 45 kDa and the corresponding bands are marked next to gel

8.2.1.6. 200 mL fermentation

Two 200 mL reactors were prepared in an identical manner, except one was harvested 56 h after the start of the methanol fed-batch phase and one was harvested after 72 h. The batch fermentation differs to shake flask expression because the former gives rise to a higher cell density in a given volume and hence more protein is produced per litre. However, a number of factors must be controlled to keep the culture healthy such as the dissolved oxygen (DO) content, temperature, pH and methanol content.

P. pastoris cells will uptake dissolved oxygen in solution for metabolic processes, whether they are grown in glycerol or methanol, so the DO% of the solution will be low. Therefore, the DO% is representative of the health of the culture. If either glycerol or methanol has been consumed, the cells will stop metabolising and will stop

consuming oxygen. As a result, the DO% of the culture will suddenly increase. This is known as a DO spike.

The fermentation was divided into two phases, the glycerol batch phase and the methanol fed-batch phase. The glycerol batch phase was the time for the culture to consume all the remaining glycerol from the start up media BMGY which is indicated by the presence of a DO spike (as shown in Figure 8.10). This phase lasted for roughly 24 h. Afterwards, the methanol phase begins where a constant methanol source is added to the culture. A low feed rate is used initially to allow the culture to adapt to the presence of methanol. Once the culture adapted to the methanol, the feed rate was slowly increased until it reached 10.9 mL h^{-1} L^{-1}. As seen in Figure 8.10, the levels of DO% becomes more erratic as the demand for resources increase as the number of cells increase. When a shortage of a resource occurs, the cells temporarily stop metabolising and a DO spike occurs.

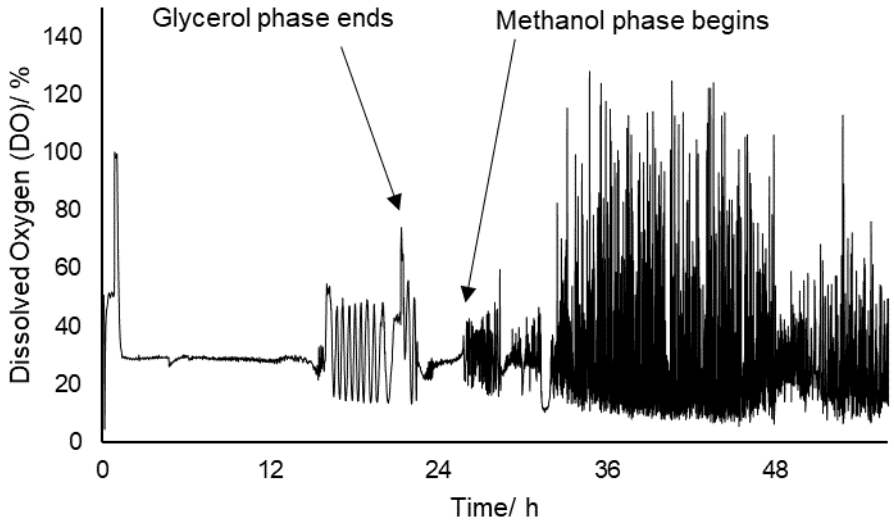


Figure 8.10 Fermentation log for the 56 h artUPO$_{yeast}$ fermentation at 200 mL scale showing the dissolved oxygen (DO) % over the course of the fermentation with the end of the glycerol phase and the beginning of the methanol phase shown

Samples of the cultures were taken at least every 24 h from the beginning of the methanol fed-batch phase. The samples were centrifuged and the supernatant was kept for SDS-PAGE analysis (Figure 8.11).

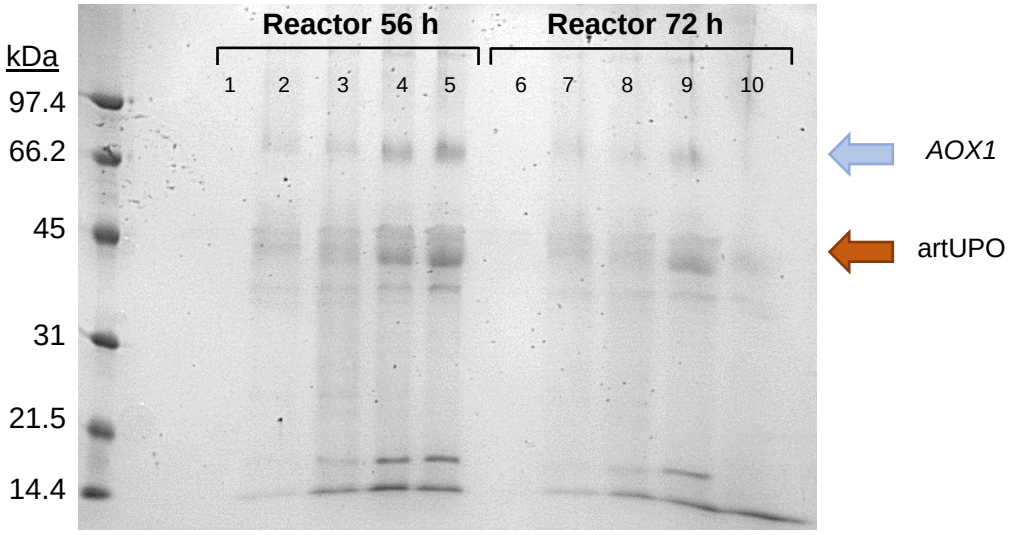


Figure 8.11 SDS-PAGE of the samples of the reactors harvested after 56 h and 72 h. Lanes 1 – 5 are the 0 h, 24 h, 32 h, 48 h and 56 h samples of reactor 56 h. Lanes 6 – 10 are the 0 h, 24 h, 32 h, 48 h and 72 h samples of reactor 72 h. On an SDS PAGE gel, artUPO$_{yeast}$ and *AOX1* appear to have molecular weights of roughly 45 kDa and 66 kDa respectively. The corresponding bands are marked next to the gel.

artUPO$_{yeast}$ is expressed in both reactors over the time course but there appears to be stronger expression in reactor 56 h for the same time sample at 48 h. In fact, very little protein is observed at 72 h in the 72 h reactor, suggesting there may be degradation in the protein over time. Only the culture in reactor 56 h was harvested and kept, whereas the culture in reactor 72 h was discarded. Another band observed at 66.2 kDa in both reactors may correspond to the intracellular alcohol oxidise (*AOX1*) found in the *P. pastoris* cells. This is not seen in the shake flask expression. *AOX1* may be seen in the fermentation culture because of a larger number of cells that may die and

lyse, causing *AOX1* to escape the cells in the culture. For use of artUPO$_{yeast}$ in further experiments, the culture was spun and the supernatant was kept.

The wet cell weight of the samples was also calculated by taking a 250 µL sample and centrifuging at 16000 × g for 5 min . The supernatant was discarded and the mass of the cell pellet was weighed. The wet cell weight could then be calculated per litre and the results are shown in Figure 8.12. The cell wet weights for both fermenters increase over time indicating cell growth, and the approximate weights at these time points match with the expected weights given in the Pichia Fermentation Process Guidelines (Invitrogen).

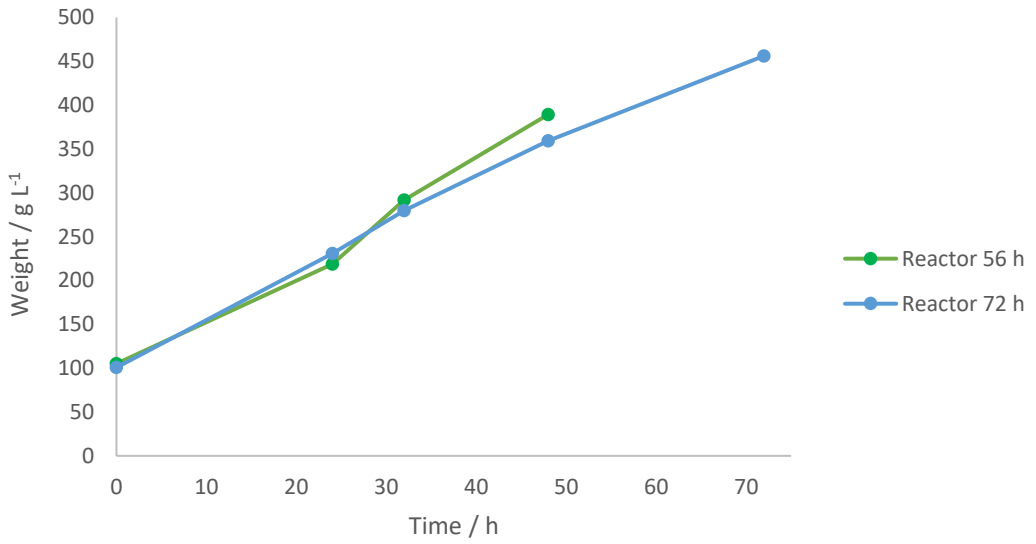


Figure 8.12 The wet cell weight of the 200 mL fermentation over time for reactor 56 h and reactor 72 h.

To produce more artUPO$_{yeast}$ in one fermentation, the process was scaled up to 2 L and following the same procedures as described previously but in a 5 L fermentation vessel. The 2 L sample was concentrated to a more manageable volume of 200 mL

using a tangential diafiltration unit. The unit also contained a 30 kDa filter which enriched the UPO supernatant.

8.2.1.7. Purification of artUPO$_{yeast}$ from *P. pastoris* in shake flasks

Nickel affinity chromatography (NiNTA) can be used to purify proteins that possess a His-tag. The pH of the artUPO$_{yeast}$ sample was adjusted directly using acid or base to pH 8.0. The sample was loaded onto the nickel column with a small amount of imidazole to reduce non-specific binding of other contaminants to the column. As the concentration of imidazole increased over the run, artUPO$_{yeast}$ was eluted due to the increasing concentration of imidazole which out competed artUPO$_{yeast}$ binding for the immobilised nickel. The purification was monitored using UV light at 280 nm as aromatic amino acids in proteins absorb UV at this wavelength. artUPO$_{yeast}$ was found in the fractions highlighted in Figure 8.13 at a molecular weight of roughly 45 kDa.

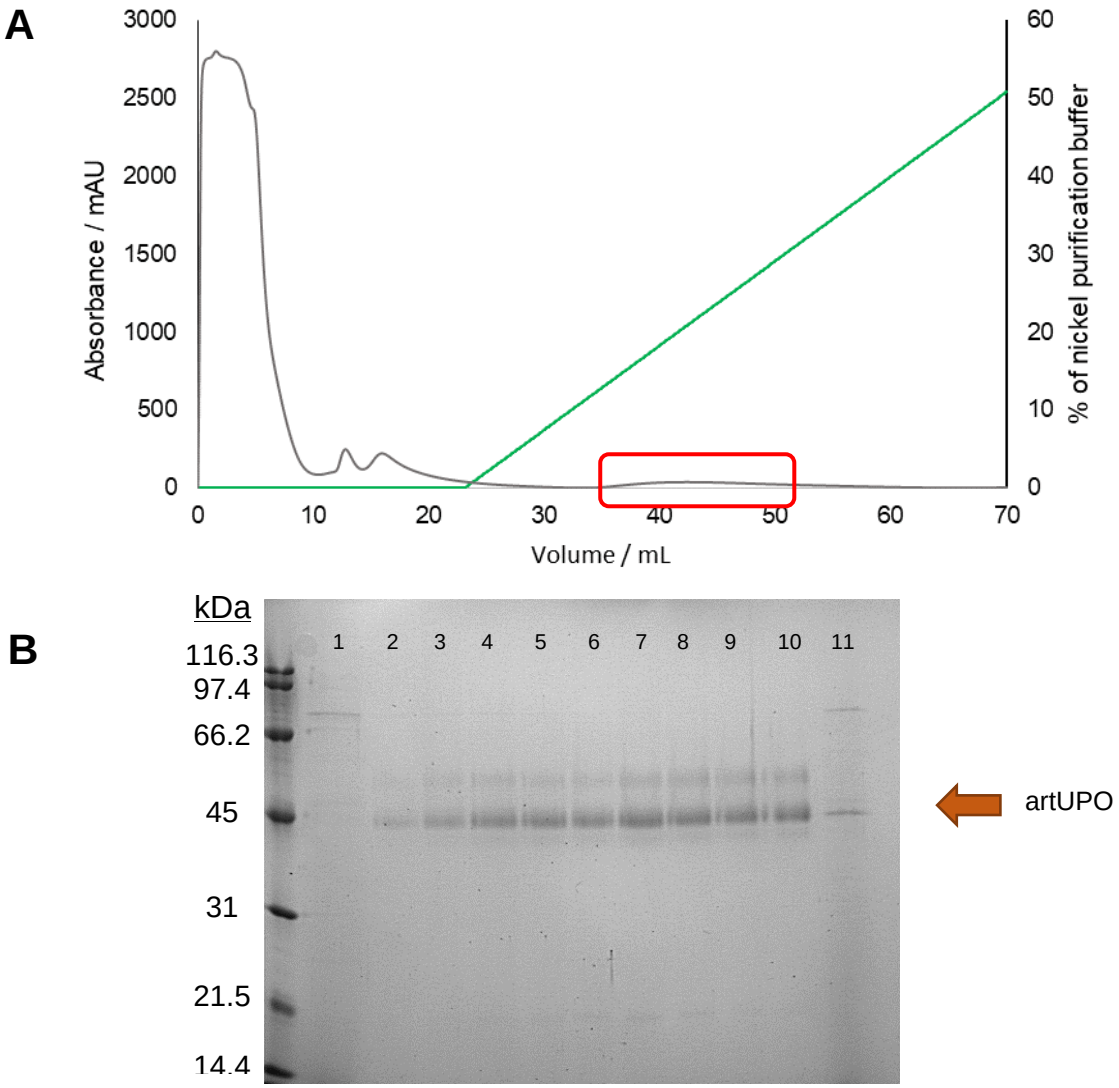


Figure 8.13 NiNTA purification of artUPO$_{yeast}$. **A**. FPLC trace of the purification where the grey line shows the absorbance at 280 nm and the green line of the % of the nickel buffer B. **B**. SDS-PAGE of the fractions (highlighted in red in A) where lane 1 is the flow-through collected during loading of the column, lane 11 is the flow through collected during the purification and lanes 2 – 10 are the fractions of interest containing artUPO$_{yeast}$. The bands corresponding to artUPO$_{yeast}$ are marked next to the gel.

To purify the protein further, size exclusion chromatography (SEC) was used (Figure 8.14). The protein and any impurities will elute in size order with the largest eluting first. Because the fractions corresponding to artUPO were too weak to observe on an SDS-PAGE gel, these fractions were pooled together and concentrated to analyse on

an SDS-PAGE gel (Figure 8.14B). From this new enriched sample, a stronger band was visible at 45 kDa, indicating the presence of artUPO$_{yeast}$. This concentrated fraction also tested positive for peroxidative and peroxygenative activity using the coloured activity assays turning from colourless to blue and yellow respectively. The yield of the 0.6 L shake flask expression was 0.63 mg L^{-1} using a Bradford assay to determine the concentration. As observed in the western blots in Figure 8.8 and 8.9, two bands are observed corresponding to UPO which suggests the presence of two glycoforms.

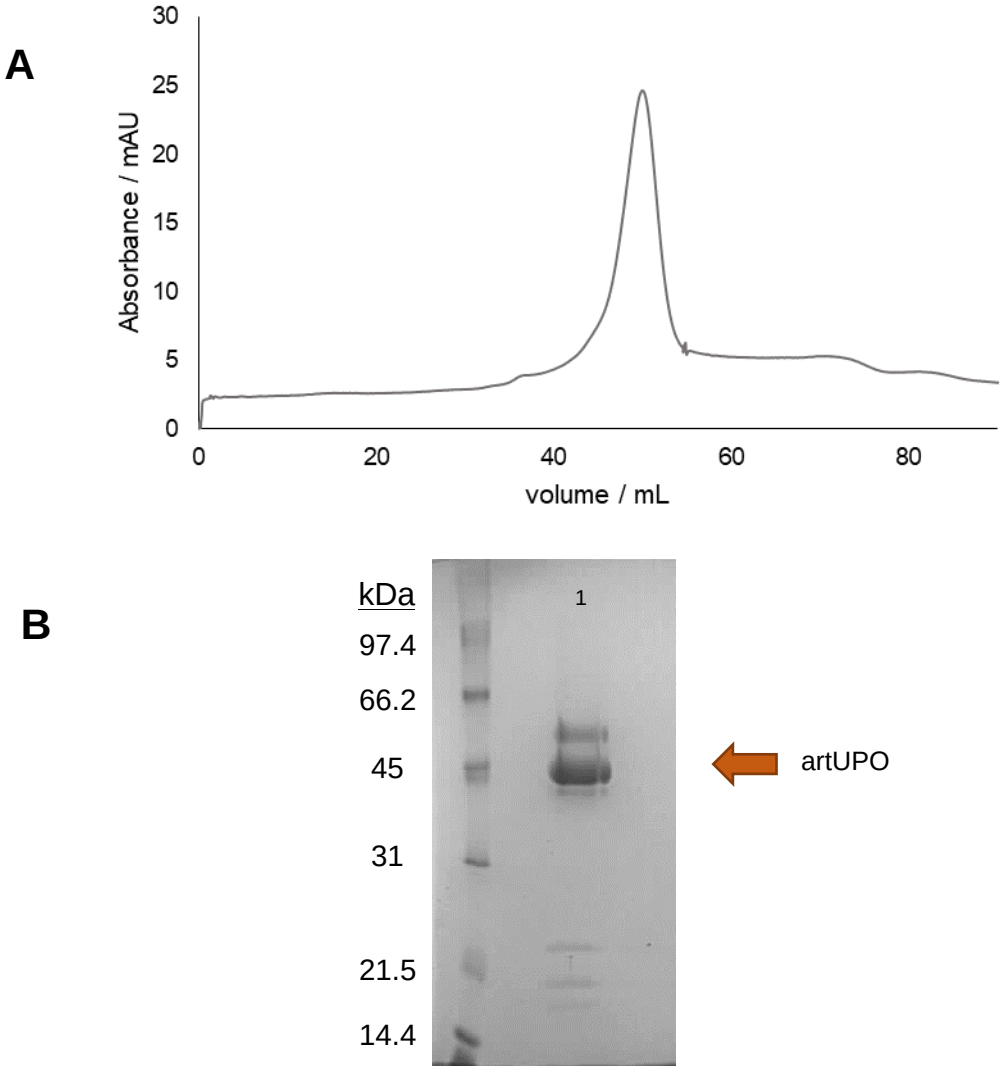


Figure 8.14 SEC purification of artUPO$_{yeast}$. **A**. FPLC chromatogram of purification. **B**. SDS-PAGE gel of the concentrated fractions in lane 1 (highlighted in **A**). The bands corresponding to artUPO$_{yeast}$ (45 kDa) are marked next to the gel.

8.2.1.8. Purifcation of artUPO$_{yeast}$ from *P. pastoris* in a batch fermentation

The pH of the fermentation culture was adjusted either with the addition of acid/base or *via* dialysis and was loaded onto a nickel affinity column but no binding was observed. This suggested that there was a problem with the His-tag on the protein as expression was shown to be successful previously (see Figure 8.11). A Western blot was run using the time course samples of the fermentation as shown in Figure 8.15.

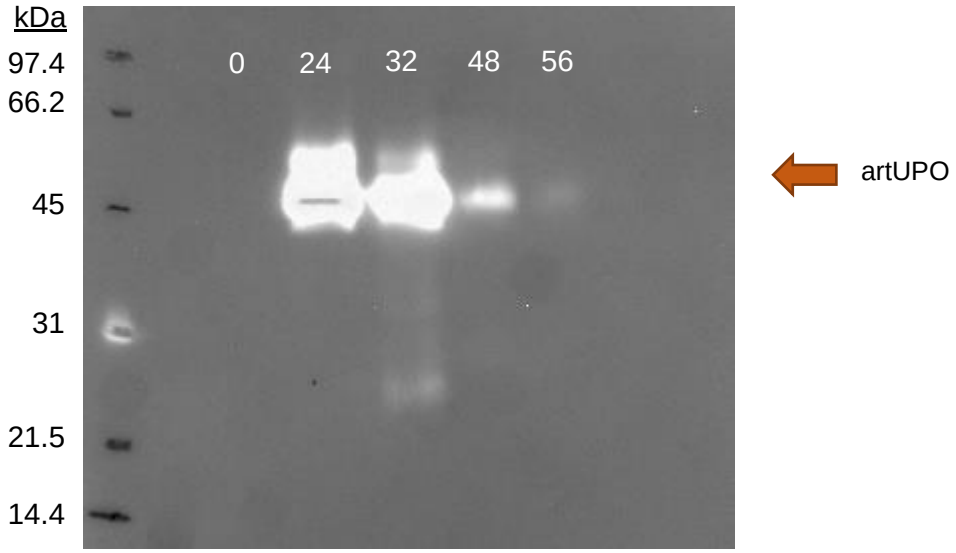


Figure 8.15 Western blot of the fermentation time course for artUPO$_{yeast}$, which is marked next to the gel

The signal for artUPO$_{yest}$ appeared to be strongest at 24 h but this signal degrades over time, but the SDS-gel in Figure 8.11 shows that artUPO$_{yeast}$ concentration increases over time. These images together suggested that the protein itself did not degrade but the His-tag was degraded or removed. This may be due to *P. pastoris* because it secretes some proteases along with artUPO$_{yeast}$. These proteases may be able to cleave the protein at the HRV-3C protease cleavage site and so artUPO$_{yeast}$ loses its His-tag.

To check that the artUPO$_{yeast}$ from the fermentation was active, colourimetric activity assays were carried out on the samples that had been pH adjusted directly or were dialysed using ABTS **118a** or NBD **119a**. These were run with a negative control with no enzymes and two positive controls, with recombinant *Aae*UPO (will now be called r*Aae*UPO) and artUPO$_{yeast}$ after dialysis (Figure 8.16). The samples showed a positive result with a colour change and so it was deemed that the enzyme was still active.

Figure 8.16 Coloured activity assays with the peroxidative activity assay with ABTS **118a** on the top row and the peroxygenative activity with NBD **119a** on the bottom row. The numbers correspond to the samples as follows: 1. Negative control 2. r*Aae*UPO, 3. artUPO$_{yeast}$ with direct pH adjustment 4. artUPO$_{yeast}$ after dialysis 5. artUPO$_{yeast}$ in crude fermentation culture 6. Purified artUPO$_{yeast}$ from shake flask expression

With these pieces of evidence that the His-tag was removed, a different approach was needed to purify the protein from the crude culture. Anion exchange (Q) chromatography is a method that purifies a protein though binding *via* electrostatic charge, rather than binding with a His-tag. To use this method, the artUPO$_{yeast}$ samples were dialysed into a salt-free buffer or passed through a desalting column, and loaded onto a Q column. Q chromatography was run with an increasing concentration of NaCl from 0 mM to 500 mM over 20 column volumes (CV). The protein eluted as the salt concentration increases because the higher salt concentrations out compete the artUPO$_{yeast}$ binding for the stationary phase in the Q column. The purification was monitored using UV light at 280 nm and artUPO$_{yeast}$ was found in the fractions highlighted in Figure 8.17. The NaCl concentration after the gradient increased to 2 M

to remove any bound impurities. The UV peak seen immediately after this increase was most likely to be residual DNA in the supernatant.

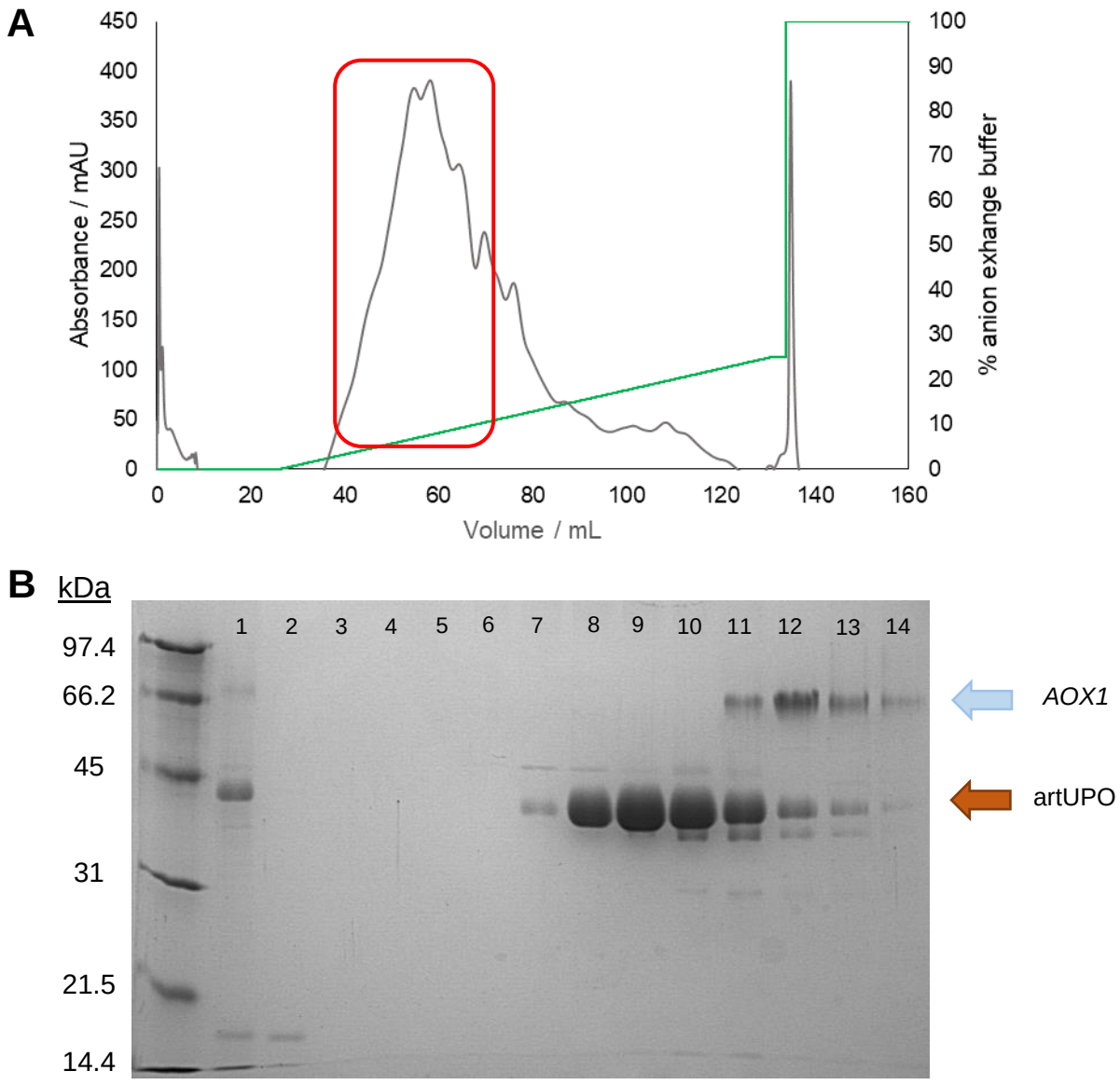


Figure 8.17 Q purification of artUPO$_{yeast}$ **A.** FPLC trace of the purification where the grey line shows the absorbance at 280 nm and the green line of the % of the anion exchange buffer. **B.** SDS-PAGE of the fractions (highlighted in red in **A**) where lane 1 is a sample taken used to load the column, lane 2 is the flow through collected after loading the protein on the column, lane 3 is the flow though collected during purification and lanes 7 – 13 are the fractions of interest containing artUPO$_{yeast}$. On an SDS PAGE gel, artUPO and *AOX1* appear at 45 kDa and 66 kDa respectively and are marked next to the gel

Lanes 7 – 10 (Figure 8.17) were pooled together and concentrated for SEC (Figure 8.18). It was expected that the suspected *AOX1* would be separated with SEC because *AOX1* is larger than artUPO$_{yeast}$. This was not observed and both proteins were eluted together.

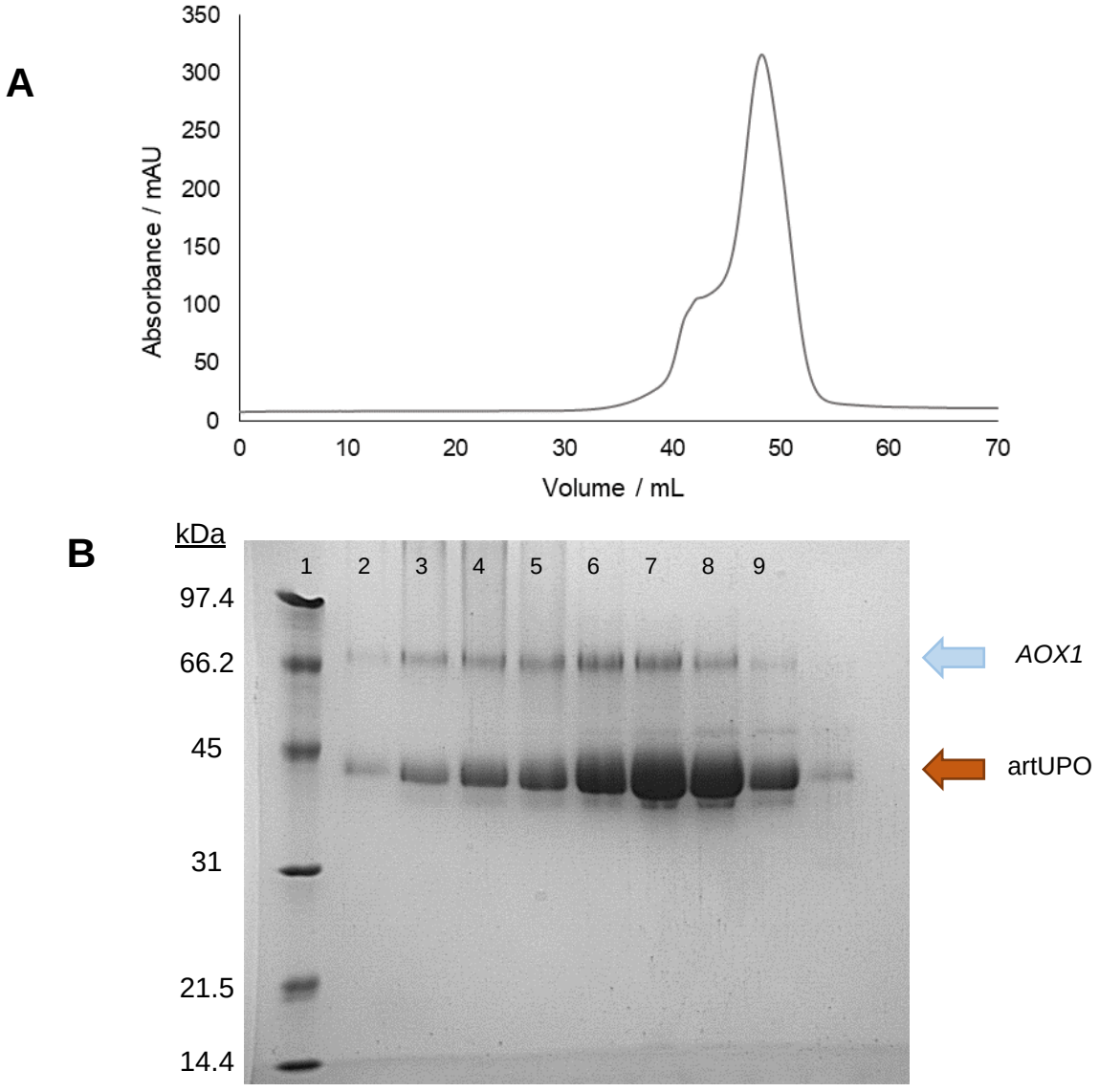


Figure 8.18 SEC of artUPO$_{yeast}$ derived from Q purification. **A**. FPLC chromatogram of purification. **B**. SDS-PAGE gel of the fractions highlighted in **A** where lanes 1 – 9 contained artUPO. The bands corresponding to artUPO$_{yeast}$ (45 kDa) and *AOX1* (66 kDa) are marked next to the gel

To improve the method and separate the AOX1 from artUPO$_{yeast}$, the Q chromatography method was changed to a slower gradient from 0 mM – 500 mM to 0 mM – 300 mM over the same number of column volumes. This improved the separation of the two proteins for a new sample of fermentation culture (Figure 8.19) and the fractions that contained only artUPO$_{yeast}$ were pooled together and concentrated for SEC (Figure 8.20).

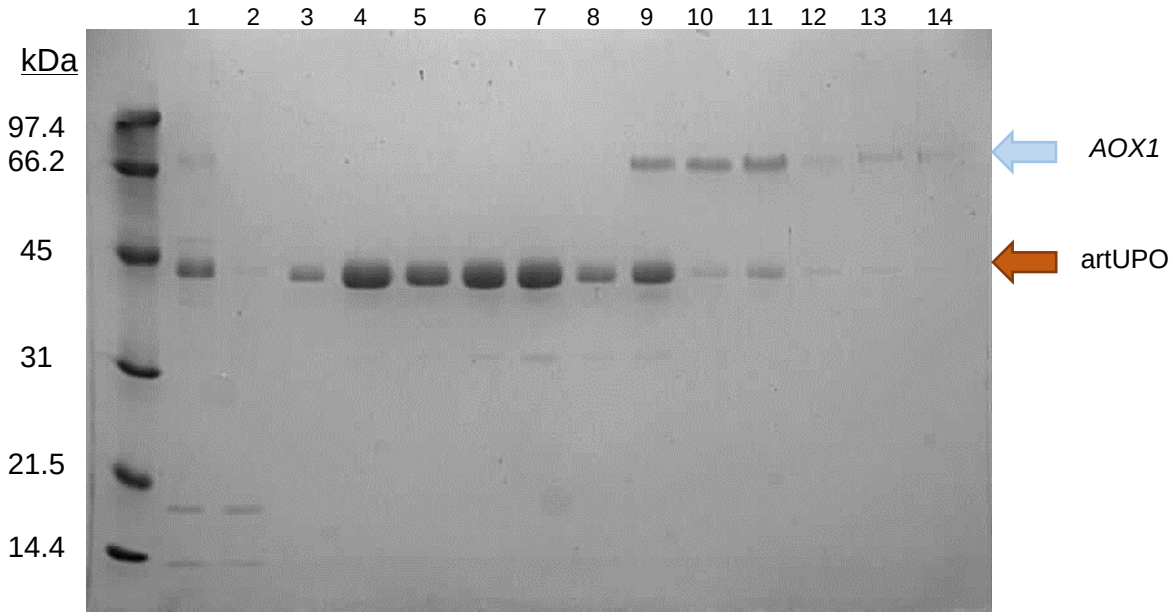


Figure 8.19 SDS-PAGE gel of improved Q purification of artUPO$_{yeast}$ from fermentation. Lane 1 is the sample of the loading material onto the column, lane 2 was the flow through collected after loading the column and lanes 3 – 8 were pooled together for SEC. The bands corresponding to artUPO$_{yeast}$ (45 kDa) and *AOX1* (66 kDa) are marked next to the gel

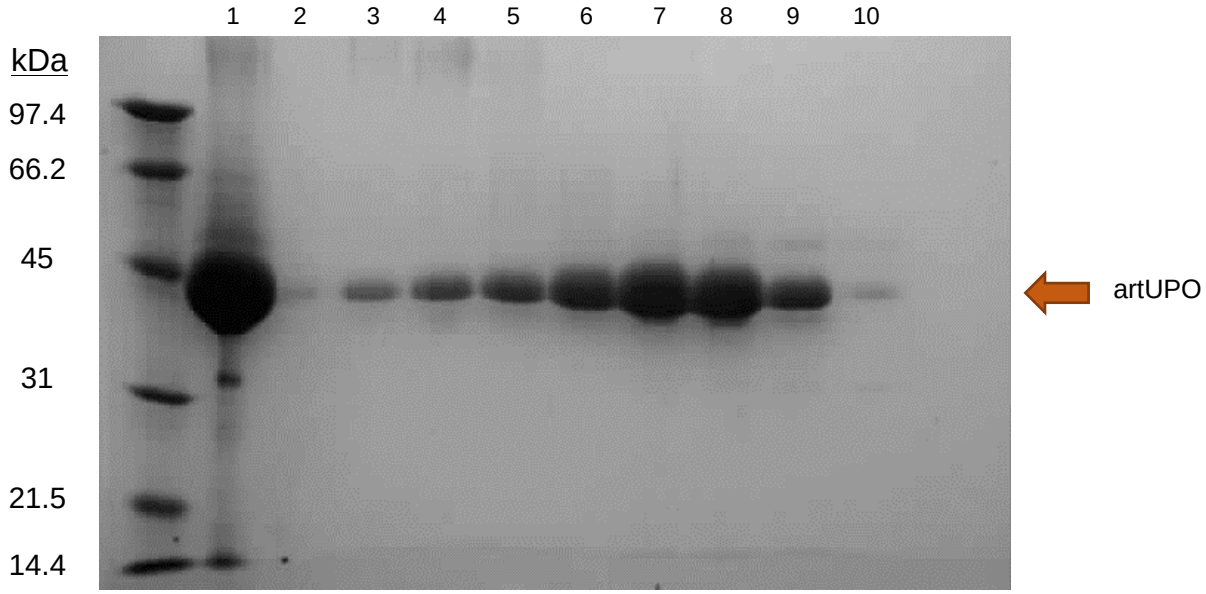


Figure 8.20 SDS-PAGE gel of SEC of artUPO$_{yeast}$ from improved Q purification. Lane 1 was the loading sample and lanes 2 – 10 contained artUPO$_{yeast}$. The bands corresponding to artUPO$_{yeast}$ (45 kDa) are marked next to the gel

The overall yield of the fermentation was 98 mg L^{-1} using a Bradford assay to determine the concentration of protein.

8.2.2. Expression and purification of artUPO in *E. coli*

8.2.2.1. Expression tests

The artUPO gene for the expression in *E. coli* was purchased from GenScript in a pET-28a(+) vector as shown in Figure 8.21. The gene was codon optimised for the expression in bacteria.

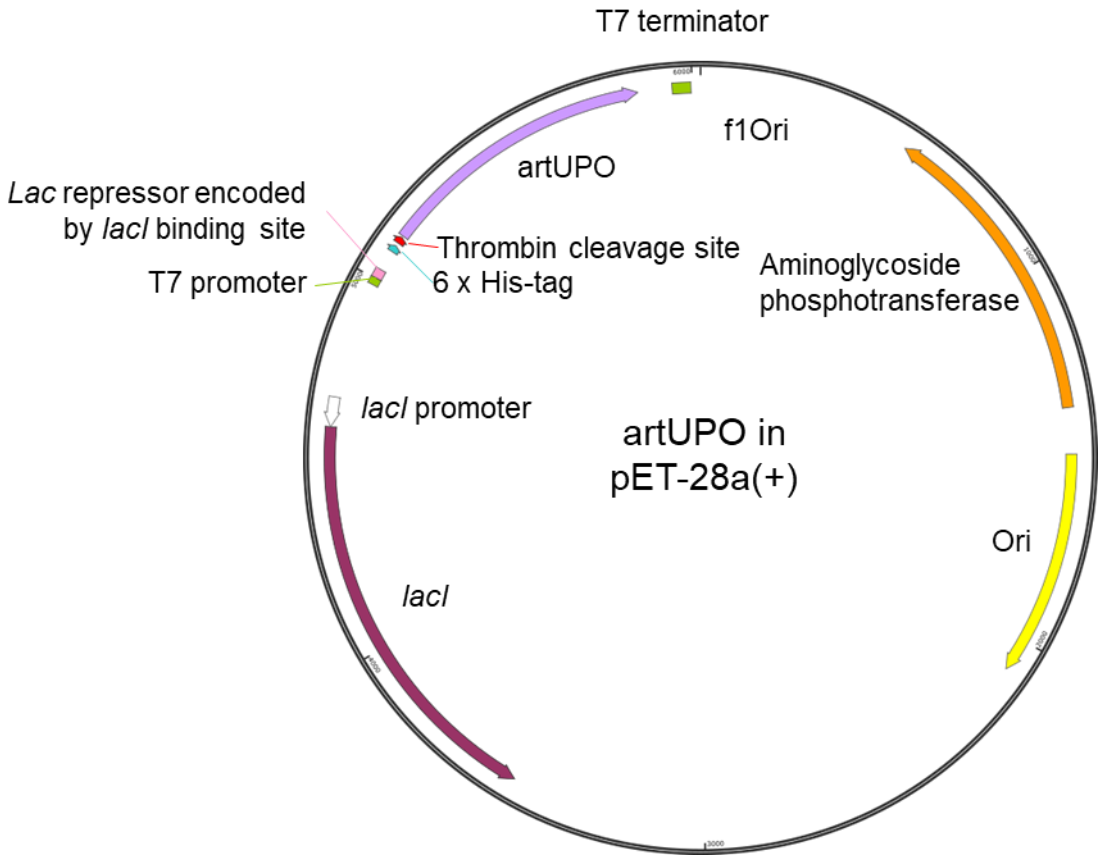


Figure 8.21 Map of expression vector pET-28a(+) containing artUPO gene. The expression of artUPO is controlled by the *lac* operon (maroon). A 6 x His-tag (blue) and a thrombin cleavage site (red) are placed at the N-termninal of the protein. The codon optimised gene for artUPO is shown in purple. Resistance for kanamycin is conferred by the aminoglycoside phosphotransferase gene (orange). The origin of replication (ori) is shown in yellow, which is the site of DNA replication and enables the plasmid to reproduce itself.

Like in the pPICZ plasmid, an *N*-terminal 6 × His-tag was introduced but a thrombin cleavage site was used instead. The expression of artUPO is under the control of a T7 polymerase promotor. This system is based on the *lac* operon which controls the switch from glucose to lactose metabolism in *E. coli*. This change is mediated by the binding of allolactose to the lac repressor (*lacI*) which inhibits its binding to target sequences. In the DE3 *E. coli* expression systems used the gene encoding T7 polymerase is present under the control of the *lac* promotor. T7 polymerase expression is normally inhibited by *lacI*, which binds to the *lac* promotor and prevents

binding of *E. coli* RNA polymerase. Expression can be induced by the addition of IPTG to the expression media. IPTG acts as a allolactose mimic and is able to bind to *lacI* which prevents inhibition of T7 polymerase expression. This allows for expression of T7 polymerase which in turn leads to expression of the gene of interest. [156]

The plasmid was used to transform three different *E. coli* strains: BL21 (DE3), Rosetta PLysS and B834. All three strains grew colonies and single colonies from each strain were picked, added to LB and gene expression was induced after the addition of IPTG, 5-aminolevulinic acid (5-ALA) and $FeSO_4$. The addition of 5-ALA and $FeSO_4$ has shown to aid protein production of heme-thiolate proteins when induction occurs. [157] This is because these compounds are precursors needed to make a prosthetic heme group for an active enzyme. *E. coli* does not naturally produce enough of these compounds and so a supplement is necessary for the production of artUPO$_{bact}$. After an overnight incubation at different temperatures, the cultures were spun 10000 × g and the soluble and insoluble fractions were separated and analysed as shown in Figure 8.22.

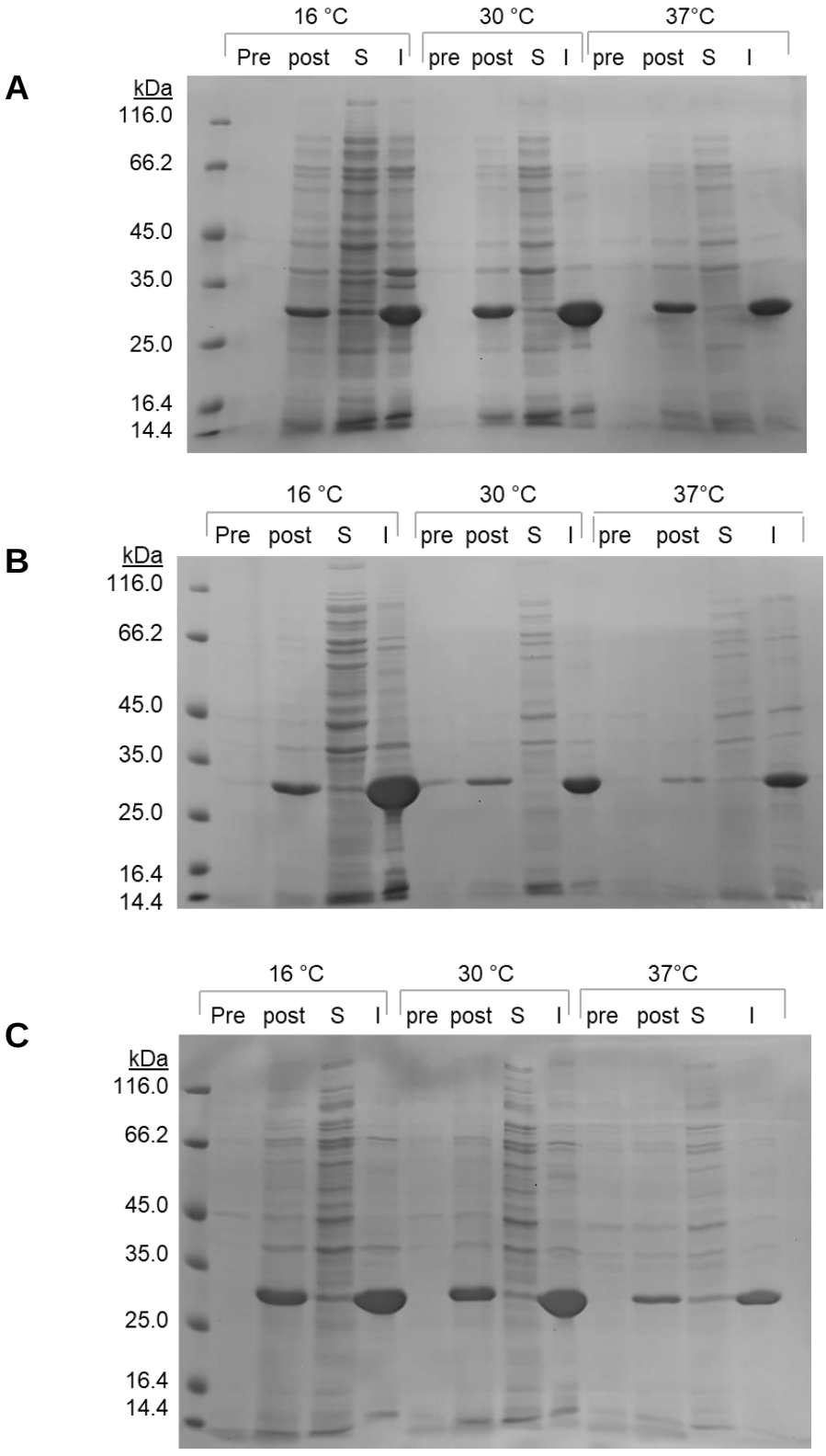


Figure 8.22 SDS-PAGE gels of expression tests in different strains of *E. coli* at different temperatures post induction where Pre = pre induction, post = post induction, S = soluble, I = insoluble. **A** – BL21 DE3 expression test, **B** – Rosetta pLyS expression test, **C** – B834 expression test

To monitor the cell growth, samples of the culture were taken and an OD_{600} reading was taken from the samples to measure the cell density. Once the OD_{600} values were in the range of 0.6 – 0.8, the culture could be induced. A lower temperature for incubation after gene expression would reduce the rate of cell growth and would prolong the time in which the cells are in the stationary phase of growth which is optimal for protein production. Higher temperatures would increase the rate at which the cells initially grow but will then increase the rate of cell death. With less active cells, the yield of the desired protein would decrease. Also most proteins are stable at 16 °C for long periods of time.

In Figure 8.22 where at higher temperatures in all the strains, a band corresponding to $artUPO_{bact}$ at 32 kDa is observed in the post induction sample but is also seen in the insoluble fraction where the protein is likely to be unfolded and has remained in the cells. At lower temperatures, $artUPO_{bact}$ can be observed in the soluble fractions where the protein is folded and can be isolated after cell lysis. As seen in Figure 8.22, $artUPO_{bact}$ can be observed in all of the soluble fractions of each strain but has the strongest band in the soluble fraction of Rosetta when incubated at 16 °C, which suggests at there is significant overexpression of $artUPO_{bact}$. This condition was selected to carry forward and scale up the culture.

8.2.2.2. 1 L shake flask culture

Rosetta and the respective suitable conditions were selected from the expression tests and the protein production was scaled up to a 1 L shake flask. An overnight culture was grown, added to 1 L LB, and was grown until an appropriate OD_{600} value was reached. The gene expression was induced as described in the methods and was

incubated overnight at 16 °C at 180 rpm. To lyse the cells, a cell disruptor was used instead of sonicating as before because the cell disruptor was able to lyse cells more consistently. The lysate was spun and the supernatant could be used for the next stage of purification. If more protein was needed, more 1 L shake flask cultures were prepared and grown at the same time.

8.2.2.3. Purification of artUPO$_{bact}$ from *E. coli* in shake flasks

To purify artUPO$_{bact}$ from the supernatant described above, a nickel affinity chromatography purification was carried out and the FPLC trace with the resulting SDS PAGE gel is shown in Figure 8.23.

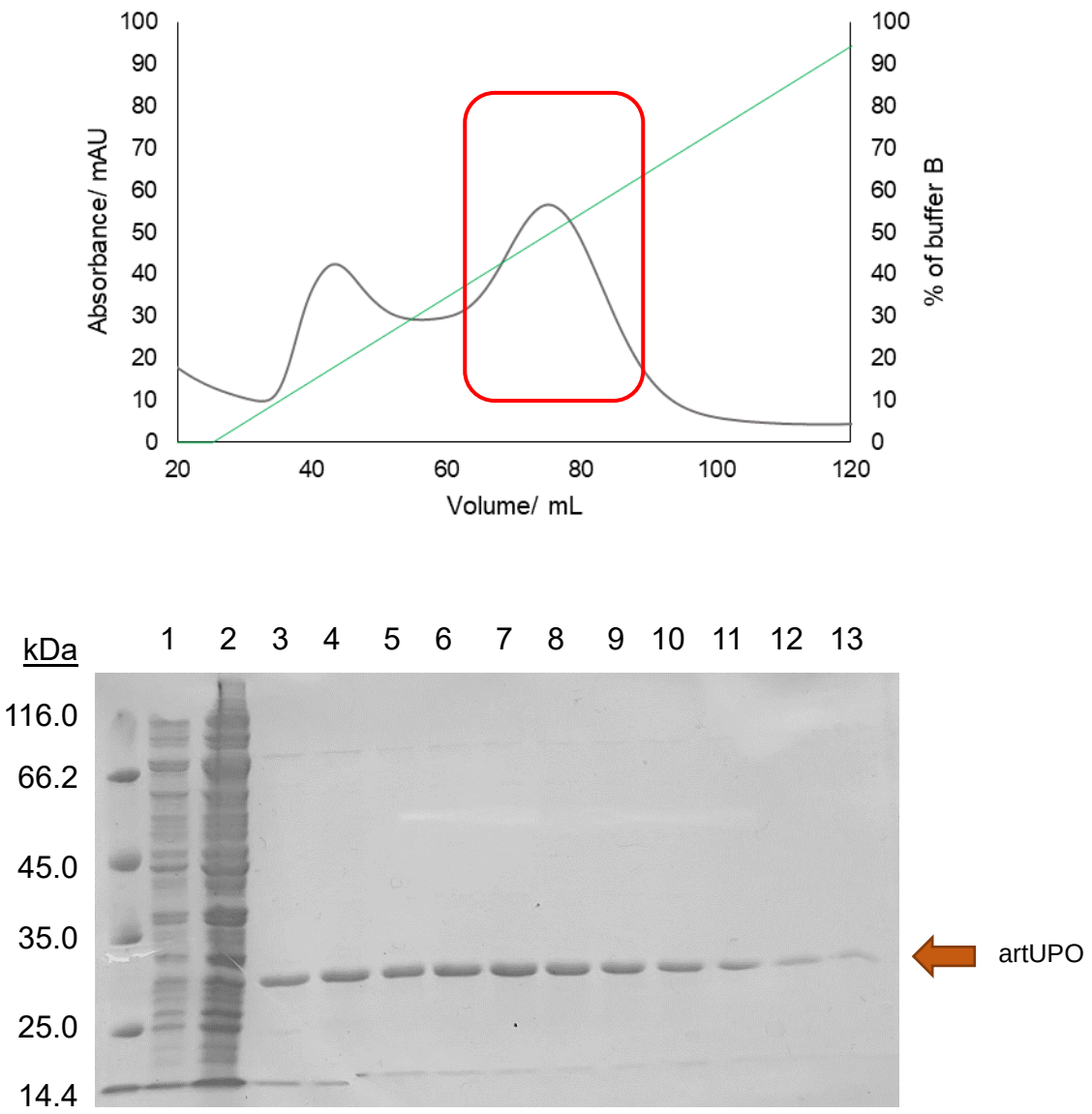


Figure 8.23. NiNTA purification of artUPO$_{bact}$ **A.** FPLC trace of the purification where the grey line shows the absorbance at 280 nm and the green line of the % of nickel buffer B. **B.** SDS-PAGE of the fractions (highlighted in red in **A**) where lane 1 is a sample taken used to load the column, lane 2 is the flow through collected after loading the protein on the column and lanes 3 – 13 are the fractions of interest containing artUPO$_{bact}$ which are marked next to the gel. artUPO$_{bac}$ on an SDS PAGE appears at 32 kDa

Two peaks appeared in Figure 8.23A at 40 mL and 75 mL. The first peak eluting at 40 was analysed using SDS PAGE but no protein bands were observed (not shown). The second peak eluting at 75 mL resulted in the SDS PAGE gel seen in Figure 8.23B.

Lanes 3-13 were pooled together and concentrated to load onto a SEC column (Figure 8.24).

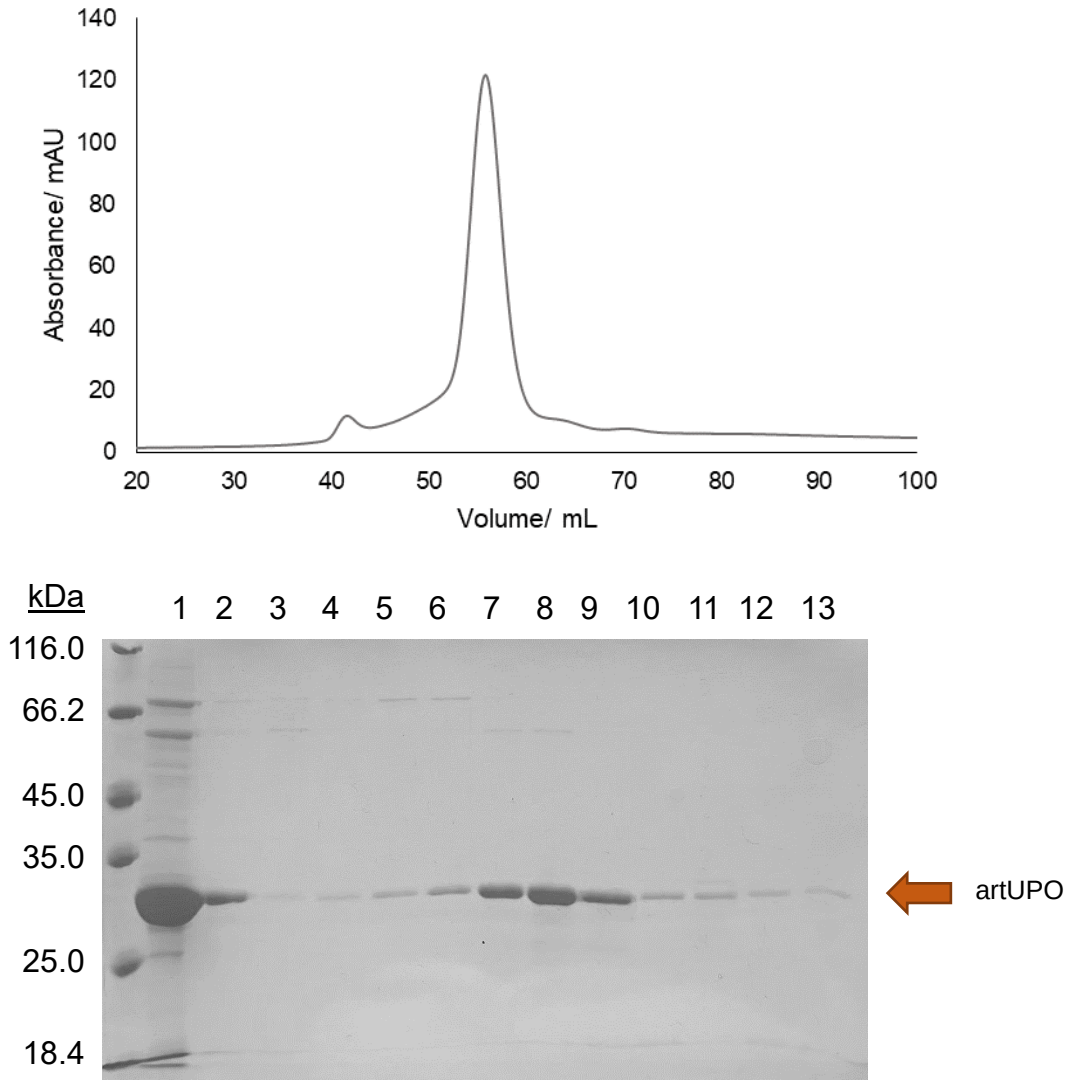


Figure 8.24. SEC of artUPO$_{bact}$ derived from NiNTA purification. **A**. FPLC chromatogram of purification. **B**. SDS-PAGE gel of the fractions highlighted in **A** where lane 1 is the load, lane two is the peak observed at 40 mL and lanes 3 – 13 were collected and contained artUPO$_{bact}$ which are marked next to the gel

A small peak in the SEC UV trace (Figure 8.24) was observed at 40 mL into the elution and appeared at a mass of 32 kDa in the SDS PAGE gel which suggests this is artUPO$_{bact}$. This peak appears in the void volume of the column which suggests this species must be very large so this peak could be an aggregate form of artUPO$_{bact}$. The

aggregate was not desired and the artUPO$_{bact}$ in lanes 3 -13 were collected and stored for further use. The overall yield for the 1 L shake expression was 0.91 mg L^{-1} and the concentration was determined by a Bradford assay.

8.2.3. Deglycosylation of artUPO$_{yeast}$ using PNGase F

PNGase F is an enzyme that is able to remove *N*-linked oligosaccharides from glycoproteins and has a molecular weight of of 36 kDa. After incubation of purified artUPO$_{yeast}$ with PNGase F, several bands are observed where PNGase F has digested the sugars on artUPO$_{yeast}$ to different extents (Figure 8.25). The species with the lowest molecular weight in the PNGase F treatment was 31 kDa; this corresponds with the molecular weight of artUPO$_{yeast}$ without glycosylation. From this SDS PAGE gel, the extent of glycosylation of artUPO$_{yeast}$ was calculated to be around 30% which differs to the value of 16% reported for the wild type *Mro*UPO expressed in *Marasmius rotula*. [158] The extent of glycosylation will vary in a protein depending on the organism expressing the protein and so differences in the amount of glycosylation were expected. The percentage calculated from the SDS PAGE gel will be somewhat inaccurate because the glycosylated forms of artUPO$_{yeast}$ will migrate at different speeds and so the true mass of these glycoforms may not be represented using this technique. Mass spectrometry experiments were carried out to attempt to resolve this issue as described in Section 8.2.4.

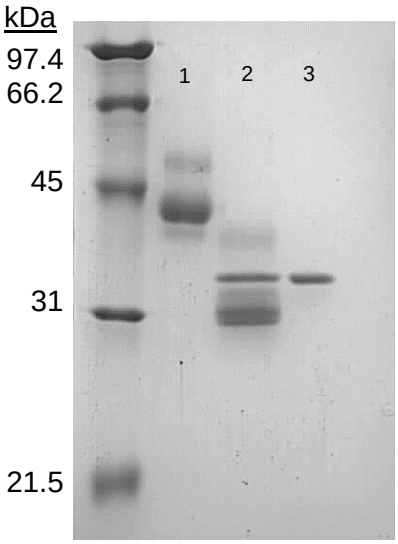


Figure 8.25 SDS-PAGE analysis of PNGase F treatment with artUPO$_{yeast}$ where lane 1 has artUPO$_{yeast}$ only, lane 2 has artUPO$_{yeast}$ and PNGase F incubated together and lane 3 has PNGase F only

8.2.4. Mass spectrometry analysis

A sample of artUPO$_{yeast}$ was submitted for ESI-MS to accurately determine the molecular weight of the protein and to identify any patterns in glycosylation. The following spectrum was generated from this experiment (Figure 8.26).

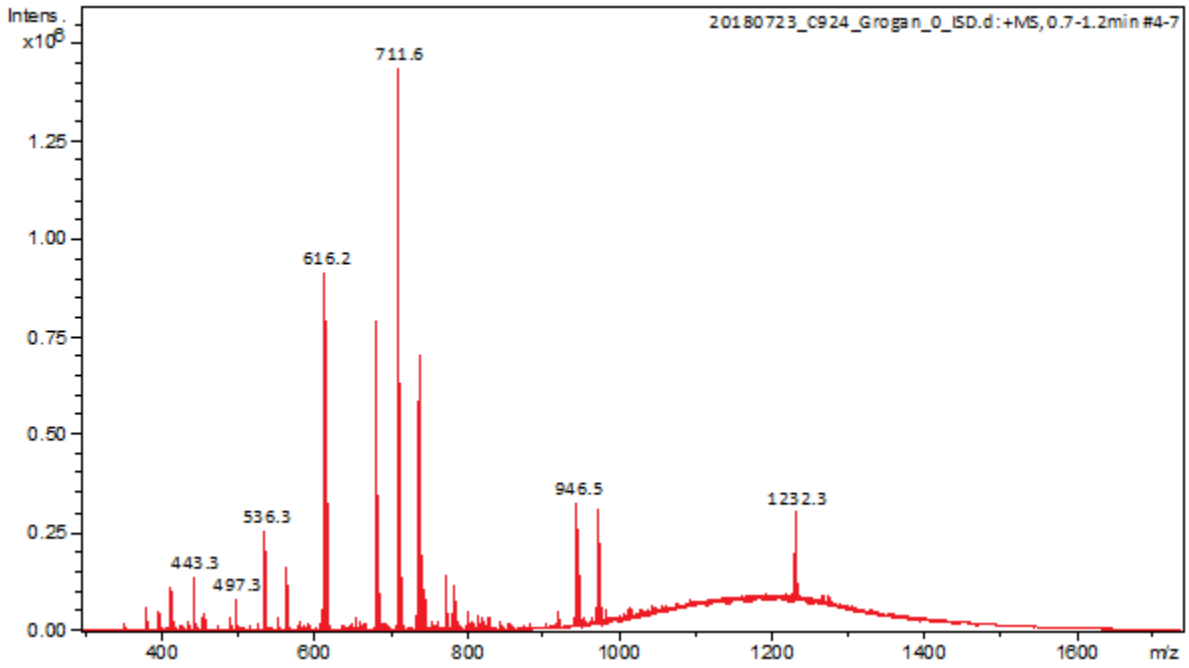


Figure 8.26 ESI-MS spectrum of artUPO$_{yeast}$

The spectrum shows a number of low mass, singly charged species and the broad hump between *m*/*z* 1000-1400 corresponds to the protein. The broadness of the peak may be due to the different glycoforms present in artUPO$_{yeast}$ or the adducts splitting the signal across multiple masses. Because this peak is very broad, it was difficult to resolve individual charge states and deconvolute to the mass of artUPO$_{yeast}$. To improve the spectrum and stabilise the protein signal, the voltage of the source ion funnel was increased from 0 eV to 80 eV where potential adducts may be removed and non-covalent interactions may be disrupted. The change in voltage gave rise to the spectrum as shown in Figure 8.27

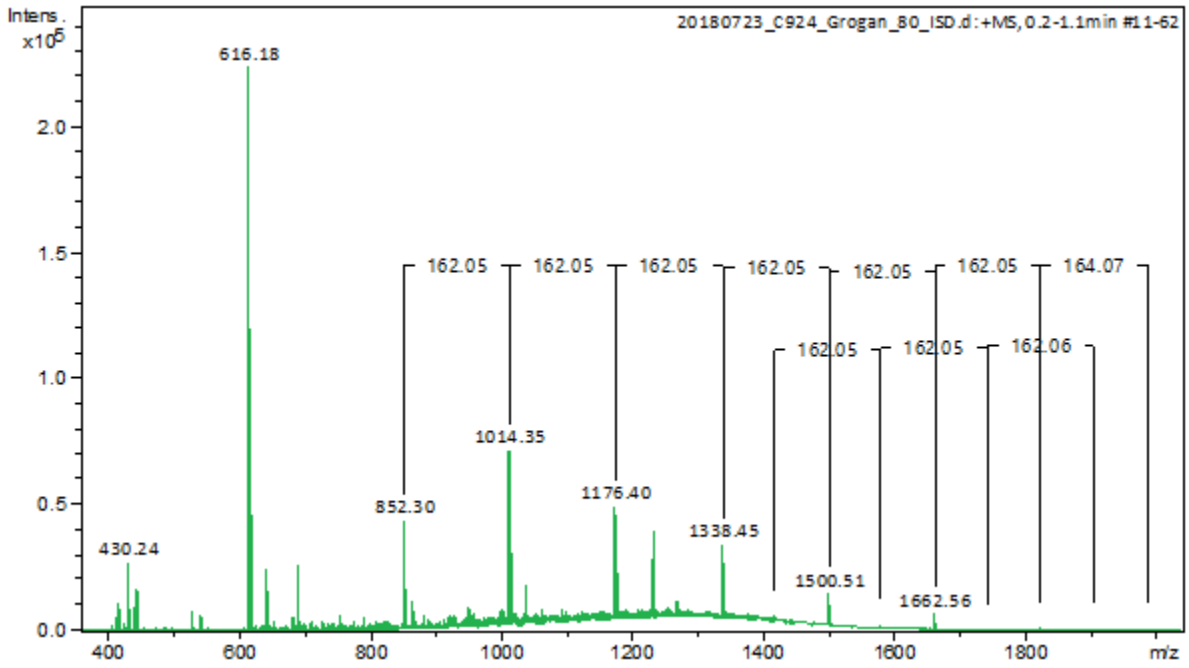


Figure 8.27 ESI-MS 80 eV spectrum of artUPO$_{yeast}$ with two glycan fragmentation patterns highlighted

In this spectrum, new singly charged peaks appeared and corresponded with glycan fragmentation patterns with 162 Da intervals. The difference in 162 Da between each peak indicates the presence of a hexose chain which is expected because *Mro*UPO contains the hexose mannose. The spectrum shows that there are two glycosylation patterns, this is also observed in SDS PAGE gels of artUPO$_{yeast}$ where two bands are seen which corresponds to two different glycoforms. A singly charged species at *m*/*z* 616.18 is observed with high intensity and corresponds to the molecular mass of heme. Unfortunately the mass of the intact artUPO$_{yeast}$ could not be measured from these experiments potentially due to the heterogeneity in the glycoforms of artUPO$_{yeast}$.

8.2.5. UV/vis scan

Both artUPO$_{yeast}$ and artUPO$_{bact}$ were scanned using UV/visible light from 250 – 800 nm and the resulting spectra are shown below overlayed (Figure 8.28).

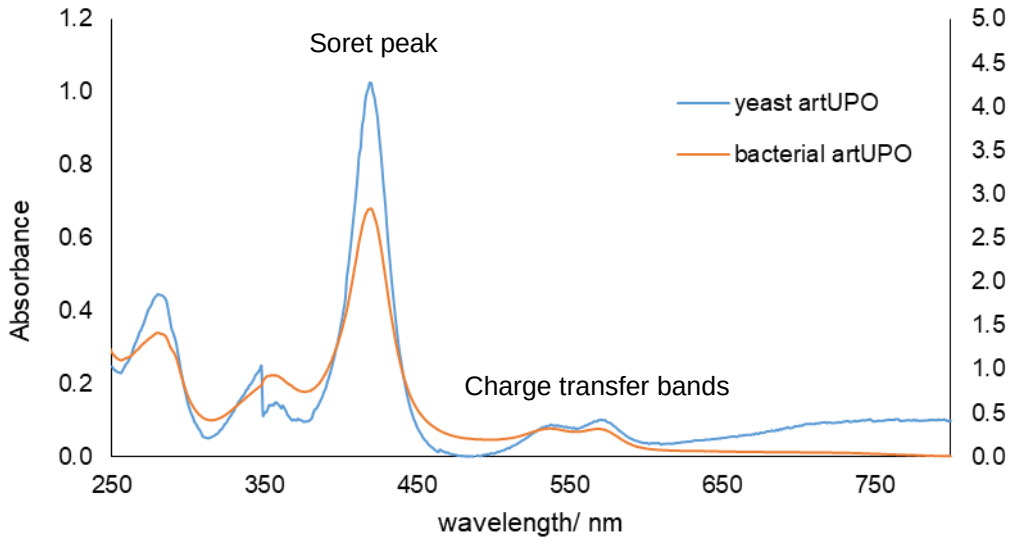


Figure 8.28 UV/vis spectra of both artUPO$_{yest}$ (in blue) and artUPO$_{bact}$ (in orange)

Both spectra share the same basic shape, where the soret peak at 418 nm and the charge transfer bands, derived from π to π* transitions in the porphyrin upon absorption, are present in both samples. [159] These features in the UV spectrum are also present heme-thiolate enzymes such as P450s, *Cfu*CPO and other UPOs. [160, 161, 96, 158] The main difference between the two different artUPO UV spectra are the Rz values which indicate roughly the purity of the sample by taking the ratio between the absorbance at the soret peak to the aborbance at 280 nm (Equation 1). The glycosylated artUPO$_{yeast}$ had a higher Rz value (~3) than the non-glycosylated artUPO$_{bact}$ (~2), suggesting good incorporation of the heme in both proteins and that both artUPOs are pure.

8.2.6. Comparison of expression systems for UPOs for the use in industry

UPOs have received a lot of interest over the last few years due to their catalytic abilities, stability and relative ease of use and so it is desirable to find the simplest system to express these proteins in the highest yields. Until very recently, UPOs were expressed heterologously in yeast such as *S. cerevisiae* or *P. pastoris*, or the fungus *M. oryzae* but had never been successfully expressed in bacteria. [96,98,103] In recent years three UPOs, *Mro*UPO, *Cvi*UPO and *Dca*UPO, were expressed in *E. coli* cells and were found to be active. [108, 109, 110] Now that two heterologous systems have been established in bacteria and yeast, it can be debated if one system is preferable over the other. In this project, artUPO was successfully expressed in both *P. pastoris* and *E. coli*.

There are key differences between the two systems when expressing UPOs. *P. pastoris* is able to carry out post-translational modifications to proteins such as glycosylation. This may be important for UPOs as it may affect its chemical or thermal stability. *E. coli* are unable to carry such post-translational modifications on recombinant proteins and so it was originally believed that they were unable to produce folded and functional UPOs. With the research carried out previously, demonstrating *E. coli* expression of *Mro*UPO, *Cvi*UPO and *Dca*UPO, and this project, it seems that short UPOs may suitable candidates for *E. coli* expression but further investigation would be needed to test more short UPOs to investigate their catalytic capabilities and industrial applicability. [108,110] The effects of glycosylation of artUPO$_{yeast}$ is discussed further in the following chapters.

Another important aspect to discuss is the applicable use of these systems in industry and which method of production is the most cost and time effective. *E. coli* is a very popular organism for recombinant protein expression because the system has been extensively researched and relatively simple to use. One step is required to introduce the plasmid into the host organism and the growth of the culture takes roughly a day in shake flasks when a starter culture is used. When expressing artUPO$_{bact}$, the yield was 0.91 mg mL^{-1}. The cells however need to be lysed to release the protein prior to purification and the culture required the addition of $FeSO_4$ and 5-ALA to produce the heme, which significantly increases the cost. For the case of *P. pastoris*, the cloning step is more complicated because the plasmid must first be grown and linearised before introducing the DNA into *P. pastoris* cells. To save time on this step, a glycerol stock can be created of the transformed *P. pastoris*. Like *E. coli* a starter culture must be established before growing either in a shake flask or fermenter. *P. pastoris* has a longer doubling time than *E. coli* (3 h compared to 30 min) and so growth to suitable OD_{600} values take longer to reach. In a fermentation, higher cell densities can be reached compared to growth in a shake flask as shown in this project where an increase in yield of artUPO$_{yeast}$ production is observed from 0.63 mg L^{-1} in shake flasks to 98 mg L^{-1} in a fermenter. *P. pastoris* in particular can reach very high cell densities in fermentation conditions and has been used to produce a variety of recombinant proteins for this reason. [162] The high yields seen are also due to the mutated signalling peptide developed by Molina-Espeja *et al*. 2014 in the pPICZ plasmid. [96] *E. coli* can also be fermented but this was not carried out in this project. Carrying out an *E. coli* fermentation for the production of artUPO$_{bact}$ would be very useful to make a fair comparison between the two hosts used in this study, as time versus yield could be evaluated. A direct comparison can be made between both organisms when shake

flasks are used. On this scale, *E. coli* appears to be a better host because it produces a higher yield of artUPO within a shorter amount of time. However, it is important to consider other factors when choosing an expression system other than yield, such as the activity and stability of the enzyme which is discussed in later chapters. Because artUPO$_{yeast}$ is secreted into the culture, the process of obtaining protein is more simple than that of *E. coli*, although the his-tag on artUPO$_{bact}$ is not cleaved in the growth of *E. coli* and so the purification is easier. From discussing these points, it is clear to see that comparing the expression systems of *E. coli* and *P. pastoris* to produce artUPO is more nuanced than it appears, however a judgement can be easier to make when the activity and stability of each artUPO is investigated, which was carried out in later chapters.

8.3. Summary of expression and purification experiments

In this chapter, an artificial UPO gene (artUPO), derived from *Mro*UPO from the organism *Marasmius routla*, was cloned *via* InFusion cloning into the pPICZ vector for expression in *Pichia pastoris* or purchased in a pET28-a(+) vector for expression in *Escherichia coli* where in the latter case, the gene was codon optimised for *E. coli* expression.

The pPICZ vector has been optimised by the Alcade group and our group previously for the expression of UPOs in *P. pastoris*; the α-factor and *c-myc* sequences were removed, four point mutations were introduced to the signalling peptide for optimum secretion, and InFusion overhangs were added to easily integrate new UPO genes into the plasmid. [96, 98, 95] The artUPO and r*Cci*UPO genes were cloned into pPICZ and were used to transform stellar cells. Colony PCR and sequencing were used to determine the success of the cloning. The plasmids were then linearised using *Sac*I and were used to transform *P. pastoris* X-33 cells *via* electroporation. Expression tests were carried out to determine the success of the electroporation and for how long the cells should be incubated after induction of the gene using methanol in a shake flask. Western blots revealed that a 72 h incubation would be the optimum expression time for artUPO$_{yeast}$ but expression of r*Cci*UPO was not successful. For artUPO$_{yeast}$ the process was scaled up to a 0.6 L shake flask growth and protein was purified using NiNTA. The resulting yield was 0.63 mg L^{-1}.

To increase the yield of artUPO$_{yeast}$, fermentations in bioreactors at a 0.2 L and 2 L scale were carried out and optimised. A western blot showed however that the His-tag

seems to disappear significantly after 48 h, potentially due to the presence of proteases released in the media, and so NiNTA purification was not suitable to purify artUPO$_{yeast}$ from the fermentation. Instead anion exchange (Q) chromatography was used and the protein was further purified with size exclusion chromatography (SEC). The yield of the fed-batch fermentations were 98 mg L^{-1}.

A pET28a-(+) plasmid was purchased containing the artUPO gene and expression tests were carried out to determine the best strain of *E. coli* to use and the optimum conditions required to produce soluble protein. Rosetta pLysS cells were chosen as the best strain and incubating the cells at 16 °C overnight after induction of the gene seemed to yield the most artUPO$_{bact}$. To aid artUPO$_{bact}$ production, $FeSO_4$ and 5-ALA were added at the time of induction. These conditions were scaled up to 1 L shake flasks and artUPO$_{bact}$ was purified using NiNTA followed by SEC which resulted in a yield of 0.91 mg L^{-1}.

Treatment of artUPO$_{yeast}$ with PNGase F showed that artUPO$_{yeast}$ has roughly 30% *N*-linked glycosylation. For artUPO$_{yeast}$, mass spectrometry experiments confirmed the presence of two glycoforms but could not clearly identify the molecular mass of either of these forms. The UV/vis spectrum showed that both artUPO contained key features of UPOs such as a soret bond corresponding to the heme group and charge transfer bands. The Reinheitszahl (Rz) value of both artUPOs indicated good heme incorporation and that both samples from protein production were pure. With artUPO from both bacterial and yeast expression systems successfully produced and purified, characterisation of the structure and the activity of the enzymes could then be achieved.

9. Characterisation of artUPO from yeast and bacteria

With purified artUPO$_{yeast}$ and artUPO$_{bact}$, investigation of the biochemical and structural characteristics of these proteins could be carried out. This allowed identification of any significant differences between artUPO$_{yeast}$ and artUPO$_{bact}$ and allowed for comparison to the other UPOs. Kinetic parameters of both the peroxigenative and peroxidative activity of both artUPO$_{yeast}$ and artUPO$_{bact}$ were determined, which allowed for a comparison of the enzymes to each other and to other well characterised UPOs. *Mro*UPO, that artUPO was derived from, has previously been shown to exist as a dimer in solution. [128] It was therefore worthwhile to investigate the oligomeric state of artUPO to see if the mutations had disrupted the oligomeric state of the protein. Analytical SEC and treatment with reducing agent experiments were carried out to determine the oligomeric state of artUPO. The thermal stability of both artUPOs was investigated using nanoDSF to see if differences such as glycosylation affected the thermodynamic stability of the enzyme.

The structures of both artUPO$_{bact}$ and artUPO$_{yeast}$ were determined using protein x-ray crystallography. Determination of the structure provided insights into the substrate specificity of artUPO, and could be used to develop the understanding of the general substrate specificity of UPOs. These structures may offer a significant contribution to the relatively small number of UPO structures currently available on the PDB (4 unique proteins) and aid in enhancing our understanding of activity and substrate specificity of UPOs. A combination of biochemical and structural information gathered from this chapter may elucidate the reasoning of the mutations created by NovoZymes between *Mro*UPO and artUPO, and may enlighten further engineering work of other UPOs.

9.1. Materials and Methods

9.1.1. Materials

Chemicals and other molecular biology materials were purchased from Alfa Aesar, Fisher Scientific, Fluorochem, Generon, Merck Chemicals Ltd, Millipore, New England Biolabs, Qiagen, Scientific Laboratory Supplies Ltd, Takara Bio, Tokyo Chemical Industry and VWR International.

9.1.2. Specific Activity Assay

To find the specific activity of a sample of a UPO with 1,2-(methylenedioxy-4-nitrobenzene) (NBD), **122**, a reaction containing 50 mM potassium phosphate, 1 mM **122**, 10 µL UPO and 2 mM H_2O_2 was monitored at 425 nm for 1 min. Absorbance against time was plotted to find the initial observed rate (k_{obs}) which was used to find the activity, U and specific activity, U mL^{-1} using Equations 2 and 3.

Equation 2. Calculation of enzyme activity

$$U / \mu mol\ min^{-1} = \frac{k_{obs}}{\varepsilon_{425} \times l} \times v_{rxn}$$

Equation 3. Calculation of specific activity of a UPO

$$U\ mL^{-1} = \frac{U}{v_{UPO}}$$

Where U = activity, k_{obs} = observed rate, ε_{425} = molar absorption coefficient of 4-nitrocatechol at 425 nm, l = pathway length, v_{rxn} = volume of the reaction, v_{UPO} = volume of enzyme.

9.1.3. Michaelis-Menten Kinetics

2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), **120** and dimethoxyphenol (DMP), were used to investigate the peroxidative activity of artUPO. 1,2-(methylenedioxy-4-nitrobenzene) (NBD) and 3,4-dimethoxybenzyl alcohol (VA) were used to investigate the peroxygenative activity of artUPO. The reactions were run with varying concentrations of substrate given in Table 9.1 and monitored using UV-vis spectroscopy.

Table 9.1 concentrations of ABTS, DMP, NBD and VA used for enzyme kinetics

ABTS / µM	DMP / µM	NBD / µM	VA / µM
10	10	10	50
50	50	15	100
75	75	20	200
100	100	50	250
150	150	75	500
200	200	100	750
300	300	125	1000
400	400	150	2000
500	500	200	3000
		250	4000
		300	5000
		400	10000
		500	

The reactions for both substrates were carried out on a 1 mL scale in 1 mL quartz cuvette.

For the ABTS assays, the reactions contained 50 mM citrate buffer at pH 4.4, ABTS with a final concentration given in Table 9.1, 1 µL artUPO (4 mg mL^{-1}) and were initiated with H_2O_2 (2 mM final). For the DMP assays, the reactions contained 50 mM citrate buffer at pH 5.5, DMP with a final concentration given in Table 9.1, 1 µL artUPO (4 mg mL^{-1}) and were initiated with H_2O_2 (2 mM final). For the NBD and VA assays, the reactions contained 50 mM potassium phosphate buffer at pH 7.0, NBD or VA with a final concentration given in Table 9.1 in acetonitrile (10% final), 1 µL artUPO (4 mg mL^{-1}) and were initiated with H_2O_2 (2 mM final). Both ABTS and NBD reactions were

monitored for 1 min at 418 nm for ABTS, 469 nm for DMP, 310 nm for VA and 425 nm for NBD.

Absorbance against time at each concentration of substrate was plotted to find the observed rate. The observed rate was converted into rate of product formation over time using the Beer-Lambert law. This new rate was plotted against concentration of substrate to find values of K_m, V_{max} and k_{cat} using the following equations.

Equation 4. Michaelis-Menten equation

$$V = \frac{V_{max}\,[S]}{K_m + [S]}$$

Equation 5. Calculation to determine the catalytic constant k_{cat}

$$k_{cat} = \frac{V_{max}}{[E_T]}$$

Where V = rate, V_{max} = maximum velocity (rate), [S] = substrate concentration, K_m = Michaelis-Menten constant, k_{cat} = turnover number, $[E_T]$ = enzyme concentration.

9.1.4. Analytical size exclusion

200 µL of artUPO (1 mg mL^{-1}) was run on a Superdex™ 200 Increase 10/300 GL in 10 mM potassium phosphate pH 7.4, 150 mM NaCl at 0.75 mL min^{-1}. This was compared to a protein standard mix (69385 Protein Standard Mix 15, Sigma Aldrich) run on the same column in the same conditions.

9.1.5. Incubation with dithiotheitol (DTT)

Samples of bacterial and yeast artUPO at 1.0 mg mL-1 were incubated for 3 min at 95 °C with SDS sample buffer that contained either 40 mM DTT or no reducing agent. These samples were then analysed using SDS-PAGE.

9.1.6. NanoDSF (Differential Scanning Fluorimetry)

$artUPO_{bact}$ and $artUPO_{yeast}$ were prepared at 2 mg mL^{-1} in 50 mM Tris, 300 mM NaCl pH 8.0 and were loaded onto a Prometheus NT 48 (NanoTemper) in 10 µL capillary tubes. The fluorescence was measured at 330 nm and 350 nm between 20 – 80 °C at 1 °C min^{-1}. The data was analysed using ThermalControl (NanoTemper).

9.1.7. Crystallisation of artUPO

9.1.7.1. 96 well plates

artUPO was purified from either a fermentation extract from yeast or a bacterial culture and was concentrated to 80 mg mL^{-1} and 15 mg mL^{-1} respectively. Initial crystallisation screens were set up in 96 well 2 drop plates using a mosquito robot with each well containing either 1:1 or 1:2 ratio of protein to buffer in a sitting drop with a total volume of 300 nL. These initial commercial screens included INDEX, PACT and clear strategy screen I (1 M bis-tris pH 6.2) and II (1 M HEPES pH 7.0).

9.1.7.2. 48 well optimisation plates

Using purified artUPO$_{yeast}$, a condition from a 96 well plate was optimised to improve crystal quality and quantity using a 48 well maxi plate. The protein and buffer were dispensed in 1:1 or 1:2 ratios to give a total of a 500 nL sitting drop. The original condition was changed by varying the salt concentration, buffer concentration, pH, polyethylene glycol (PEG) concentration and other additive concentration.

For a hit found in INDEX E10 (0.1 M bis tris methane pH 6.5, 45% (*v/v*) polypropylene glycol (PPG) 400), two optimisation plates were set up, with artUPO$_{yeast}$ as described in Section 9.1.7.2, varying pH, percentage of PPG 400 and different additives at different concentrations such as glycerol, butan-1,4-diol, isopropanol, glucose and ethylene glycol.

9.1.7.3. Optimisation using seeding

From a hit in the well INDEX H11 (0.1 M potassium thiocyanate, 30% *w/v* PEG monomethyl ether (MME) 2000), crystals were resuspended in total of 50 µL mother liquor and vortexed for 30 s with a Hampton Research Seed Bead to create a seed stock. This seed stock was then used for either screening with a brand new screen or to optimise the original condition where the crystal was formed. These seeding experiments were dispensed using an Oryx 8 liquid handling robot.

For the new 96 well two drop screen, 0.30 µL of artUPO$_{yeast}$ (60 mg mL^{-1}) and 0.10 µL of seed stock were mixed with 0.20 µL of buffer in the commercial screen PEG ION, resulting in a 1:2 protein to buffer ratio for the top drop of a condition. Remains of the

seed stock remained on the needle after dispensing the top drop and so the needle was used to dispense trace amounts of seed stock in the bottom drop.

Further optimisation of the condition found in H11 of the INDEX screen was carried out in a 48 well MAXI plate with 0.3 µL artUPO$_{yeast}$ (60 mg mL^{-1}), 0.05 µL of seed stock and 0.45 µL buffer which varied in potassium thiocyanate concentration and percentage of PEG MME 2000.

9.1.7.4. Data collection, structure solution and refinement

All artUPO crystals were fished and flash-frozen in liquid nitrogen. The crystals were initially tested in house using a Rigaku MicroMax-007 HF x-ray Generator to collect initial data and determine the quality of the crystals. Selected artUPO crystals were sent to the Diamond Light Source and the data was collected on either beamline I03 or I04-1. For both artUPO$_{yeast}$ crystals the data were processed and integrated using XDS and scaled using SCALA within the Xia2 pipeline. [163,164] The data for the artUPO$_{bact}$ crystal was integrated and processed using XDS and XSCALE, indexing with peaks found from all images, within the Xia2 3dii pipeline at Diamond. [163, 165] The artUPO$_{yeast}$ structures were solved by molecular replacement using MOLREP with *Mro*UPO from *Marasmius rotula* (75% sequence identity, PDB code 5FUJ) and were refined using iterative cycles of REFMAC and COOT using the CCP4i or CCP4i2 interface. [166-169] The artUPO$_{bact}$ structure was solved and refined as previously described but using artUPO$_{yeast}$ as the model for molecular replacement.

9.1.7.5. Co-crystallisation and soaking of artUPO with ligands

Crystals of artUPO$_{yeast}$ were grown in 0.1 M potassium thiocyanate, 30% PEG MME with a 2:1 ratio of 80 mg mL^{-1} artUPO$_{yeast}$ and mother liquor, resulting in a 1 µL drop, for three months using a mosquito robot to set up the 48 well crystal tray. These crystals were broken using a seed bead (Hampton Research) and used to make a seed stock. The stock was seeded into a 48-well plate using an Oryx robot with varied conditions of 0.05 – 0.2 M KSCN and 25 – 35% PEG MME with a 1:1 ratio of artUPO$_{yeast}$ (15% seed stock, 85% 60 mg mL^{-1} artUPO) and mother liquor. From these higher quality crystals, either the crystals were soaked with an excess amount of the desired ligand overnight (soaking) or the crystals were used as seed stock for screening crystallisation with artUPO$_{yeast}$ and 5% solvent containing 10 mM of the ligand (co-crystallisation).

The following ligands were used in co-crystallisation: 12-hydroxydodecanoic acid (HDDA) in DMSO, cortisone in 50:50 DMSO and methanol, ethylbenzene in DMSO and thioanisole in DMSO.

9.1.8. Substrate channel dimension calculations using MOLEonline

To calculate the substrate channel dimensions of artUPO, the sequence was inputted into MOLEonline 2.0 (accessed September 2021). [170]

9.1.9. AlphaFold 2

Alphafold2-multimer models were made using ColabFold with MMseqs2 without using a template. [171-174]

9.2. Results and Discussion

9.2.1. Michaelis-Menten kinetics

Four different assays, with ABTS **118a**, DMP **120a**, NBD **199a** or VA **121** were used to investigate both peroxidative (one electron transfer) and peroxygenative (two electron transfer) activities of artUPO$_{bact}$ and artUPO$_{yeast}$ (Figure 9.1). The reactions with all four substrates at varying concentrations were monitored using UV-vis spectroscopy. Plotting the rate of product formation against each concentration resulted in Michaelis-Menten curves for each substrate (Figure 9.2 and 9.3). From these curves, values of K_m, V_{max} and k_{cat} were determined and are given in Table 9.2.

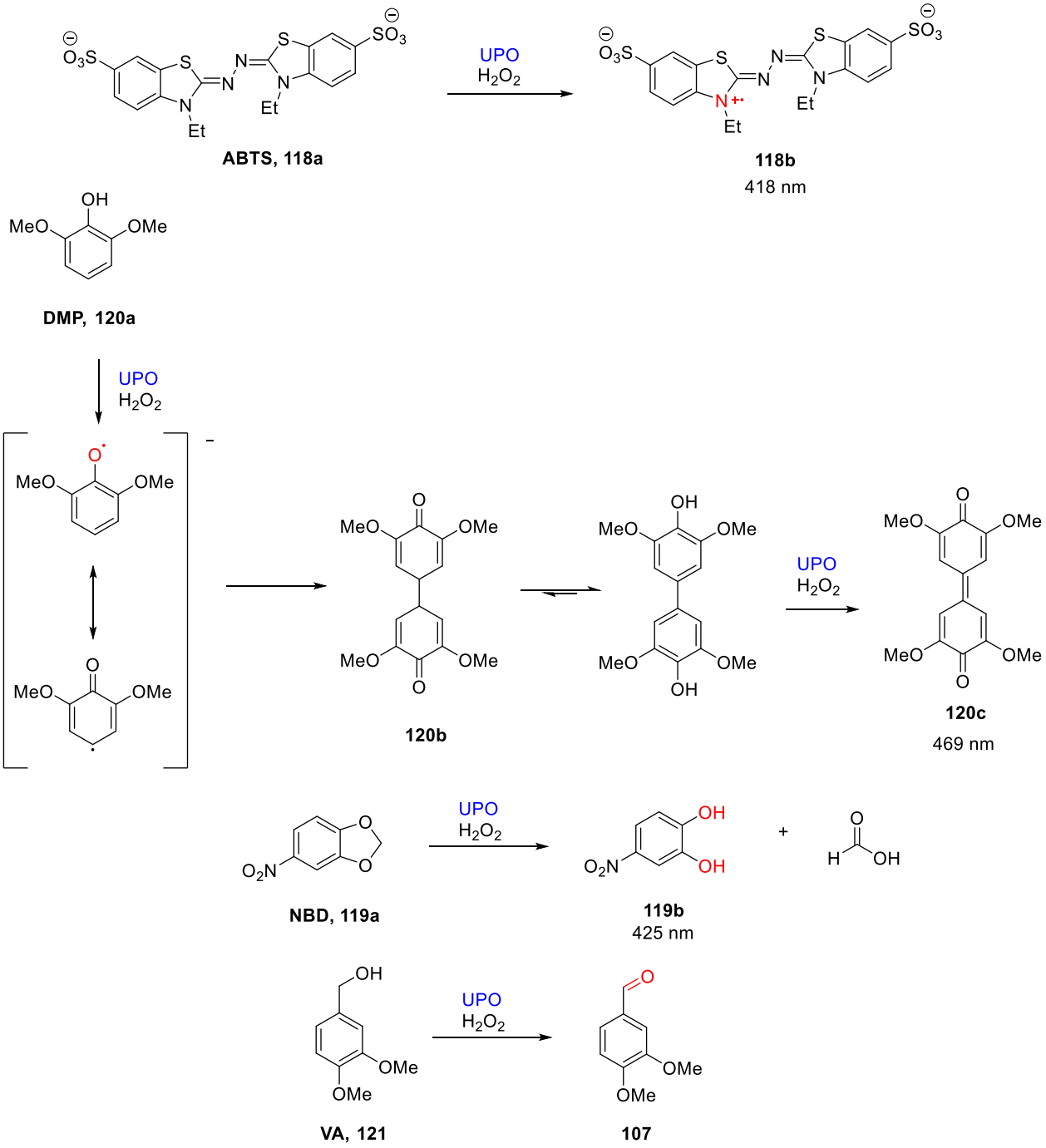


Figure 9.1 Oxidation of ABTS **118a** *via* one electron transfer, DMP **120a** *via* one electron transfer, NBD **119a** *via* two electron transfer and VA **121** *via* two electron transfer

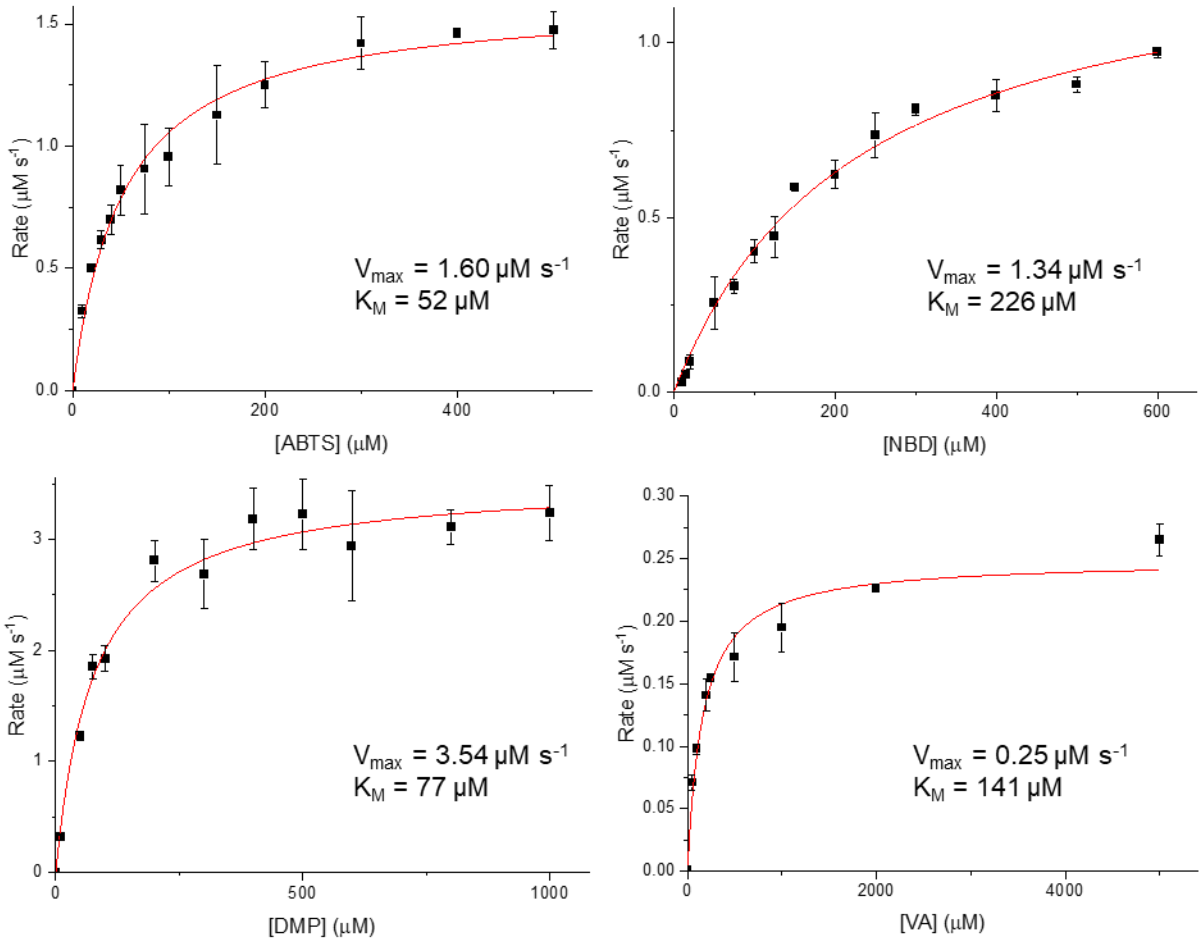


Figure 9.2 Michaelis-Menten curves of artUPO$_{yeast}$ with **A**. ABTS **118a**, **B**. NBD **119a**, **C**. DMP **120a** and **D**. VA **121**. All data were carried out in triplicates

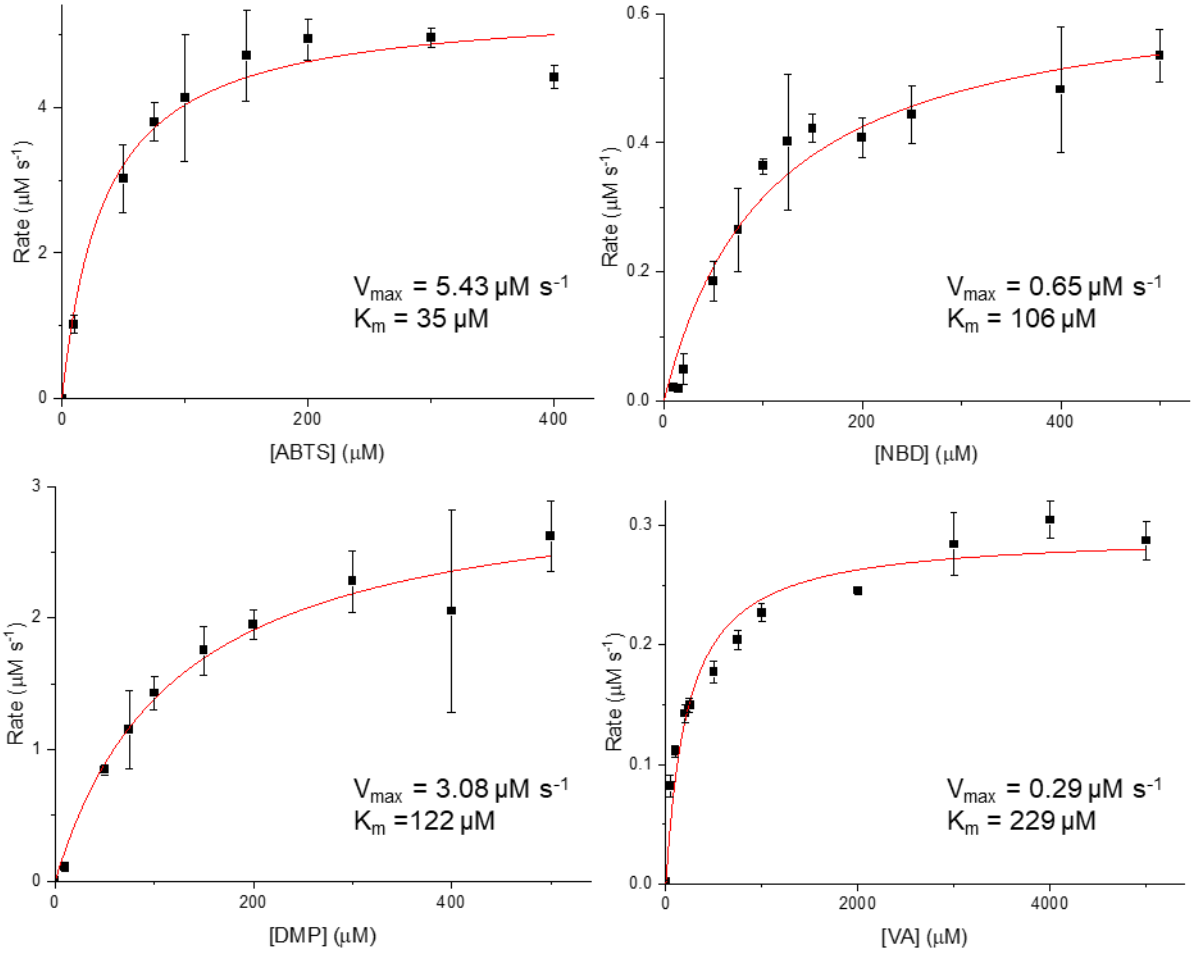


Figure 9.3 Michaelis-Menten curves of artUPO$_{bact}$ with **A**. ABTS **118a**, **B**. NBD **119a**, **C**. DMP **120a** and **D**. VA **121**. A. All data were carried out in triplicates

Table 9.2 Kinetic parameters of both bacterial and yeast artUPO with ABTS, NBD, VA and DMP

		V_{max} / μM s^{-1}	K_m / μM	k_{cat} / s^{-1}	k_{cat}/K_M / s^{-1} mM^{-1}
$artUPO_{yeast}$	ABTS	1.60 ± 0.04	52 ± 4.6	13	256
	NBD	1.34 ± 0.06	226 ± 20	11	50
	VA	0.25 ± 0.01	141 ± 15	2	14
	DMP	3.54 ± 0.11	77 ± 11	29	383
$artUPO_{bact}$	ABTS	5.43 ± 0.27	35 ± 7.6	45	1306
	NBD	0.65 ± 0.05	106 ± 23	5	51
	VA	0.29 ± 0.01	229 ± 40	2	11
	DMP	3.08 ± 0.20	122 ± 22	26	209

The Michaelis-Menten constant, K_M, corresponds to the affinity of a particular substrate towards the enzyme which relates to the rate of the reaction (Equation 1). The turnover number, k_{cat}, is a useful constant for a given substrate and enzyme, however this constant is dependent on substrate concentration and is therefore difficult to compare values of k_{cat} with other enzymes in other studies. Values of k_{cat}/K_m relate to the catalytic efficiency of a reaction and is independent of substrate concentration making this a useful value for the comparison of different enzymes.

Comparing the two activities in general, both forms of artUPO have higher values of k_{cat}/K_M values for the peroxidative activity than the peroxygenative activity. These results were expected because the enzyme (*Mro*UPO) from which artUPO was derived from also has higher peroxidative activity than peroxygenative activity (Table 9.3). [158]

For both types of activity, two substrates were chosen where one of the reactions would be carried out in water and buffer, and the other would have an additional 10% acetonitrile (*v*/*v*) in the reaction. This was to investigate the effect of organic solvent on the rate of reaction when comparing the same reaction type. Mostly no significant difference was observed when comparing the two substrates for the same reaction type despite changes in solvent content (i.e. ABTS vs DMP for peroxidative activity or NBD vs VA for peroxygenative activity) in artUPO$_{yeast}$ reactions. A large difference in catalytic efficiency was observed in the peroxidative reactions with ABTS and DMP using artUPO$_{bact}$ where the reaction seemed to favour the lack of solvent. artUPO$_{bact}$ appeared to have a much higher activity with ABTS but not with DMP when compared to artUPO$_{yeast}$. The reason for this difference remains unknown but further screening of the effects of solvent on the rate of reaction may provide more information on this effect. Differences in activity due to the presence of solvent may be caused by the stability of the unglycosylated enzyme, as glycosylation has already been shown to be important to the overall stability of UPOs. [104] However, k_{cat}/K_M values for the peroxygenative activity of artUPO are very similar between the two expression systems. This suggests that glycosylation has little effect on specifically the peroxygenative activity of artUPO, hence differences in peroxygenative activity may not need to be a consideration when choosing an expression host.

Both artUPOs can be compared to the kinetic activity of *Mro*UPO, the UPO artUPO is derived from, and the most studied UPO *Aae*UPO. The kinetic parameters of *Mro*UPO and *Aae*UPO are given in Table 9.3.

Table 9.3 Kinetic parameters of wild type *Mro*UPO and *Aae*UPO

		K_m / μM	k_{cat} / s^{-1}	k_{cat}/K_M / s^{-1} mM^{-1}	Reference
*Mro*UPO	ABTS	71	25	353	Gröbe et al[158]
	VA	279	49	176	
	DMP	133	70	529	
*Aae*UPO	ABTS	37	283	7648	Ullrich *et al*[77]
	NBD	1190	457	384	Poraj-Kobielska[175]
	VA	1001	269	269	Ullrich *et al*[77]
	DMP	298	108	362	Ullrich *et al*[77]

All UPOs tend to have a higher k_{cat}/K_M values for the substrates ABTS and DMP, implying they have higher peroxidative activity than peroxygenative activity. This is potentially unfavourable as the peroxygenative activity allows UPOs to activate C-H bonds, which is usually desirable. [176] Both forms of artUPO have relatively similar kinetic parameters to *Mro*UPO for the peroxidative activity and are within the same order of magnitude. However, differences are observed when comparing parameters using VA as a substrate. k_{cat} for *Mro*UPO with VA is higher than the k_{cat} values for both artUPOs with VA and, therefore, the *Mro*UPO k_{cat}/K_M value with VA is much higher. This suggests that the mutations in artUPO compared to *Mro*UPO do not to affect the peroxidative activity but do affect the peroxygenative activity.

Compared to *Aae*UPO, the kinetic parameters for both artUPOs are significantly lower for all substrates, suggesting that both artUPOs have less peroxidative and peroxygenative activity than *Aae*UPO. However, the differences in activity between artUPO and *Aae*UPO observed in these assays are not necessarily representative of

activities towards target compounds. Screening of a wider range of compounds may give a better idea of broad range activity.

9.2.2. Analytical Size Exclusion Chromatography (SEC)

Analytical SEC can estimate the molecular weight of a protein moiety when compared to known standards. SEC separates compounds by size with a stationary phase of cross linked agarose and dextran, where larger molecules will travel through the column faster and have an earlier elution time compared to smaller molecules. Purified samples of both bacterial and yeast artUPO at 1 mg mL^{-1} were run and compared to a protein standard mix (Sigma Aldrich) (Figure 9.4).

To calculate the molecular mass of both bacterial and yeast artUPO species, a plot of the log of the molecular weights of the standards against the elution volume were made and the molecular weight of the artUPO species was extrapolated from this plot (Figure 9.5).

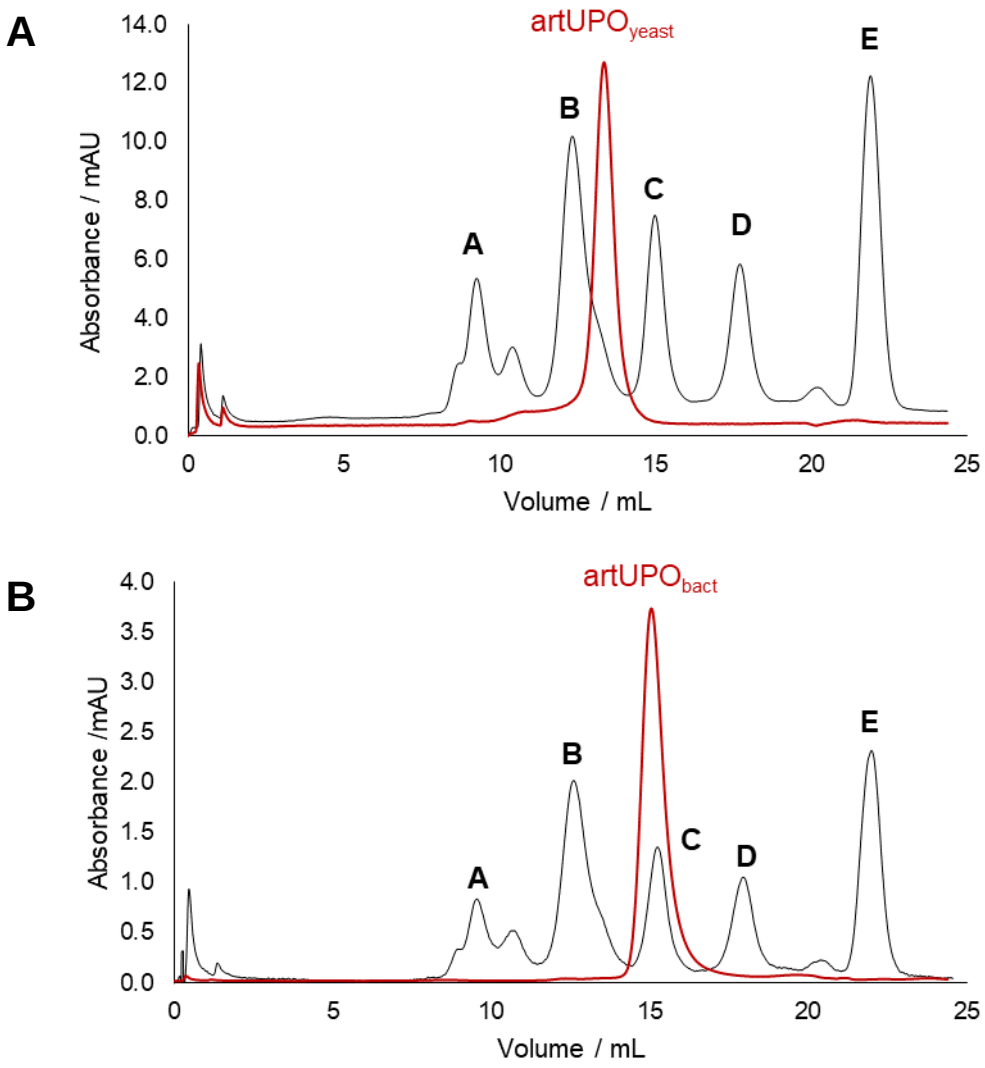


Figure 9.4 Analytical SEC chromatograms overlaid where the black trace is the protein standard mix and the red trace is either **A**. artUPO$_{yeast}$ or **B**. artUPO$_{bact}$. For the protein standard mix, A: Thyroglobulin 670.0 kDa, B: γ-globulins from bovine blood 150.0 kDa, C: Albumin chicken egg grade VI (ovalbumin) 44.3 kDa, D: Ribonuclease A type I-A from bovine pancreas 13.7 kDa, E: *p*-aminobenzoic acid (pABA) and other buffers and preservatives.

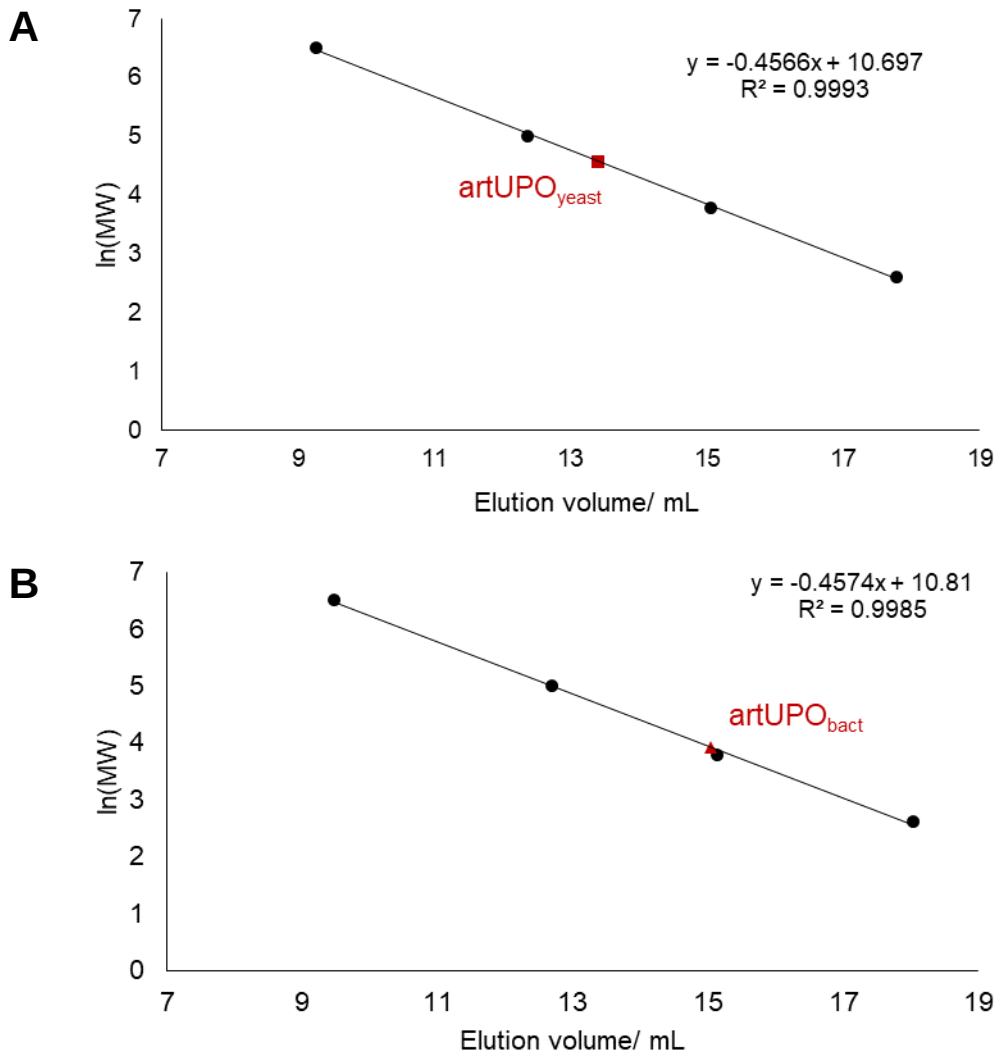


Figure 9.5 ln(MW) of the standard compounds against elution volume. **A**: standards run with artUPO$_{yeast}$ **B**: standards run with artUPO$_{bact}$

The calculated molecular weight of artUPO from the amino acid sequence is 33.2 kDa, although this does not account for glycosylation. From SDS PAGE analysis, when the protein is denatured, artUPO$_{yeast}$ appears to have a molecular weight of roughly 45 kDa (Figure 8.20) but artUPO$_{bact}$ runs at the more expected molecular weight of 33 kDa (Figure 8.24). Figure 9.4 shows that on a SEC column compared to a standard, artUPO$_{yeast}$ elutes between γ-globulin (150.0 kDa) and Ovalbumin (44.3 kDa), and the molecular weight was calculated to be 98 kDa. This implies that artUPO$_{yeast}$ exists as a dimer in solution. artUPO$_{bact}$ elutes at a similar volume as Ovalbumin (44.3 kDa) and

the MW was calculated to be 51 kDa, which suggests that artUPO$_{bact}$ may exist as either as a dimer or as a mixed population of both monomer and dimer.

9.2.3. Incubation of artUPO with and without reducing agent

To investigate the presence of a disulfide bond in artUPO produced in both bacteria and yeast, an SDS PAGE gel was run with both artUPO samples with either the presence of DTT (dithiothreitol) or without DTT in the SDS sample buffer. The sample buffer also did not contain β-mercaptoethanol. If artUPO contains a disulfide bridge, the DTT should reduce this bond and a change in molecular weight should be observed compared to a sample without reducing agent. In this experiment the samples were incubated with or without DTT for 3 min with the SDS sample buffer (without β-mercaptoethanol) at 95 °C. The samples were then loaded onto an SDS-PAGE gel and run, with the results shown in Figure 9.5.

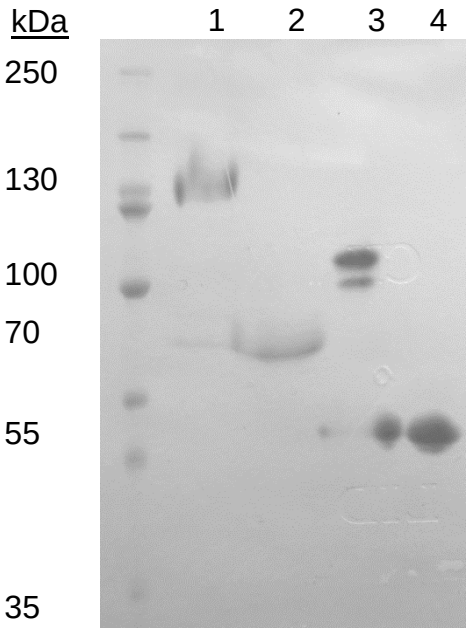


Figure 9.6 SDS-PAGE analysis of DTT treatment with artUPO where lane 1 has artUPO$_{yeast}$ without DTT, lane 2 has artUPO$_{yeast}$ with DTT, lane 3 has artUPO$_{bact}$ without DTT and lane 4 has artUPO$_{bact}$ with DTT

In Figure 9.6 where in both bacterial and yeast artUPO samples with DTT (lanes 2 and 4 respectively), the molecular weight is lower by half compared to its respective samples without DTT treatment (lanes 1 and 3). This implies that artUPO exists as a dimer and the units are held together with a disulfide bond.

9.2.4. NanoDSF

To investigate the differences in thermal stability of artUPO$_{bact}$ and artUPO$_{yeast}$, NanoDSF experiments were carried out to determine the melting temperature (T_m) and the results are shown in Figure 9.7.

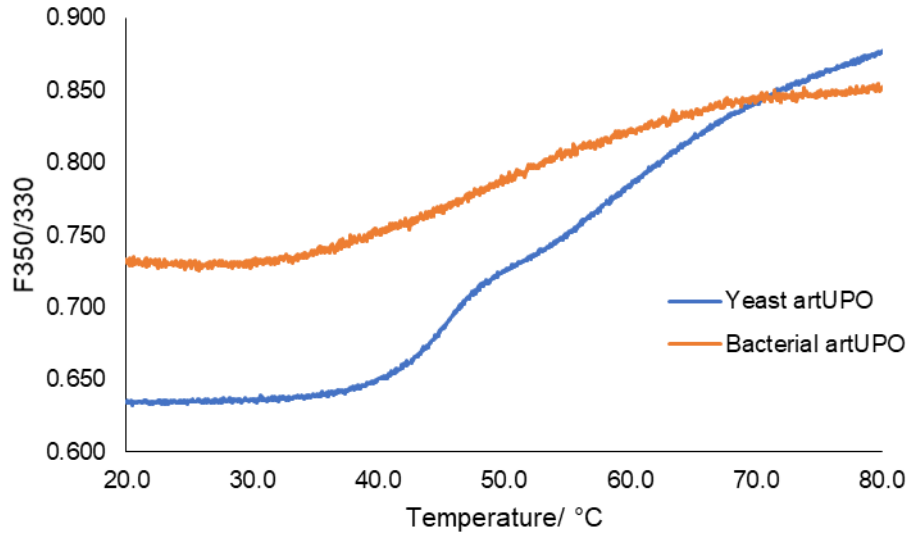


Figure 9.7 Ratio of the fluorescence at 350 nm/330 nm against temperature from 20 – 80 °C with 2 mg mL^{-1} of **blue**: $artUPO_{yeast}$ and **orange**: $artUPO_{bact}$. The artUPOs were in 50 mM Tris pH 8.0, 300 mM NaCl and the temperature ramp was 1 °C min^{-1}

As the temperature increases, a protein will start to unfold, denature and internal residues will become exposed to the solvent. In nanoDSF, the fluorescence of tryptophan and tyrosine are monitored at 350 nm and 330 nm. As the tryptophan and the tyrosine residues become exposed, their emission wavelength shifts from 330 nm to 350 nm. [177] The ratio of 350 nm emission to 330 nm emission therefore increases as the protein becomes denatured. At a given temperature, a sharp increase in F350/330 will be observed, giving rise to a sigmoidal curve. The melting temperature T_m, the temperature at which half of the protein is unfolded, can be derived from the inflection of the curve and gives an indication of the thermodynamic stability of the protein. [178] NanoDSF, compared to DSF, does not require labelling and uses a small amount of protein and so this technique was suitable to examine the thermostability of both artUPOs. [179]

In Figure 9.7, the curve for artUPO$_{bact}$ has a sigmoidal shape as would be typical for most proteins, and has a derived T_m of 47.8 °C. However, the curve for artUPO$_{yeast}$ appears to have two points of inflection, which suggests there may be two different stages of unfolding. The T_m derived from the first inflection point was 45.5 °C but the second inflection point could not be clearly identified.

9.2.5. Crystallisation of artUPO

9.2.5.1. Crystal formation and optimisation of yeast artUPO

Initial crystallisation screens were set up with yeast artUPO$_{yeast}$ (ranging from 10 to 80 mg mL^{-1}) in 96 well plates to find conditions that would promote crystallisation of the protein. Two months after plates were set, crystals appeared in different conditions in the INDEX screen when the protein concentration was 80 mg mL^{-1}. These crystals were red which was indicated the presence of heme (Figure 9.8).

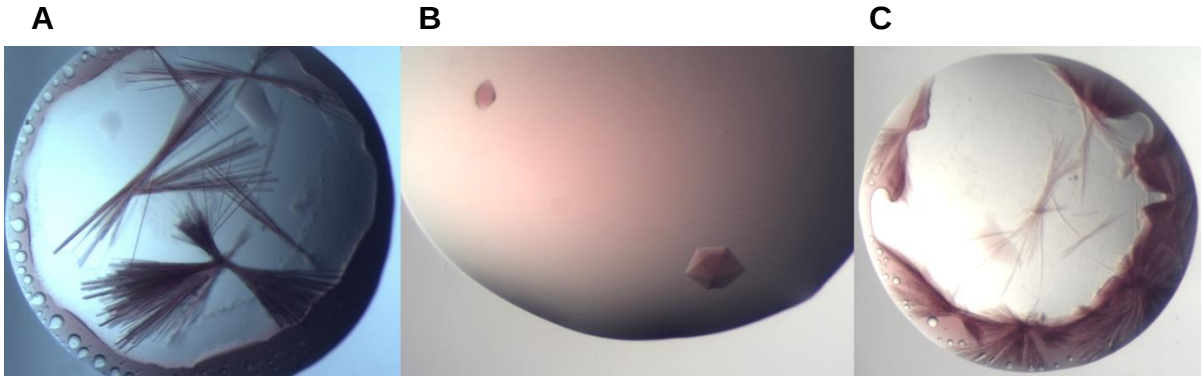


Figure 9.8 artUPO crystals found two months later after initial set up in INDEX screen **A**. 0.1 M KSCN, 30% PEG MME 2000, 1:2 protein to buffer **B**. 2.4 M sodium malonate 1:1 protein to buffer **C**. 0.1 M bis tris methane pH 6.5, 20% PEG MME 5000, 1:2 protein to buffer

Crystals in Figure 9.8A (INDEX H11) were broken down into smaller fragments using a seed bead to make a seed stock. This stock was used in a new 96 well plate screen, PEG ION, and was used in a 48 well optimisation plate, based the original conditions used to make the crystals for seed stock (Figure 9.7A).

In crystallisation, there are various states where the protein to be crystallised is either stable, metastable, labile or in the precipitation state. In the stable state, the concentration of protein is undersaturated and neither nucleation nor growth of protein crystals is possible. In the metastable and labile states the protein is saturated, and in the labile state, nucleation and growth of crystals are possible. However, in the metastable state, nucleation is not possible but crystal growth from seeds is possible. [180] The purpose of creating a seed stock was to screen for conditions where the protein would be in the labile state, and the nucleation step of protein crystal growth can be skipped. By using a seed stock along with purified protein in new conditions, higher quality crystals than the original hit can be grown.

In the new PEG ION plate, multiple crystals were formed within two weeks of setup (Figure 9.9).

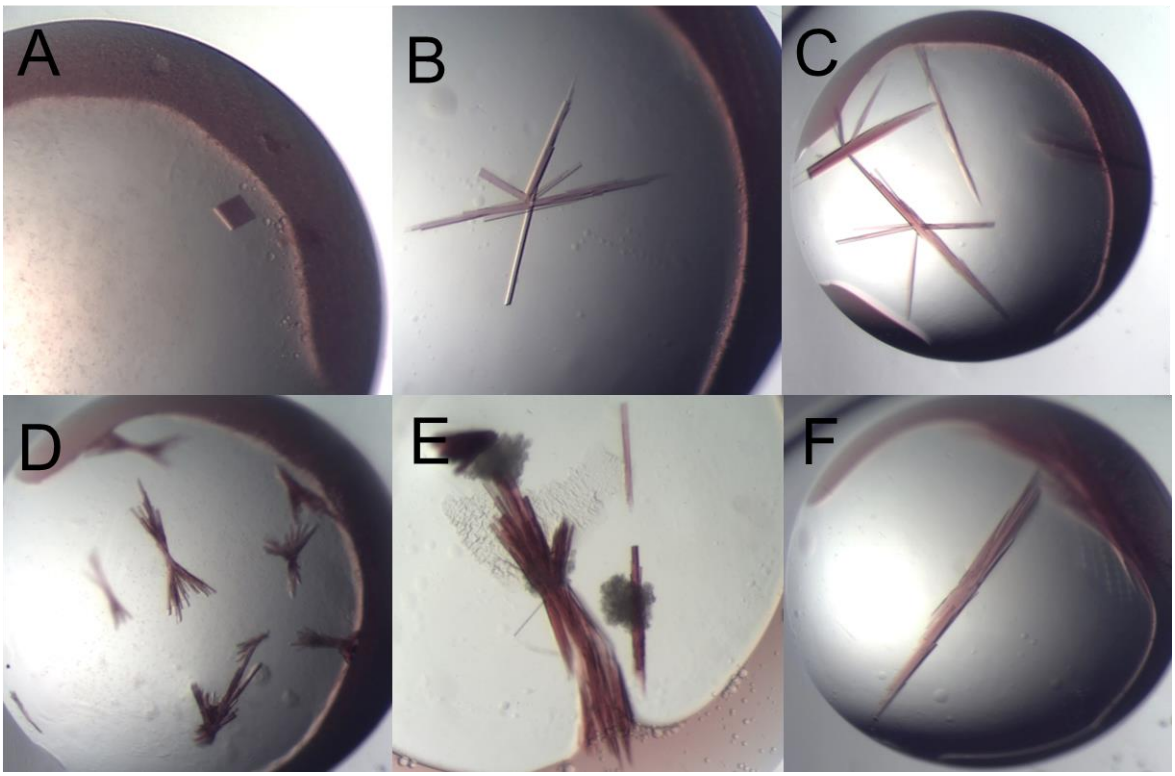


Figure 9.9 Selection of yeast artUPO crystals from PEG ION screen set up with seed stock **A**. 0.2 M calcium chloride dihydrate pH 5.1, 20% PEG 3350 **B**. 0.2 M ammonium iodide pH 6.2, 20% PEG 3350 **C**. 0.2 M lithium acetate dihydrate pH 7.9, 20% PEG 3350 **D**. 0.2 M Ammonium formate pH 7.3, 20% PEG 3350 **E**. 0.2 M ammonium phosphate dibasic dihydrate pH 9.1 20% PEG 3350 **F**. 0.1 M tryptone, 1 mM NaN_3, 0.05 M HEPES sodium pH 7.0, 20% PEG 3350

The crystals in Figure 9.9A and 9.9F were tested in house by x-ray diffraction which resulted in the diffraction patterns in Figure 9.10A and 9.10B respectively. These samples appeared to diffract to a higher resolution than the previous crystal, however rings in the pattern suggest that the protein sample contains ice which is not desirable for structure determination. Despite this, both crystals were sent to the Diamond Light Source synchrotron and data was collected for the crystal shown in Figure 9.8F. This resulted in a 2.0 Å resolution structure with an R_{factor} of 0.19 and an R_{free} value of 0.22 (Table 9.4).

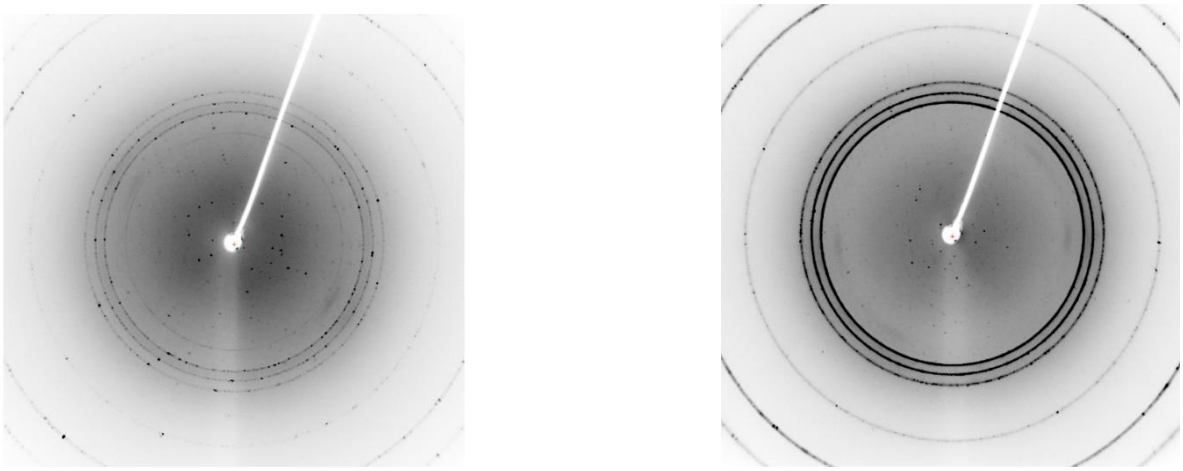


Figure 9.10 x-ray diffraction patterns of crystals tested in house in **left**: Figure 9.9A, **right**: Figure 9.9F

The optimisation plate with crystal seed resulted in crystals after a few weeks. Some of the crystals were soaked with either 12-hydroxydodecanoic acid or cortisone as an attempt to gain a structure with a ligand in the active site. Both of these compounds were chosen because they are both products of biotransformations that *Mro*UPO has been able to catalyse. [93,105] These crystals were fished, with no cryoprotectant, and were tested in house. The crystals from Figure 9.11C were then sent to the Diamond Light Source synchrotron. Figure 9.12 shows the x-ray diffraction patterns obtained from Diamond.

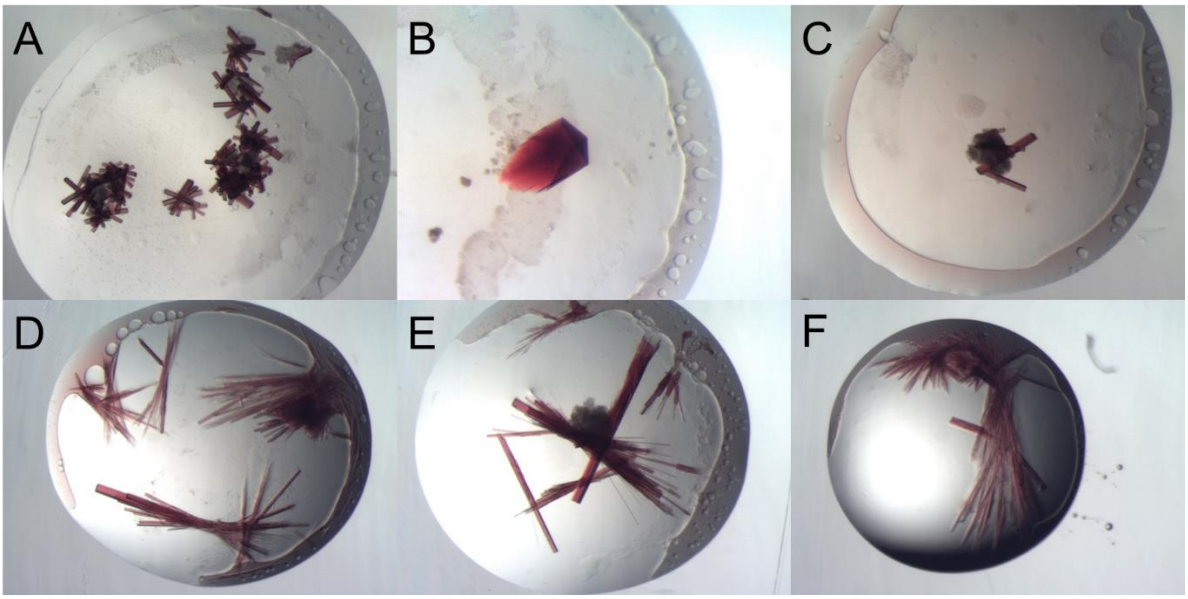


Figure 9.11 Selection of artUPO crystals from INDEX H11 optimisation screen set up with seed stock **A**. 0.1 M KSCN, 25% PEG MME 2000 **B**. 0.15 M KSCN, 25% PEG MME 2000 **C**. 0.15 M KSCN, 25% PEG MME 2000 **D**. 0.1 M KSCN, 30% PEG MME 2000 **E**. 0.1 M KSCN, 30% PEG MME 2000 **F**. 0.1 M KSCN, 35% PEG MME 2000

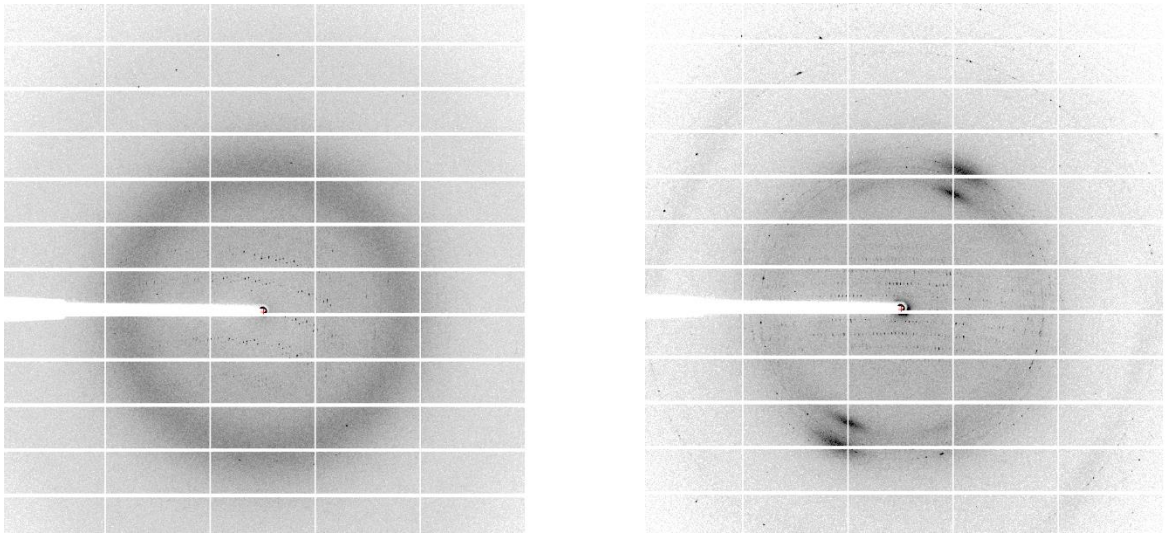


Figure 9.12 x-ray diffraction patterns of crystals in Figure 9.11C tested at the Diamond Light Source synchrotron

Although no ligand was seen the resulting structure, the new structure obtained had a higher resolution and a different space group to the first artUPO$_{yeast}$ structure (Table 9.4).

9.2.5.2. Crystal formation of bacterial artUPO

Initial screens were set up using artUPO$_{bact}$ (15 mg mL^{-1}) in 96 well plates in the same way as described previously. After 12 days, red crystals appeared in the screens INDEX, PACT and CSS 1+2 (Figure 9.13). The crystals in Figure 9.13A, B and C were sent to the Diamond Light synchotron and the crystals in Figure 9.13C resulted in a 2.07 Å resolution structure with an R_{factor} of 0.21 and an R_{free} value of 0.26 (Table 9.4).

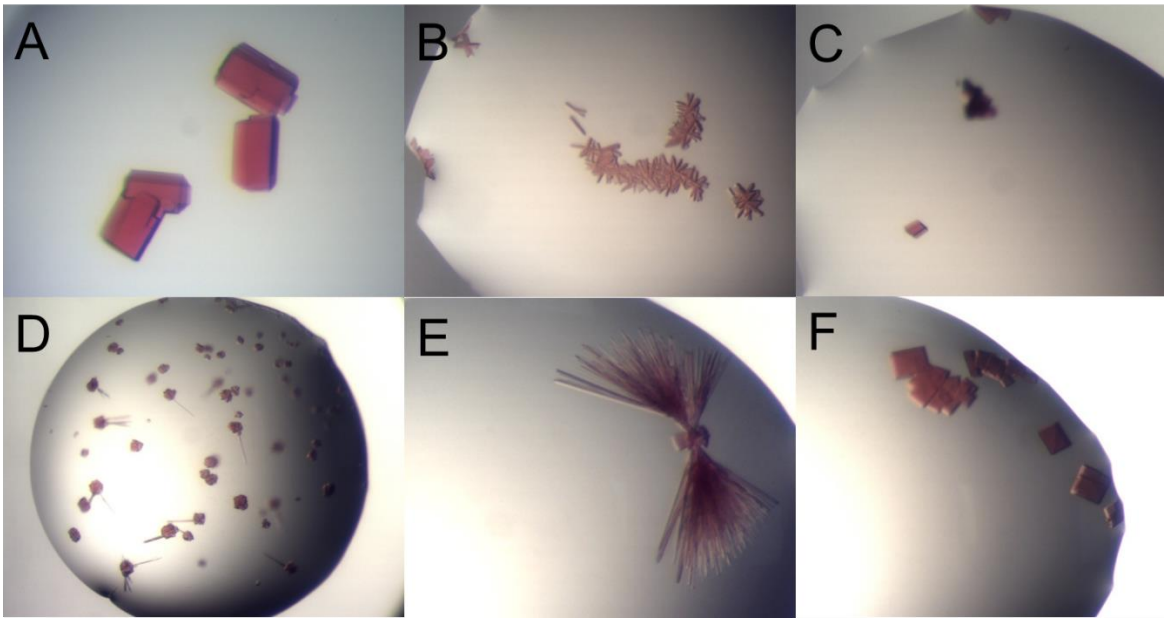


Figure 9.13 Selection of bacterial artUPO crystals from initial screens **A**. 0.1 M bis-tris pH 6.2, 1.5 M $(NH_4)_2SO_4$, 1:1 protein to buffer **B**. 0.1 M bis-tris pH 6.5, 0.2 M NH_4OAc, 25% PEG 3350, 1:1 protein to buffer **C**. 0.1 M HEPES pH 7.5, 25% PEG 3350, 1:1 protein to buffer **D**. 0.1 M PCTP buffer pH 5.0, 25% PEG 1500, 1:2 protein to buffer **E**. 0.1 M MMT pH 4.0, 25% PEG 1500, 1:2 protein to buffer **F**. 0.1 M PCTP buffer pH 5.0, 25% PEG 1500, 1:1 protein to buffer

9.2.5.3. Solving, building and refinement of the artUPO structures

For the artUPO$_{yeast}$, two structures were refined in two different space groups: one in *P* $2_12_12_1$ with a resolution of 2.01 Å and the other in *C* 2 2 2_1 with a resolution of 1.21 Å. The *P* $2_12_12_1$ crystal has two molecules in the asymmetric unit whereas the *C* 2 2 2_1 crystal has one. Both structures were solved by molecular replacement using MOLREP with the dimer of *Mro*UPO as a model found in the Protein Data bank but this structure is not currently published (PDB code 5FUJ). This structure was an appropriate model to use for molecular replacement because it shares 75% sequence identity to artUPO. The solvent content in both crystal forms was found to be 55%. The structures were built and refined using iterative cycles in Coot and REFMAC.

For the artUPO$_{bact}$, the structure was determined in the space group *P* $2_12_12_1$ with a resolution of 2.07 Å. It has two molecules in the asymmetric unit and was found to have a solvent content of 37%. The structure was solved using MOLREP with the structure from yeast artUPO. and was built and refined as previously described using Coot and REFMAC.

Table 9.4 Data collection and refinement statistics for all artUPO crystals, numbers in brackets refer to data for highest resolution shells

	artUPO$_{yeast}$ 1	artUPO$_{yeast}$ 2	artUPO$_{yeast}$
Beamline	Diamond I03	Diamond I04-1	Diamond I03
Wavelength (Å)	0.976254	0.915870	0.9762
Resolution (Å)	47.82-2.01 (2.06-2.01)	77.27-1.21 (1.23-1.21)	38.34-2.07 (2.13-2.07)
Space Group	*P* $2_12_12_1$	*C* 2 2 2_1	*P* $2_12_12_1$
Unit cell (Å)	a = 50.34; b = 75.82; c = 151.70 $\alpha = \beta = \gamma = 90.00$	a = 50.85; b = 74.17; c = 154.54 $\alpha = \beta = \gamma = 90.00$	a = 46.03; b = 58.81; c = 151.46 $\alpha = \beta = \gamma = 90.00$
No. of molecules in the asymmetric unit	2	1	2
Unique reflections	39487 (2858)	89172 (4348)	25922 (1956)
Completeness (%)	99.9 (99.9)	99.9 (100.0)	99.9 (100.0)
R_{merge} (%)	0.08 (0.50)	0.04 (0.43)	0.04 (0.833)
$R_{p.i.m.}$	0.04 (0.28)	0.02 (0.28)	0.02 (0.35)
Multiplicity	8.0 (7.8)	7.7 (6.3)	12.8 (11.5)
$<I/\sigma(I)>$	16.5 (4.4)	23.6 (3.8)	12.5 (2.9)
Overall *B* factor from Wilson plot (Å^2)	20	11	32
$CC_{1/2}$	1.00 (0.96)	1.00 (0.92)	1.00 (0.92)
R_{cryst}/ R_{free} (%)	19.3/22.3	16.3/17.7	20.8/25.5
r.m.s.d 1-2 bonds (Å)	0.010	0.019	0.008
r.m.s.d 1-3 angles (°)	1.648	2.062	1.531
Avge main chain B (Å^2)	29	13	31
Avge side chain B (Å^2)	32	16	33
Avge water B (Å^2)	34	17	27

9.2.5.4. Investigation of the artUPO structures

In all structures of artUPO, each monomer contains eight α-helices, a small antiparallel β-sheet pair and a prosthetic heme group. The helices are arranged in a globin fold where a hydrophobic pocket is formed for the heme group to bind to the protein. The monomers are joined by a disulfide bond Cys232 on each subunit (Figure 9.14).

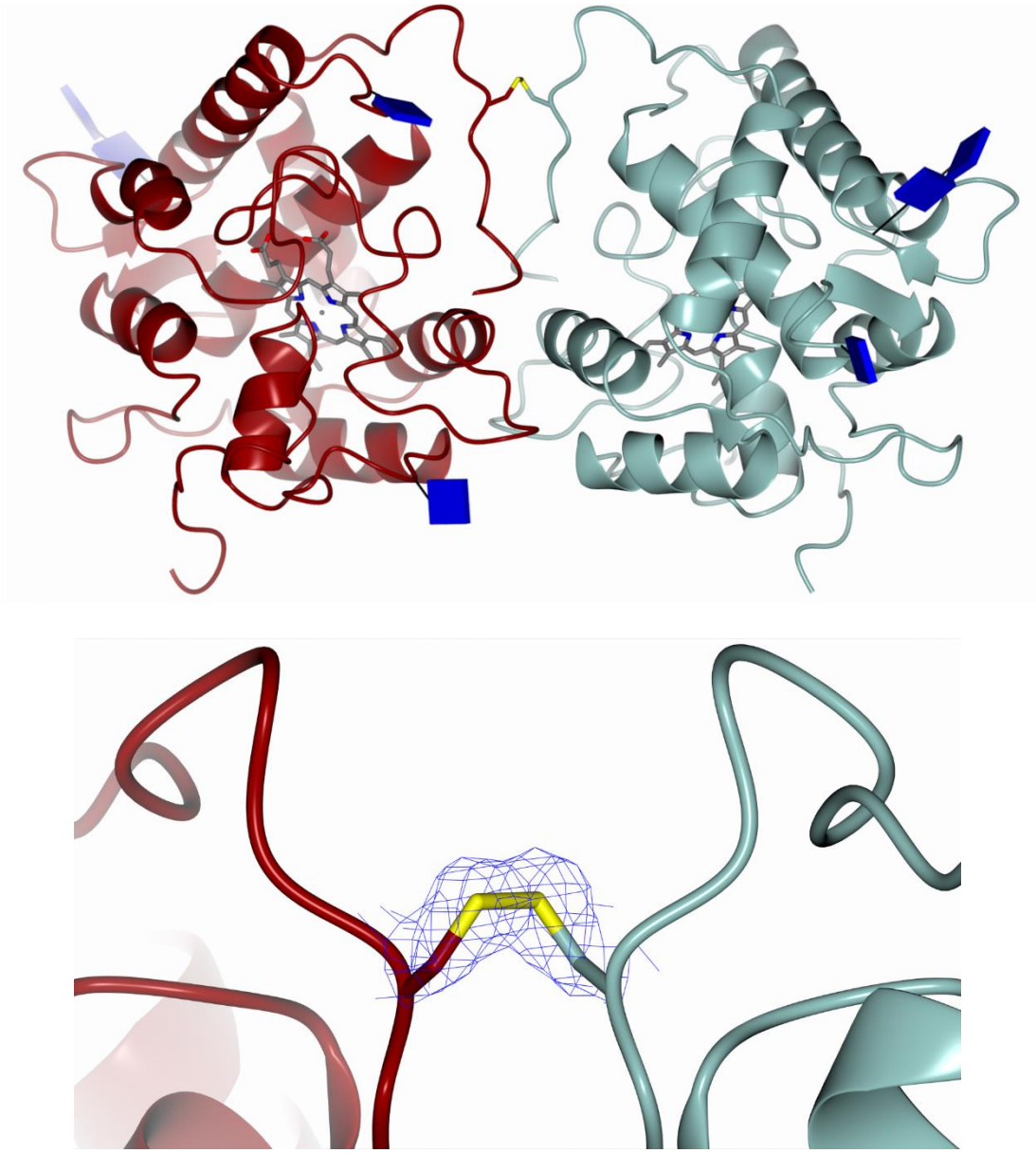


Figure 9.14 **Top**: Dimer of artUPO$_{yeast}$ 1 molecule in the space group P $2_12_12_1$ with the chains coloured in dark red and sea green, sugars represented in glycoblocks (N-acetylglucosamine GlcNAc) in dark blue) and the heme centres represented as cylinders (carbon: grey, nitrogen: blue, oxygen: red, sulfur: yellow) **Bottom**: disulfide bond between cysteine residues (Cys232) in each monomer unit with the $2mF_o-DF_c$ map shown, contoured to 1.0 σ

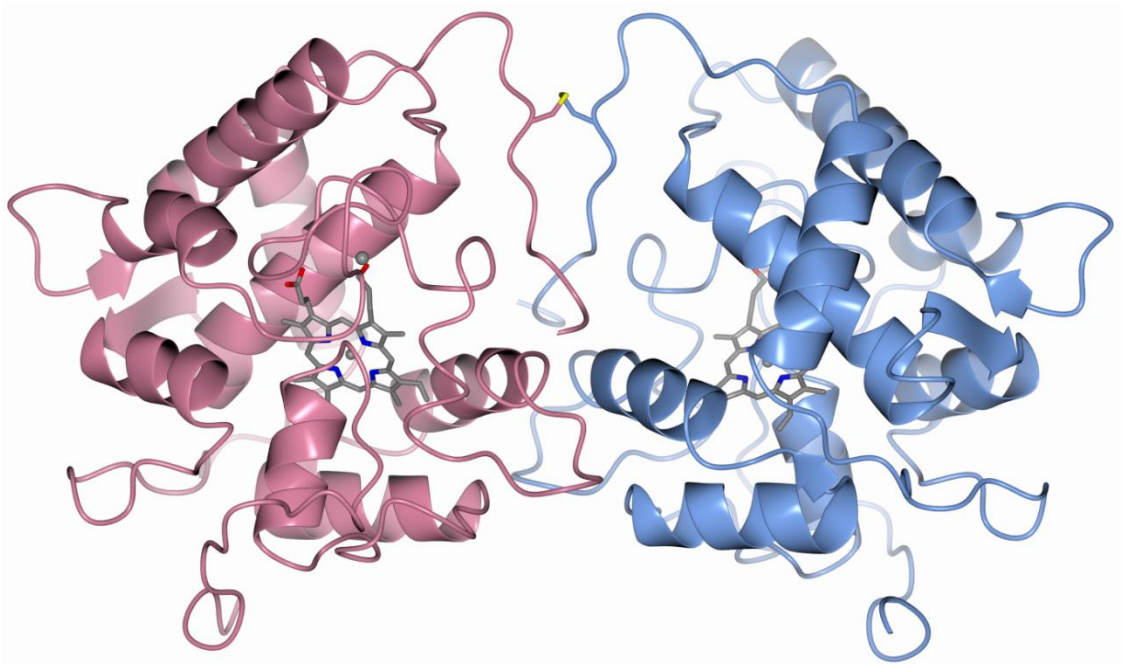


Figure 9.15 Dimer of artUPO$_{bact}$ molecule in the space group P $2_12_12_1$ with the chains coloured in pink and blue. No glycosylation is observed

In both artUPO$_{yeast}$ structures each monomer has at least two *N*-linked glycosylation sites which are observed at Asn129 and Asn174, linked *via* *N*-acetylglucosamine. In the artUPO$_{yeast}$ structure, two additional *N*-glycosylation sites are seen at Asn42 and Asn150 on one of the chains, but not the other (Figure 9.16). Because sugar units can be large and flexible, there is a possibility that more glycosylation sites exist but they are not visible by x-ray crystallography as only rigid and ordered moieties can be seen using this technique. Because *E. coli* are unable to carry post-translational glycosylation, no glycosylation is observed in the bacterial artUPO structure (Figure 9.15).

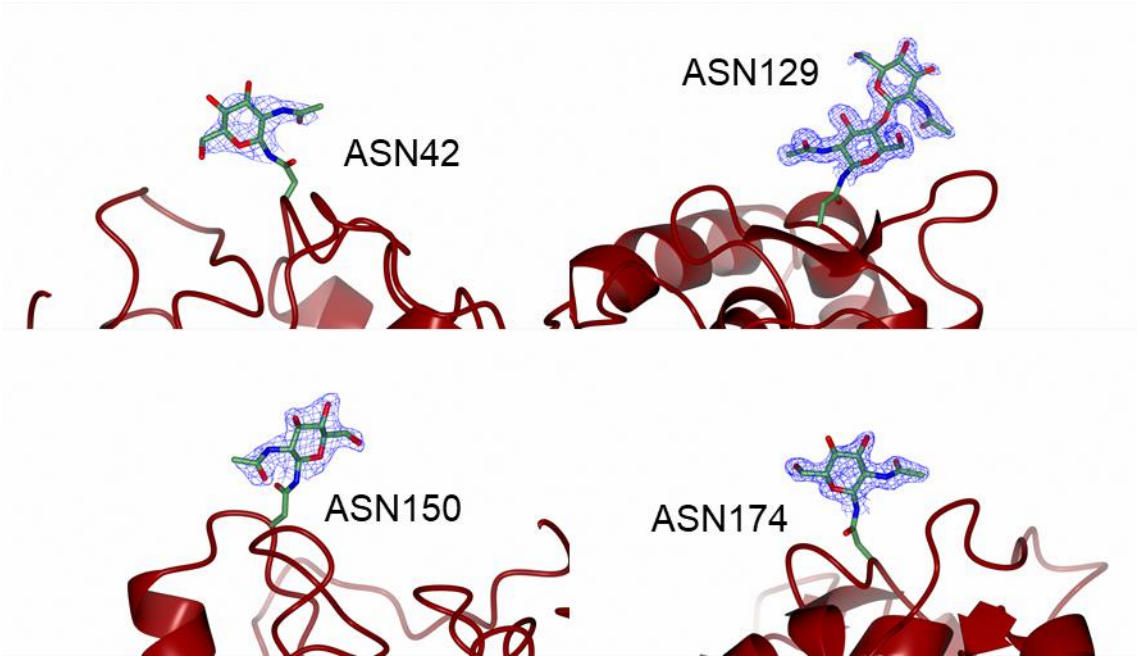


Figure 9.16 Glycosylation sites of artUPO expressed in yeast where the $2mF_O-DF_c$ map contoured to 1.0 σ is shown over the glucosamines and the linked asparagine residues are labelled. The *N*-glycosylation sites at Asn42 and Asn150 are only observed in the $artUPO_{yeast}$ 1 structure on one of the chains. Key moieties are highlighted with green carbon atoms

The cysteine (Cys24) residue, coordinated to the heme iron, is located on top of the proximal helix in the *N*-terminal half of the protein and resides between two proline residues. This Pro-Cys-Pro motif is observed in all UPOs and is also found in CPO. [86, 21] The proximal helix is believed to be stabilised through two main sets of interactions in order to stabilise the sulfur-iron interaction. [21,181] Firstly, the pairs Ser22 and Asn28, and Arg21 and Asp32 interact to stabilise the helix itself. The second interaction is through sulfur-amide hydrogen bonds where the sulfur atom of Cys24 interacts with both amide groups of Gly26 and Leu27 (Figure 9.17).

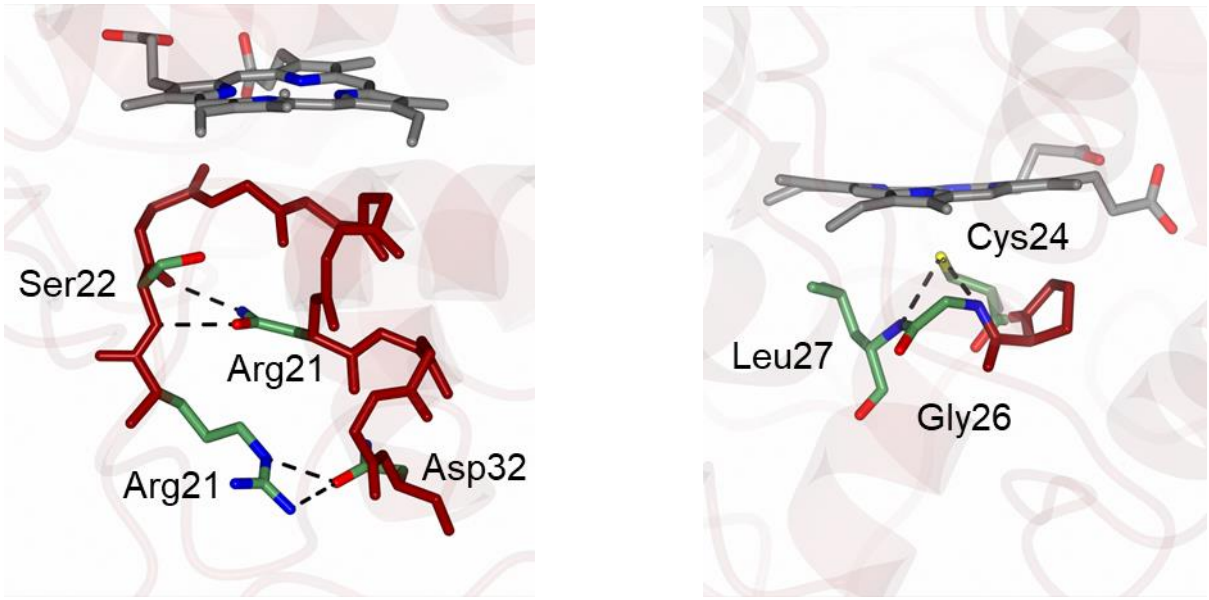


Figure 9.17 **Left**: Hydrogen bonding that occurs between the proximal helix and the previous residues for stabilisation **Right**: Sulfur-amide hydrogen bonding which stabilises the proximal cysteine residue. Key residues are highlighted and labelled in green, other residues are shown with just the backbone in red and the hydrogen bonds are shown as dashed lines

A glutamic acid (Glu164) and a histidine (His93) residue near the active site act as an charge stabilising pair in the mechanism, which are analogous to the glutamic acid and arginine pair found in *Aae*UPO (Figure 9.18). A magnesium ion also sits in a cationic site where it is coordinated to two water molecules, one of the heme propionates, Glu92, Ser96 and the carbonyl group on the backbone of His93 which forms an octahedral complex (Figure 9.18).

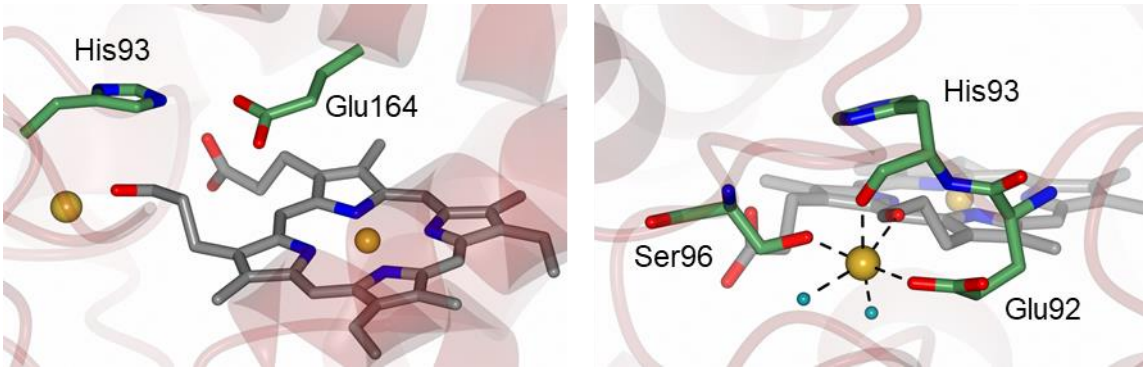


Figure 9.18 **Left**: charge stabilising pair His93 and Glu164 are located in the distal side of the heme and are involved in catalysis. **Right**: Structural Mg^{2+} ion sits near the heme group in and is coordinated with two water molecules (blue), a heme proprionate group and the residues Glu92, His93 and Ser96

The substrate channel from the solvent to the distal side of the heme consists of 17 residues (4 Ile, 4 Leu, 2 Val, 2 Ala, 2 Thr, Lys, Ser and Phe) which are predominantly aliphatic and non-polar in character (Figure 9.19). The channel is 7.9 Å long with a bottle neck of 2.7 Å located above the heme group (calculated using MOLEonline). [170] The diameter of the substrate channel is 8.66 Å (Figure 9.19).

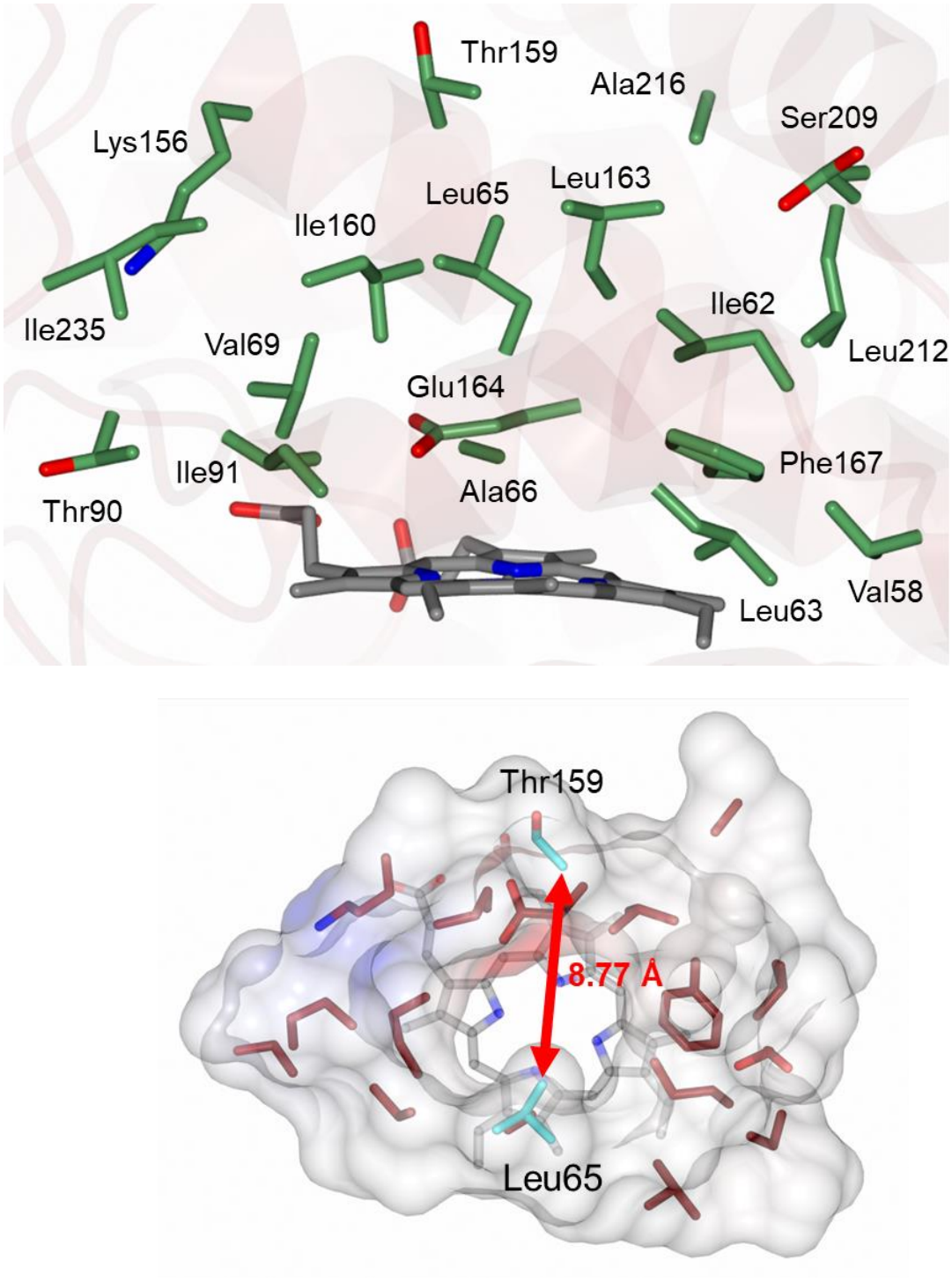


Figure 9.19 **Top**: Residues involved in the formation of the substrate channel highlighted in green **Bottom**: A bird's eye view of the substrate channel with the residues that make the tunnel highlighted in red with an electrostatic surface added to visualise the tunnel. The diameter of the entrance of the tunnel was measured using the distance between (Thr159) and (Leu65) labelled in blue which was 8.77 Å

9.2.5.5. Comparison of the structure of artUPO with other related enzymes

As artUPO is derived from *Mro*UPO with 75% sequence identity, it is expected that both enzymes will structurally be very similar. This similarity can be seen when the secondary strctures of artUPO and *Mro*UPO are superimposed as shown in Figure 9.20A. The root mean square deviation (RMSD) when both structures are superimposed is low with a value 0.57 Å over 231 common C_α atoms. *Aae*UPO, the most studied UPO, is in 'Subfamily-I' whereas *Mro*UPO (and so artUPO) is in 'Pog-family'.[86] Because of the differences in subfamilies, differences in structure and motifs are expected between artUPO and *Aae*UPO. When the structures are superimposed, artUPO and *Aae*UPO show a RMSD of 1.48 Å over 196 common C_α atoms. *Aae*UPO has two extra α-helices at the *C*-terminus compared to artUPO that sit around the top of the entrance of the substrate channel.[89] An α-helix in artUPO that is located on top of the protein (Ser209-Ala222) is replaced by a much shorter α-helix in *Aae*UPO (Glu245-Gln249), which changes the position and shape of the substrate channel (Figure 9.20B). The most recent UPO structure published to date is derived from the fungus *Hypoxylon* sp. EC38 (*Hsp*UPO) and was recombinantly expressed in *P. pastoris*.[90] *Hsp*UPO has a 38% sequence identity with artUPO and has an RMSD of 1.36 Å over 196 common C_α atoms. Like artUPO and *Mro*UPO, *Hsp*UPO is a "short" UPO and shares the same overall fold. *Cfu*CPO is a related enzyme to UPOs (EC 1.11.1.10) as it shares the same class and subclass and is currently the only other known heme thiolate protein other than UPOs. When superimposed with artUPO it has an RMSD of 1.43 Å over 201 common C_α atoms. The low RMSD values from the overlay of these structures suggest a highly conserved fold over these enzymes which indicates the importance of the globin fold to the function of these heme-containing enzymes.

Like artUPO, *Mro*UPO exists as a dimer with a disulfide bond connecting the two units. *Hsp*UPO and *Cfu*CPO are monomers, and do not contain any disulfide bonds. [21,90] *Aae*UPO exists as a dimer and contains an internal difsulfide bond between Cys278 and Cys319 in the *C*-terminal which is believed to stabilise the protein. [89,182] All the UPOs, except artUPO$_{bact}$, have *N*-glycosylation sites at asparagine residues and CPO has *O*-glycosylation sites. All enzymes have a structural cation that is coordinated with one of the heme propionates. In this cationic site, *Cfu*CPO has a manganese ion whereas the UPOs have a magnesium ion. [90, 21, 89]

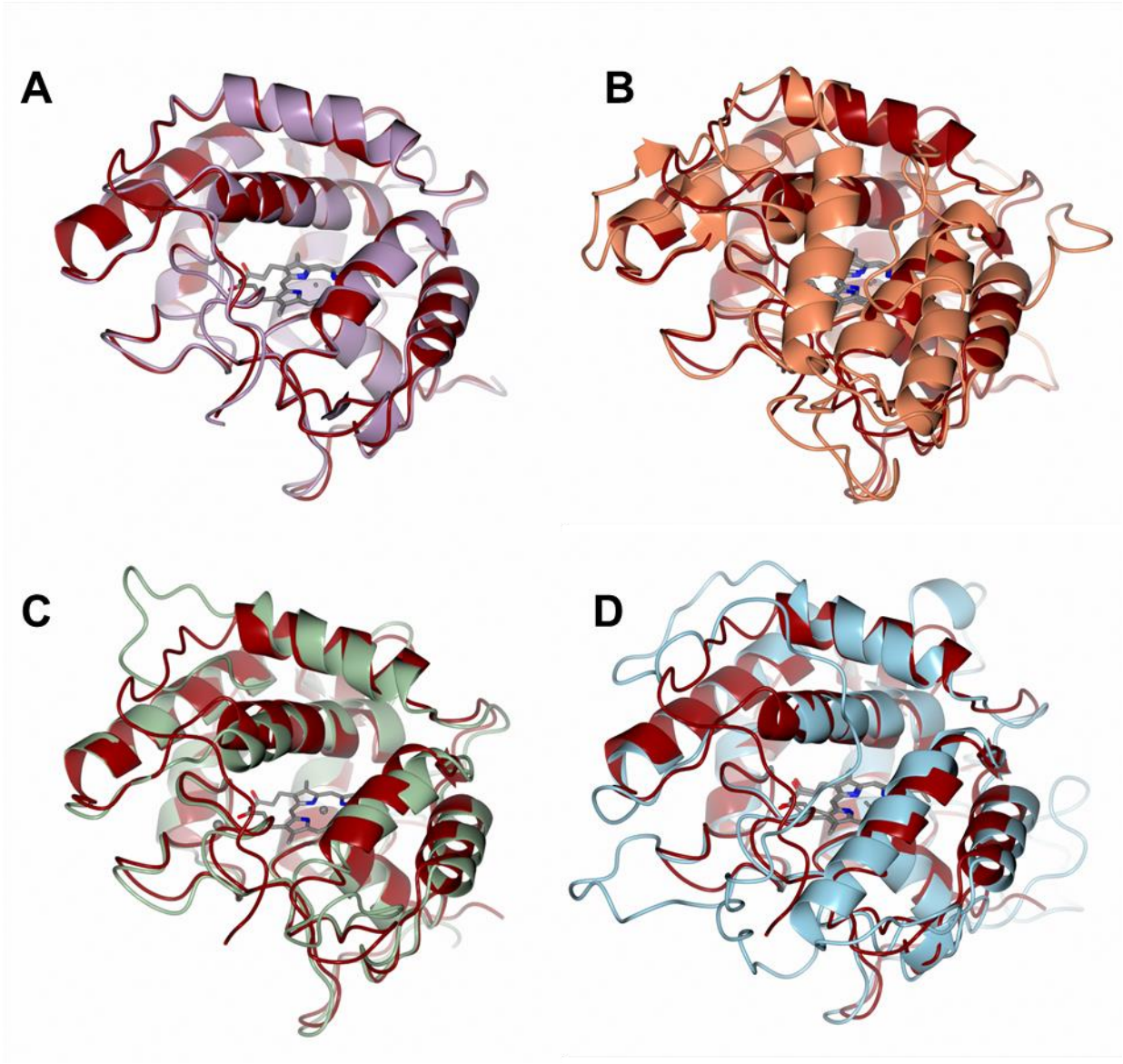


Figure 9.20 Overlay of the secondary structure of artUPO$_{yeast}$ 2 in red with **A** *Mro*UPO (5FUJ) in purple RMSD = 0.57 Å/ 231 Cα atoms **B** *Aae*UPO (2YOR) in orange RMSD = 1.48 Å/ 196 Cα atoms **C** *Hsp*UPO (7O1R) in green RMSD = 1.36 Å/ 196 Cα atoms and **D** *Cfu*CPO (2CPO) in blue RMSD = 1.43 Å/ 201 Cα atoms

The substrate channel residues directly above the heme in the active site in artUPO are mostly aliphatic in character except for Phe167. Comparing the active site of artUPO with *Mro*UPO, *Aae*UPO, *Hsp*UPO and *Cfu*CPO, a phenylalanine residue is conserved in all enzymes but the position of this residue varies in *Hsp*UPO (Figure

9.21). This is also the case with an alanine residue where it's position is conserved in all except *Hsp*UPO.

The substrate channel of *Aae*UPO has a different character to that of artUPO as most the residues on the distal side of the heme in *Aae*UPO are aromatic (Figure 9.21). This explains why *Aae*UPO can accept aromatic substrates whereas artUPO and *Mro*UPO prefer aliphatic substrates. In addition, the active site residues in artUPO and *Mro*UPO are smaller than the phenylalanine residues found in *Aae*UPO. This allows *Mro*UPO, and potentially artUPO, to accommodate larger substrates than *Aae*UPO can accept such as tetradecanedioc acid and cortisone **70**. [128, 105] The active site residues in *Cfu*CPO also differ from artUPO. A valine and isoleucine in artUPO is replaced by an isoleucine and phenylalanine respectively which may allow for the docking of the native substrate monochlorodimedone **44**. [18]

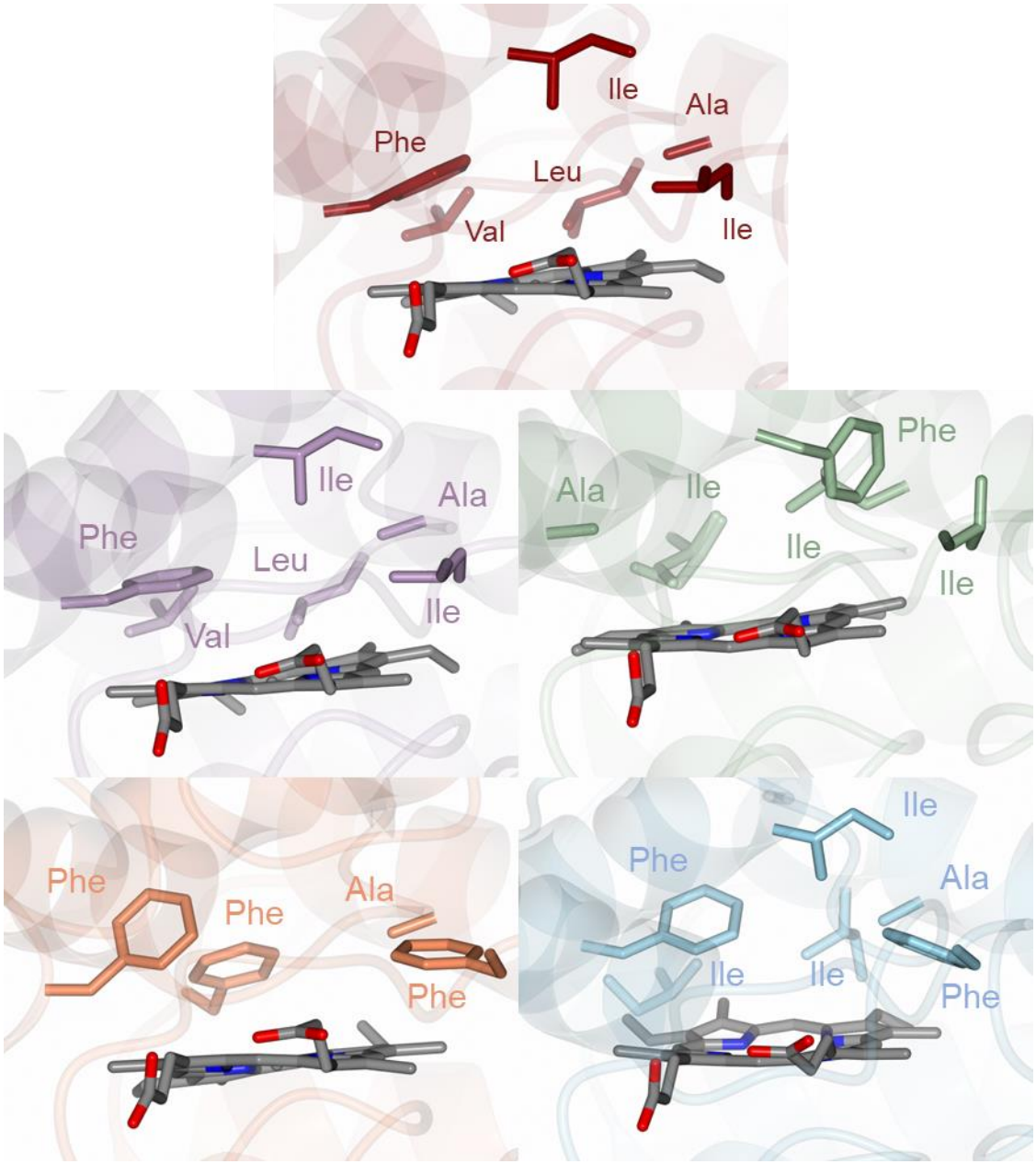


Figure 9.21 The active site resides 5.0 Å above the heme group highlighted in artUPO in red, *Mro*UPO in purple (5FUJ), *Hsp*UPO in green (7O1R), *Aae*UPO in orange (2YOR) and *Cfu*CPO in blue (2CPO). An isoleucine residue in the *Hsp*UPO structure has two conformers which is shown in this figure

The glycosylation sites of the UPOs are compared as shown in Figure 9.22.

```
AaeUPO   ------LPPGPLENSSAKLVNDEAHPWKPLRPGDIRGPCPGLNTLASHGYLPRNGVATPV   54
HspUPO   MKSLSFSLALGFGSTLVYSAPSPSSGWQAPGPNDVRAPCPMLNTLANHGFLPHDGKGITV   60
artUPO   --------------SQDIVDFSQHPWKAPGPNDLRSPCPGLNTLANHGFLPRNGRNITI   45
MroUPO   ----------------------SAHPWKAPGPNDSRGPCPGLNTLANHGFLPRNGRNISV   38

AaeUPO   QI-INAVQEGLNFDNQAAVFATYAAHLVDGNLITDLLSIGRKTRLTGPDPPPPASVGGLN   113
HspUPO   NKTIDALGSALNIDANLST---------------LLFGFAAT-TNPQP-NATFFDLDHLS   103
artUPO   PMIVQAGFDGYNVQPDILI---------------LAAKVGL---LTSP-EPDTFTLDDLK   86
MroUPO   PMIVKAGFEGYNVQSDILI---------------LAGKIGM---LTSR-EADTISLEDLK   79

AaeUPO   EHGTFEGDASMTRGDAFFGNNHDFNETLFEQLVDYSNRFGGGKYNLTVAGELRFKRIQDS   173
HspUPO   RHNILEHDASLSRQDSYFGPADVFNEAVFNQTKSFW--TG-DIIDVQMAANARIVRLLTS   160
artUPO   LHGTIEHDASLSREDFALGDNLHFNEAIFNTLANSN--PGSDVYNITSAGQVLKDRLADS   144
MroUPO   LHGTIEHDASLSREDVAIGDNLHFNEAIFTTLANSN--PGADVYNISSAAQVQHDRLADS   137

AaeUPO   IATNPNFSFVDFRFFTAYGETTFPANLFVDGRRDDGQLDMDAARSFFQFSRMPDDF--FR   231
HspUPO   NLTNPEYSLSDLGSAFSIGESAAYIGILG--DKKSATVPKSWVEYLFENERLPYELGFKR   218
artUPO   LARNPNVTNTGKEFTIRTLESAFYLSVMG--NATTGEAPKNFVQIFFREERLPIEEGWKR   202
MroUPO   LARNPNVTNTDLTATIRSSESAFFLTVMSAGDPLRGEAPKKFVNVFFREERMPIKEGWKR   197

AaeUPO   APSPRSGTG----VEVVIQAHPMQPGRNVGKINSYTVDPTSSDFSTPCLMYEKFVNITVK   287
HspUPO   PNDPFTTDDLGDLSTQIINAQHFPQSPG--KVEK--------RGDTRCPY----------   258
artUPO   STTPITSDTLNPIAGQISEASNWKPNPD------------------QCPW----------   234
MroUPO   STTPITIPLLGPIIERITELSDWKPTGD------------------NCGA----------   229

AaeUPO   SLYPNPTVQLRKALNTNLDFFFQGVAAGCTQVFPYGRD   325
HspUPO   -GYH-----------------------------------   261
artUPO   -IVLSPNL-------------------------------   241
MroUPO   -IVLSPEL-------------------------------   236
```

Figure 9.22 Alignment of *Aae*UPO, *Hsp*UPO, artUPO and *Mro*UPO amino acid sequences. The glycosylation sites of each enzyme are highlighted in yellow

It is important to note that for all UPO x-ray crystallography structures, the extent of glycosylation may not be a true representation of the actual glycosylation on each protein due to the flexible nature of sugars. However, it is still useful to compare what sites can be observed.

In all UPOs, *N*-linked glycosylation is observed through asparagine residues. Conservation of glycosylation sites between *Aae*UPO, *Hsp*UPO and *Mro*UPO appears to be low, indicating that the exact position of glycosylation may not be essential. It was somewhat expected that artUPO and *Mro*UPO would share some of the same glycosylation sites due to the similarity of these enzymes as highlighted in Figure 9.22.

However, a key difference is observed in artUPO at residue Arg174. In artUPO, the sequence NAT, a consensus *N*-glycosylation sequence, results in glycosylation at the asparagine residue.[183] However, the sequence for the corresponding area in *Mro*UPO is completely different (PLR). This change suggests that an *N*-linked glycosylation site was introduced to potentially increase the stability of artUPO compared to *Mro*UPO. In artUPO at residue Arg150, the asparagine residue is glycosylated but the corresponding residue, Arg143, in *Mro*UPO is not glycosylated. This does not necessarily imply that there is no glycosylation in this site, however this could be examined by mass spectrometry.

Table 9.5 summarises the discussion of the comparison between the UPOs and CPO and includes the sequence identity to artUPO, backbone alignment and active site residues on the distal side of the heme.

Table 9.5 Comparison of artUPO with *Mro*UPO, *Aae*UPO, *Hsp*UPO and *Cfu*CPO, where the sequence identity, backbone alignment and the active site residues are summarised

Enzyme	Sequence identity /%	Backbone alignment (RMSD) / Å/C_α	Active site residues
artUPO	100.00	-	F, V, I, L, A, I
*Mro*UPO	75.00	0.57/231	F, V, I, L, A, I
*Aae*UPO	35.78	1.48/196	F, F, A, F
*Hsp*UPO	38.30	1.35/213	A, I, F, I, I
*Cfu*CPO	25.35	1.43/201	F, I, I, I, A, F

9.2.6. artUPO and *Mro*UPO

artUPO was developed from *Mro*UPO by NovoZymes pressumably to improve *Mro*UPO's function and/or stability. By aligning the sequences, the similarities and differences can be highlighted (Figure 9.23).

```
artUPO    SQDIVDFSQHPWKAPGPNDLRSPCPGLNTLANHGFLPRNGRNITIPMIVQAGFDGYNVQP   60
MroUPO    -------SAHPWKAPGPNDSRGPCPGLNTLANHGFLPRNGRNISVPMIVKAGFEGYNVQS   53
                 * ********** *.*********************::****:***:*****

artUPO    DILILAAKVGLLTSPEPDTFTLDDLKLHGTIEHDASLSREDFALGDNLHFNEAIFNTLAN   120
MroUPO    DILILAGKIGMLTSREADTISLEDLKLHGTIEHDASLSREDVAIGDNLHFNEAIFTTLAN   113
          ******.*:*:*** * **::*:******************.*:***********.****

artUPO    SNPGSDVYNITSAGQVLKDRLADSLARNPNVTNTGKEFTIRTLESAFYLSVMG--NATTG   178
MroUPO    SNPGADVYNISSAAQVQHDRLADSLARNPNVTNTDLTATIRSSESAFFLTVMSAGDPLRG   173
          ****:*****:**.** :****************.   ***: ****:*:**.  :   *

artUPO    EAPKNFVQIFFREERLPIEEGWKRSTTPITSDTLNPIAGQISEASNWKPNPDQCPWIVLS   238
MroUPO    EAPKKFVNVFFREERMPIKEGWKRSTTPITIPLLGPIIERITELSDWKPTGDNCGAIVLS   233
          ****:**::******:**:***********   *.**  :*:* *:***. *:*  ****

artUPO    PNL   241
MroUPO    PEL   236
          *:*
```

Figure 9.23 Amino acid sequence alignment of artUPO and *Mro*UPO using Clustal Omega. Below the sequences, amino acids that are shared between the two proteins are shown as a *, amino acids that are not conserved but share very similar properties are shown as : and amino acids that have some similarity are shown as . Alpha helices and beta strands are highlighted in blue and purple respectively.

artUPO has 75% identity with *Mro*UPO and so most of the artificial protein is the same as the wild type. However, significant differences in the amino acid sequence between the two UPOs are seen. Firstly near the entrance of the substrate channel, Leu149 and Ser156 in *Mro*UPO are substituted with Lys156 and Leu163 in artUPO respectively (Figure 9.24). In each of these substitutions, the polarity of the residue changes where the first substitution changes from non-polar to polar, and the second

changes from polar to non-polar. The potential reasoning for these mutations could be to change or improve the substrate specificity of *Mro*UPO, however further mutagenesis studies would need to be carried out to investigate the effects of these substitutions.

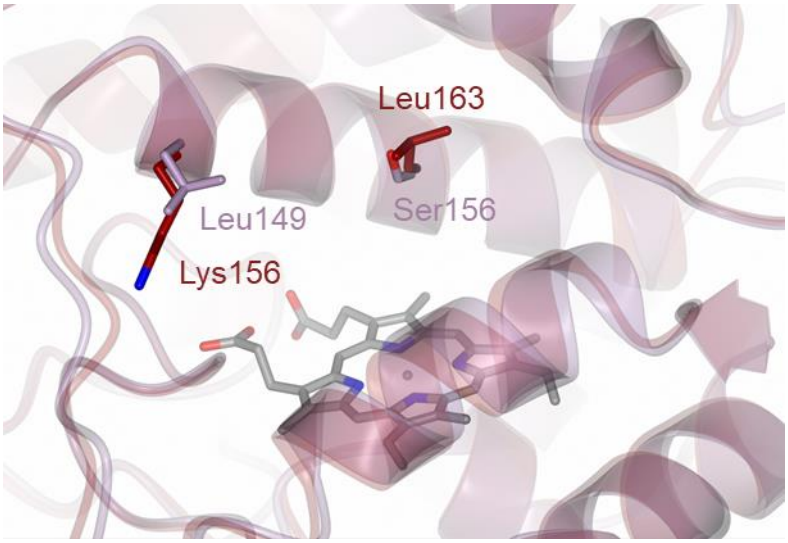


Figure 9.24 Mutations in the substrate channel, where residues Leu149 and Ser156 in *Mro*UPO are replaced with Lys156 and Leu163 respectively

Many of the other mutations found in artUPO are located away from the active site which suggests that many of the changes made by NovoZymes may not have been directly to modulate the activity, but may be designed to improve the stability of the enzyme. Examples of potentially stabilising mutations include some of the residues on the surface of the protein being mutated from hydrophobic to hydrophilic residues, such as changing leucine residues to lysine, threonine or serine (some examples seen in Figure 9.25A and 9.25B). More examples of potential stabilising substitutions include an alanine being mutated to a phenylalanine which may stabilise a neighbouring tryptophan residue (Figure 9.25C), and an alanine mutated to a

tryptophan, located near the Cys232 disulfide bond, may aid in stabilising the dimer interface (Figure 9.25D).

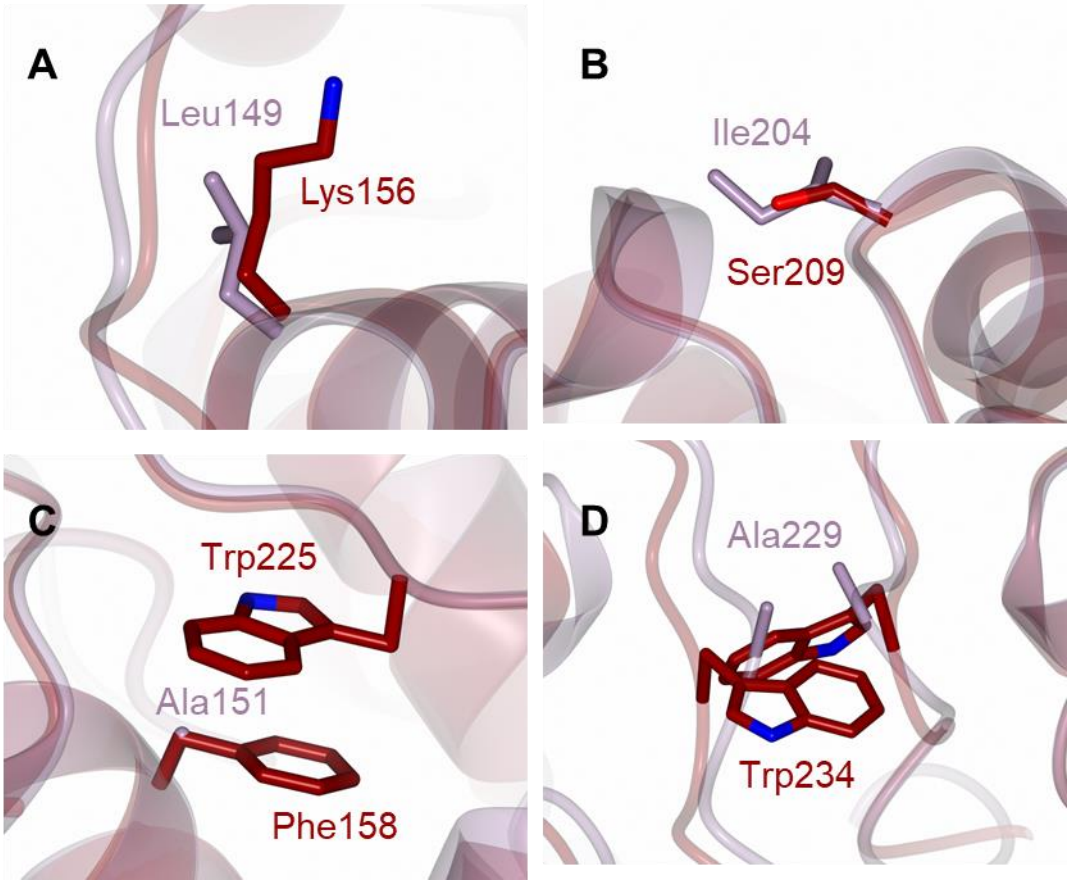


Figure 9.25 Examples of mutations that were potentially carried out to create artUPO (dark red) from *Mro*UPO (purple) where **A** and **B** are examples of non-polar residues found on the surface of the protein are changed to polar residues in artUPO, **C** change from alanine to phenylalanine which may interact with a tryptophan residue in the protein and **D** change from alanine to tryptophan in the dimer interface, the tryptophan residues may interact and aid in stabilisation

As well as point mutations, a difference between artUPO and *Mro*UPO is seen in a loop located before the short pair of β-sheets (Figure 9.26). In artUPO, this flexible loop is shorter (Met172 – Thr177) than the analogous loop in *Mro*UPO (Met165 –

Arg172). The reason for shortening this loop could potentially reduce the flexibility of the UPO and hence may increase stability. Extensive mutagenesis studies on the flexible loops of *Aae*UPO mutant PaDa-I have been carried out by the Alcade group. [91,111,176] These studies have revealed that the both the activity and stability of the UPO can be affected by point mutations in the loops, and so these results may have been a basis for the modification of the loop by NovoZymes.

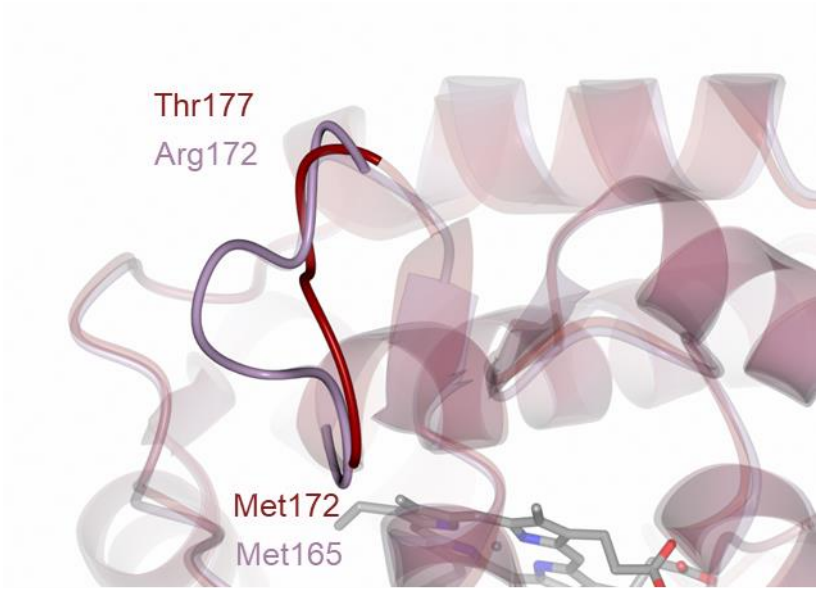


Figure 9.26 a loop between an α-helix and a short β-sheet in artUPO (Met172 – Thr177) is shorter than the corresponding loop in *Mro*UPO (Met165 – Arg172), which may increase the stability of artUPO

Differences between *Mro*UPO and artUPO can also be observed from not just comparing the changes in amino acid sequence, but also comparing the 3D structures. Figure 9.27 shows, in both artUPO and *Mro*UPO, the surface area of one chain (chain A) being buried by the other chain (chain B). The higher value of surface area buried by artUPO compared to *Mro*UPO suggests a greater stability in dimer interface due to the changes in interface residues. [184,185]

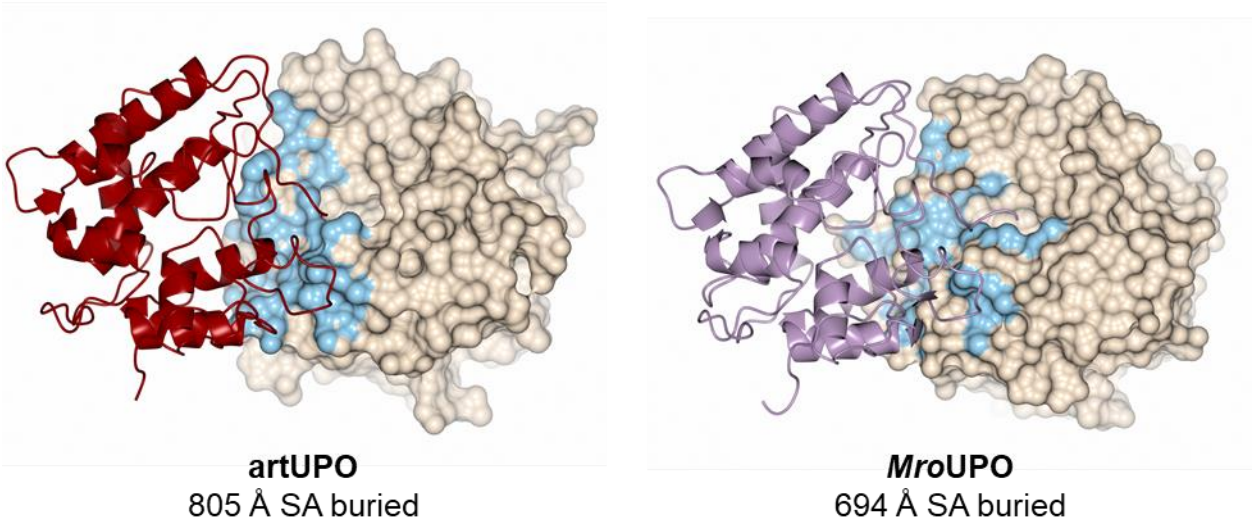


Figure 9.27 Overall structures of **left:** artUPO (red chain B) and **right:** *Mro*UPO (purple chain B) showing the surface area of chain A (cream surface) being buried (blue surface) by chain B (shown in respective colours and ribbons)

Ideally experiments would be carried out to investigate the changes in stabilisation from *Mro*UPO to artUPO such as NanoDSF.

9.2.6.1. artUPO and AlphaFold 2

In the last few months of the project AlphaFold 2 was released. [172,186] AlphaFold 2, developed by DeepMind and EMBL-EBI, uses a deep learning AI algorithm to highly accurately predict the structure of proteins based upon a training set of structures that have already been deposited in the protein data bank or are publicly disclosed. To investigate AlphaFold's capabilities, the amino acid sequence of artUPO was inputted, stating it was a dimer, and several structures were generated. The structure with the highest confidence values was selected and is shown in Figure 9.28. (Other outputs such as the predicted aligned error (PAE) plot are shown in the appendix).

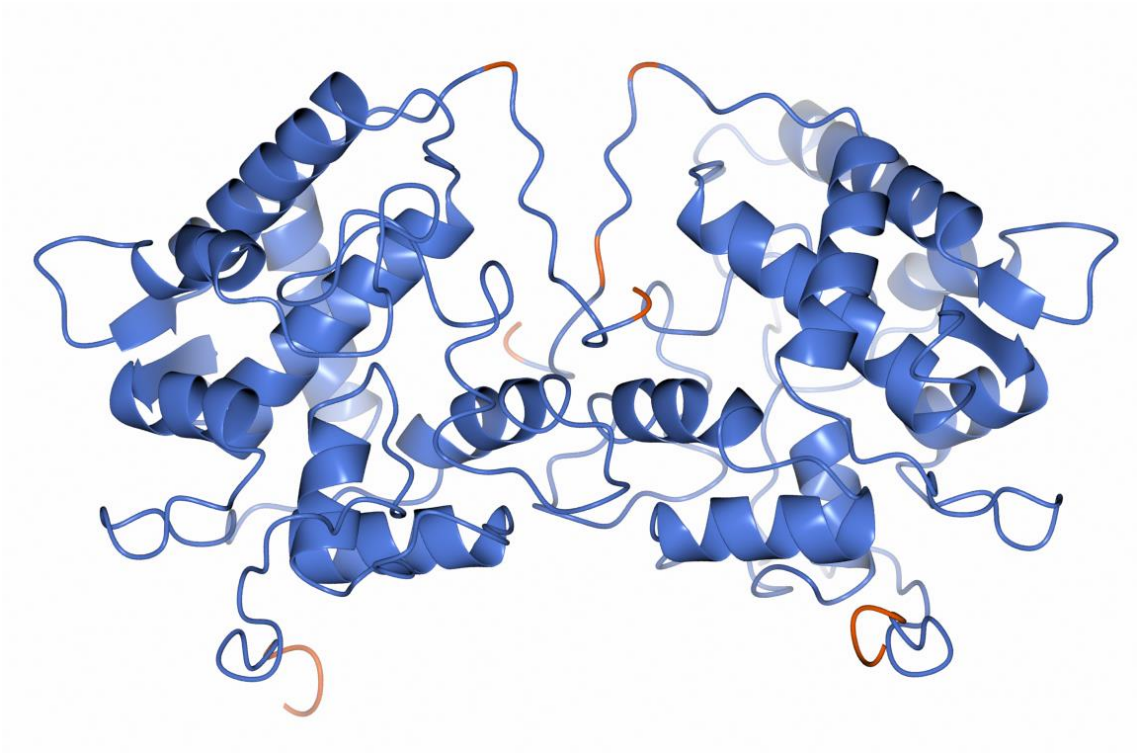


Figure 9.28 Structure of artUPO generated by AlphaFold 2 coloured by pLDDT score where dark blue is pLDDT > 90, light blue is pLDDT > 70, yellow is pLDDT > 50 and red is pLDDT < 50. Max PAE = 31.75, PTM = 0.93

The structure generated by AlphaFold has a score for each residue called a pLDDT (predicted local distance difference test) which indicates the confidence of the residue in the model. As shown in Figure 9.28, most of the model has a pLDDT score of >90 which indicates high confidence of the model. When superimposed with the crystal structure of artUPO$_{bact}$, the SSM alignment has a very low RMSD of 0.47 Å over 470 common C$_\alpha$ atoms, and the least squares fit RMSD is 0.89 Å over 1760 common heavy atoms (all atoms except hydrogen). These values demonstrate a highly accurate prediction of the structure which shows the power of AlphaFold predictions to aid in structure guided evolution. The non-natural nature of artUPO does however present a problem in using AlphaFold, which was not trained to predict structures of unnatural proteins. This problem can be seen when comparing the orientation of the side chain of Trp234 (Figure 9.29).

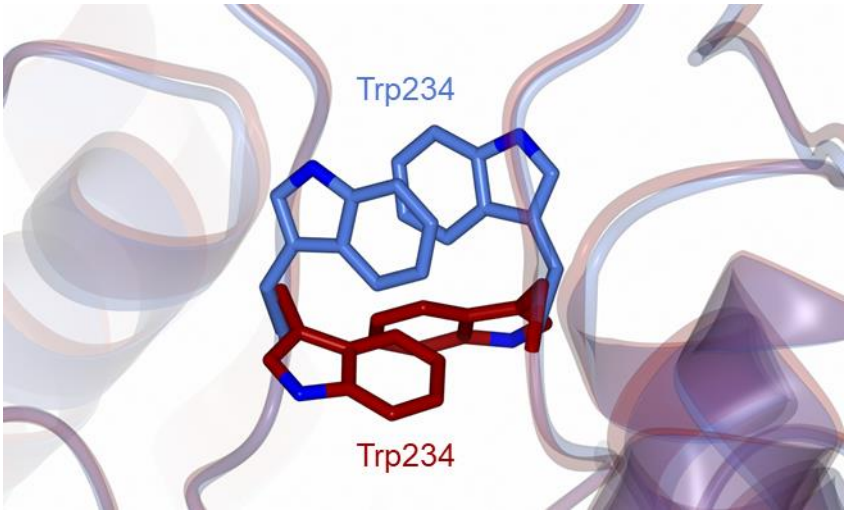


Figure 9.29 Differences in the orientation of residue Trp234 in **red**: protein x-ray crystal of artUPO$_{bact}$ and **blue**: AlphaFold 2 model of artUPO

In this example, the orientation of the tryptophan residues in the artUPO dimer interface in the AlphaFold model does not align with the tryptophan residues in the crystal structure of artUPO. This demonstrates that AlphaFold, although can be used as a guide for structure based evolution, should not be used to predict the effect of point mutations. The reliance of AlphaFold on multiple sequence alignments means that it is not appropriate to be used for predicting the effects of mutations. Other issues that arose from the prediction included a missing disulfide bridge between the two units at Cys232 and missing heme groups.

9.3. Summary of the biochemical characterisation of artUPO derived from yeast and bacteria

Purified artUPO$_{bact}$ and artUPO$_{yeast}$ was used to determine the Michaelis-Menten constants using substrates to determine the peroxidative activity and the peroxygenative activity. A summary of these experiments can be found in Table 9.2. Overall, both forms of artUPO performed similarly as the catalytic turnovers for peroxidative and peroxygenative reactions were within the same order of magnitude. artUPO$_{bact}$ however did have higher peroxidative activity with ABTS than with DMP which were in buffer and 10% ACN respectively. This may suggest that the peroxidative activity of artUPO$_{bact}$ is more sensitive to solvent. Compared to *Mro*UPO and *Aae*UPO, both artUPO forms were less active in both types of activity.

NanoDSF experiments revealed that the T_m for artUPO$_{bact}$ and artUPO$_{yeast}$ were 47.8 °C and 45.5 °C respectively, however the fluorescence curve of artUPO$_{yeast}$ suggested that a second state of unfolding may occur. Further investigation with other UPOs may provide information about the thermal stability of UPOs and how they could be potentially used in an industrial setting. Through analytical SEC and treatment with dithiotheitol, it was revealed that both artUPOs were dimers in solution and the units were joined together by a disulfide bond. These results were also seen in the x-ray crystal structures of artUPO$_{bact}$ and artUPO$_{yeast}$. For all the crystals, data was collected at Diamond Light Source and were processed using either Xia2 or Xia2 3dii. The structures were solved using molecular replacement with either *Mro*UPO (5FUJ) as a model or artUPO$_{yeast}$ as a model for artUPO$_{bact}$. Two structures were obtained for artUPO$_{yeast}$, both of these were formed using a seed stock that were grown in the

conditions 0.1 M KSCN, 30% PEG MME 2000 and addition of 80 mg mL^{-1} $artUPO_{yeast}$. The first $artUPO_{yeast}$ crystal ($artUPO_{yeast}$ 1) grew in 0.1 M tryptone, 1 mM NaN_3, 0.05 M HEPES sodium pH 7.0, 20% PEG 3350 with 60 mg mL^{-1} and the second ($artUPO_{yeast}$ 2) grew in 0.15 M KSCN, 25% PEG MME 2000. $artUPO_{yeast}$ 1 was resolved to 2.01 Å in the space group *P* $2_12_12_1$ and had R_{factor} and R_{free} values of 19% and 23% respectively. $artUPO_{yeast}$ 2 was resolved to 1.21 Å in the space group *C* 2 2 2_1 and had R_{factor} and R_{free} values of 16% and 18% respectively. $artUPO_{bact}$ crystals were grown in 0.1 M HEPES pH 7.5, 25% PEG 3350 in initial screens without optimisation with the addition of 15 mg mL^{-1} of enzyme. The crystal resolved to 2.17 Å in the space group *P* $2_12_12_1$ and had R_{factor} and R_{free} values of 21% and 26% respectively.

artUPO is a globular protein made of eight α-helices and a pair of β-sheets that surround a heme group. Features such as a His-Glu stabilising pair and a structural Mg^{2+} ion were observed which is also seen in *Mro*UPO. Glycosylation was found in both $artUPO_{yeast}$ structures where GlcNAc moieties were covalently bonded to arginine residues, but the level of glycosylation was not identical in each structure. $artUPO_{bact}$, as expected, did not have any glycosylation. The substrate channel of artUPO is very similar to *Mro*UPO where most of the residues were non-polar and aliphatic in character. Key differences between artUPO and *Mro*UPO were observed, which included two mutations in the substrate channel to potentially improve substrate recognition, and other mutations on the surface of artUPO which improve the stability of the enzyme. The next stage for $artUPO_{bact}$ and $artUPO_{yeast}$ was to investigate their catalytic capabilities for potential use in industry and to potentially make a judgement about which expression system is more applicable for further use.

10. Characterisation of the Catalytic Capabilities of artUPO

In order to assess the potential of artUPO for industrial applications, the substrate scope of the enzyme was investigated using achiral and chiral GC. This will allow for identification of the types of compounds that can be accessed by artUPO and will hopefully expand the range of compounds that can be accessed through biocatalytic methods. Of particular interest is the potential of UPOs to carry out stereoselective oxidations with high *ees* due to the challenge of producing enantiopure compounds using traditional chemistry methods. [2,10] The stereoselectivity of artUPO$_{yeast}$ was also investigated. In order to better compare artUPO to previously studied UPOs, the substrate screens were also carried out alongside with r*Aae*UPO, which is already characterised. [95] This comparison may allow for a better assessment of the potential industrial applications of artUPO. In this chapter, three reaction types were investigated: hydroxylations, sulfoxidations and epoxidations.

A small initial screen was carried out to probe the reaction capabilities of the enzyme. This screen was then expanded into a much larger screen that investigated the effect of phenyl substituents on the hydroxylation, sulfoxidation and epoxidation activity of artUPO$_{yeast}$. Optimisation of reaction conditions were carried out, including substrate loading and reaction time. Time courses were then run over 6 h to monitor the formation of oxidation products over time. Biotransformations were also carried out using artUPO$_{bact}$ in order to assess the effects of expression system on the activity of the enzyme. This work identifies a potential niche for artUPO in biocatalysis and provides inspiration for future work to be carried out with artUPO.

10.1. Materials and Methods

10.1.1. Materials

Chemicals and other molecular biology materials were purchased from Alfa Aesar, Fisher Scientific, Fluorochem, Generon, Merck Chemicals Ltd, Millipore, New England Biolabs, Qiagen, Scientific Laboratory Supplies Ltd, Takara Bio, Tokyo Chemical Industry and VWR International.

10.1.2. Compound key

The following figures and tables describe the compounds used and the numbers that have been assigned to them.

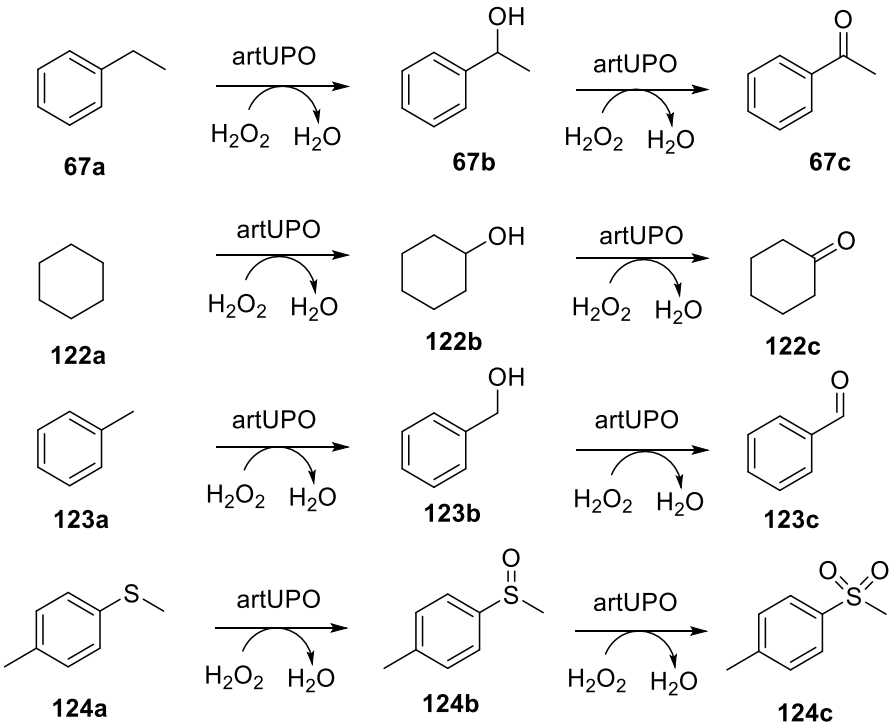
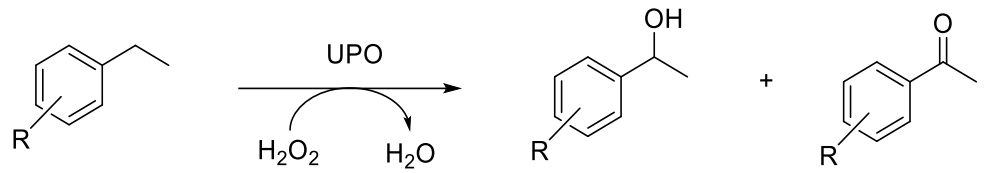
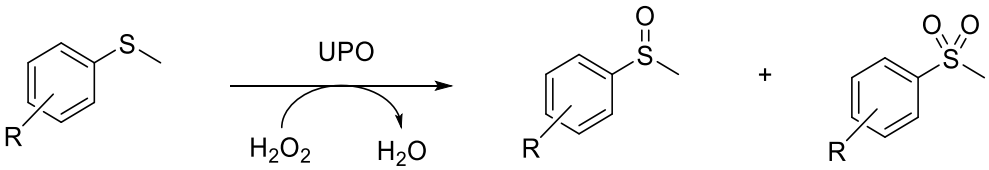


Figure 10.1 Compounds in initial biotransformation screens, involving ethylbenzene, **67a**, cylcohexane, **122a**, toluene, 1**23a** and methyl *p*-tolyl sulfide, **124a**



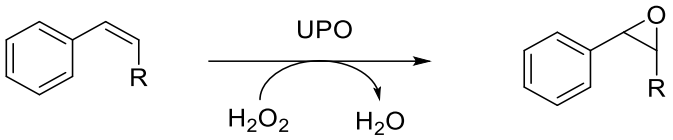
Compound	R	Alcohol	Ketone
67a	H	**67b**	**67c**
125a	2-Me	**125b**	**125c**
126a	3-Me	**126b**	**126c**
127a	4-Me	**127b**	**127c**
128a	4-F	**128b**	**128c**
129a	4-Cl	**129b**	**129c**
130a	4-OMe	**130b**	**130c**
131a	2-OMe	**131b**	**131c**
132a	3-OMe	**132b**	**132c**
133a	4-CN	**133b**	**133c**

Figure 10.2 Hydroxylation screen of alkylbenzenes catalysed by UPOs



Compound	R	Sulfoxide	Sulfone
134a	H	**134b**	**134c**
135a	2-Me	**135b**	**135c**
136a	3-Me	**136b**	**136c**
124a	4-Me	**124b**	**124c**
137a	4-F	**137b**	**137c**
138a	4-Cl	**138b**	**138c**
139a	4-OMe	**139b**	**139c**
140a	4-CN	**140b**	**140c**

Figure 10.3 Sulfoxidation screen of thioethers catalysed by UPOs



Compound	R	Epoxide
141a	H	**141b**
88a	Me	**88b**

Figure 10.4 Epoxidation screen of styrene **141a**, and *cis*-β-methylstyrene, **88a**

10.1.3. Initial Biotransformations with artUPO$_{yeast}$

Reactions with **67a**, **122a**, **123a** and **124a** were carried out on a 2 mL scale with 50 mM potassium phosphate buffer pH 7.0, 200 µL artUPO$_{yeast}$ supernatant from the *P. pastoris* fermentation (1.5 U mL^{-1} NBD) and 10 mM substrate. Prior to use, the substrates **67a**, **122a** and **123a** were dissolved in acetonitrile (ACN) and **124a** was dissolved in ethanol (EtOH), and added to the buffer to give a final solvent concentration of 10% ACN (*v*/*v*) or 1% EtOH (*v*/*v*). The reactions were started with the addition of 2 mM H_2O_2 and was sequentially added every 30 min until 10 mM H_2O_2 was reached. The reactions were incubated at 25 ºC with a stirrer bar at 300 rpm. When taking time points, 200 µL of the reaction was taken and extracted with 200 µL of ethyl acetate for analysis using GC (see Section 10.1.6). Time points were taken at 0, 1, 2 and 24 h after the start of the reaction.

To investigate the effect of stirring in the reaction, the conditions as described above at a 2 mL scale were used, investigating hydroxylations with **67a** and sulfoxidations with **124a**. The reactions were mixed with either a stirrer bar at 500 rpm or shaken at

700 rpm. The effect of room temperature was also investigated, and the reactions were run at either 20 °C or 25°C. Time points were taken at 0, 1, 2 and 24 h after the start of the reaction and analysed using GC (see Section 10.1.6).

10.1.4. Biotransformation screens on a 2 mL scale

Reactions with substrates **67a, 125a – 130a**, **124a**, **134a – 139a** were carried out on a 2 mL scale with 50 mM potassium phosphate buffer pH 7.0, either artUPO$_{yeast}$ (0.5 U mL^{-1} NBD) from the supernatant of the *P. pastoris* fermentation or lyophilised r*Aae*UPO (0.5 U mL^{-1} NBD) from a fermentation as described in a previous project carried out by Mielke, [95] 10 mM substrate in either 10% ACN with **67a**, **125a – 130a** or 1% EtOH for **124a, 134a – 139a**. For the reactions with styrene derivatives **88a** and **141a**, the compounds were dissolved in ACN and added to the buffer to give a final substrate concentration of 1 mM and 10% ACN (*v*/*v*). The reactions were initiated with the addition of 2 mM H_2O_2 for reactions with compounds **67a**, **124a – 130a, 134a – 139a** or 1 mM H_2O_2 for reactions with compounds **88a** and **141a**. The reactions were incubated at 20 °C at 700 rpm in a shaker for 2 h. A 200 µL sample at the end was taken and was extracted with 200 µL ethyl acetate. The organic layer was dried with anhydrous $MgSO_4$ and analysed by GC.

10.1.5. Biotransformation screens on a 5 mL scale

Reactions with substrates **67a, 88a**, **125a – 133a**, **124a**, **134a – 141a** were carried out on a 5 mL scale in 50 mM potassium phosphate buffer pH 7.0. Three different UPOs were used with the same activity (0.5 U mL^{-1} NBD), either $artUPO_{yeast}$, an enriched $artUPO_{bact}$ sample purified by NiNTA (see Section 8.1.4) or r*Aae*UPO as described in the previous section. The starting concentration of the starting materials for the 5 mL reactions were either 2 mM or 10 mM and were solubilised in either 10% ACN for **67a**, **88a**, **125a – 133a**, **140a** and **141a** or 1% EtOH for **124a**, **134a – 140a**. The reactions were initiated with 2 mM H_2O_2 and was sequentially added every 30 min until 10 mM H_2O_2 was reached. 300 µL samples were taken 0.0 h, 0.5 h, 1.0 h, 1.5 h, 2.0 h, 4.0 h and 6.0 h after the first addition of 2 mM H_2O_2. Reactions were shaken at 250 rpm in an orbital shaker. These samples were extracted with 300 µL ethyl acetate, dried with $MgSO_4$ and analysed by GC.

For chiral analysis of the reactions, 5 mL scale reactions were carried out in the same way as the time courses with the same addition times for H_2O_2. Samples were taken at the beginning of the reaction and 24 h after. The reactions were extracted with 5 mL of ethyl acetate, dried using anhydrous $MgSO_4$, were concentrated to roughly 100 µL and were analysed using GC.

10.1.6. GC analysis

Samples for biotransformations with compounds **67**, **88**, **122** – **141** were analysed using an Agilent Technologies 789B GC system. Achiral GC was carried out with a HP-5ms column (30 m × 0.25 mm internal diameter, 0.25 μm film thickness) from Agilent J&W and using hydrogen as a carrier gas at 1.2 mL min^{-1}. The inlet was heated to 250 °C. The oven temperatures and the split used for each method used are described in Table 10.1.

Samples for the chiral analysis of the alcohols in Figure 10.2 were analysed using a BetaDex 120 column (30 m × 0.25 mm internal diameter, 0.25 μm film thickness) from Supelco and using hydrogen as a carrier gas at 1.2 mL min^{-1}. The inlet was heated to 250 °C. The oven temperature and the split used for this method used is described in Table 10.1. Samples for the chiral analysis of the sulfoxides in Figure 10.3 were analysed using a BGB-175 column (30 m × 0.25 mm internal diameter, 0.25 μm film thickness) from BGB Analytik using hydrogen as a carrier gas at 1.2 mL min^{-1}. The inlet was heated to 250 °C. The oven temperature and the split used for this method used is described in Table 10.1. To determine absolute configuration of the oxidised compounds, literature of *Aae*UPO biotranformations with either the same or very similar substrates have stated that these reactions occur with a preference to form the *R* enantiomer of both alcohols and sulfoxides. [95,127,144,187,188] Other analysis of similar compounds in the literature by chiral GC was also examined to assign the configuration of each enantiomer in the screens. [189,190]

Table 10.1 GC methods for achiral and chiral analysis of biotransformations

Method name	Type of analysis	Method	Split	Column used	Compounds used for method
1	Achiral	100 °C for 2 min, 20 °C min^{-1} to 160 °C, 40 °C min^{-1} to 220 °C	20:1	HP-5ms	**67, 88, 125 – 133, 141**
2	Achiral	150 °C for 7 min, 20 °C min^{-1} to 220 °C, hold for 3 min	10:1	HP-5ms	**124, 134-140**
3	Achiral	40 °C for 1 min, 30 °C min^{-1} to 230 °C, hold for 2 min	20:1	HP-5ms	**122**
4	Achiral	100 °C for 2 min, 10 °C min^{-1} to 120 °C, 20 °C min^{-1} to 160 °C, 30 °C min^{-1} to 200 °C, hold for 1 min	20:1	HP-5ms	**123**
5	Chiral	120 °C for 35 min, 40 °C min^{-1} to 220 °C		Betadex 120	**67b, 125b – 133b**
6	Chiral	180 °C for 15 min, 40 °C min^{-1} to 220 °C		BGB-175	**124b, 134b – 140b**
7	Chiral	90 °C for 30 min, 40 °C min^{-1} to 220 °C	10:1	Cyclodex B	**88b, 141b**

The retention times for each compound in both achiral and chiral GC analysis are given in the tables below.

Table 10.2 Achiral GC analysis of compounds in biotransformations and methods used

Entry no.	Compound no.	Compound	Method	Retention time / min
1	**122a**	cyclohexane	3	1.68
2	**123a**	toluene	4	1.95
3	**67a**	ethylbenzene	1	1.83
4	**125a**	2-ethyltolune	1	2.53
5	**126a**	3-ethyltolune	1	2.39
6	**127a**	4-ethyltoluene	1	2.40
7	**128a**	1-ethyl-4-fluorobenzene	1	1.89
8	**129a**	1-chloro-4-ethylbenzene	1	3.11
9	**130a**	4-ethylanisole	1	3.59
10	**131a**	2-ethylanisole	1	3.04
11	**132a**	3-ethylanisole	1	3.27
12	**133a**	4-ethylbenzonitrile	1	4.06
14	**123b**	benzyl alcohol	4	3.69
15	**67b**	phenylethanol	1	3.12
16	**125b**	1-(2-methylphenyl)ethanol	1	4.11
17	**126b**	1-(3-methylphenyl)ethanol	1	3.95
18	**127b**	1-(4-methylphenyl)ethanol	1	3.97
19	**128b**	4’-fluorophenylethanol	1	3.23
20	**129b**	4’-chlorophenylethanol	1	4.79
21	**130b**	2’-methoxyphenylethanol	1	4.68
22	**131b**	3’-rethoxyphenylethanol	1	4.86
23	**132b**	4’-methoxyphenylethanol	1	5.17
24	**133b**	4-(1-hydroxyethyl)benzonitrile	1	5.62
25	**67c**	acetophenone	1	3.79
26	**129c**	4’-chloroacetophenone	1	4.29
27	**123c**	benzylaldehyde	4	3.04
28	**123d**	benzoic acid	4	5.27

29	**134a**	thioanisole	2	1.68
30	**135a**	methyl-*o*-tolyl sulfide	2	1.73
31	**137a**	methyl-*m*-tolyl sulfide	2	2.03
32	**124a**	methyl-*p*-tolyl sulfide	2	2.03
33	**137a**	4-fluorothioanisole	2	1.63
34	**138a**	4-chlorothioanisole	2	2.42
35	**139a**	4-methoxythioanisole	2	2.87
36	**140a**	4-(methythio)benzonitrile	2	3.55
37	**134b**	methylphenyl sulfoxide	2	2.90
38	**135b**	methyl *o*-tolyl sulfoxide	2	3.56
39	**136b**	methyl *m*-tolyl sulfoxide	2	3.94
40	**124b**	methyl *p*-tolyl sulfoxide	2	4.18
41	**137b**	4-fluorophenyl methyl sulfoxide	2	2.72
42	**138b**	4-chlorophenyl methyl sulfoxide	2	5.04
43	**139b**	4-methoxyphenyl methyl sulfoxide	2	7.38
44	**140b**	4-(methylsulfinyl)benzonitrile	2	6.95
45	**134c**	methyl phenyl sulfone	2	3.21
46	**135c**	1-methyl-2-(methylsulfonyl)benzene	2	3.96
47	**136c**	1-methyl-3-(methylsulfonyl)benzene	2	4.62
46	**124c**	4-(methylsulfonyl)toluene	2	4.81
49	**137c**	4-fluorophenyl methyl sulfone	2	2.85
50	**138c**	4-chlorophenyl methyl sulfone	2	5.41
51	**139c**	4-methoxyphenyl methyl sulfone	2	8.18
52	**140c**	4-(methylsulfonyl)benzonitrile	2	7.63
53	**141a**	styrene	1	2.20
54	**88a**	*cis*-β-methylstyrene	1	2.15
55	**141b**	styrene oxide	1	3.41
56	**88b**	(1S,2S)-(−)-1-phenylpropylene oxide	1	3.14

n.d. = not determined

Table 10.3 Chiral GC analysis of oxidation products (alcohols, sulfoxides and epoxides) and methods used

Entry no.	Compound no.	Compound	Method	Retention time (min)	
				R	S
1	**67b**	phenyl ethanol	5	6.65	7.07
2	**125b**	1-(2-methylphenyl)ethanol	5	12.00	13.24
3	**126b**	1-(3-methylphenyl)ethanol	5	10.40	10.76
4	**127b**	1-(4-methylphenyl)ethanol	5	9.94	10.40
5	**128b**	4'-fluorophenylethanol	5	7.35	7.71
6	**129b**	4'-chlorophenylethanol	5	22.18	23.75
7	**130b**	4'-methoxyphenylethanol	5	28.36	27.40
8	**131b**	2'-methoxyphenylethanol	5	n.d.	n.d.
9	**132b**	3'-methoxyphenylethanol	5	n.d.	n.d.
10	**133b**	4-(1-hydroxyethyl)benzonitrile	5	n.d.	n.d.
11	**134b**	methylphenyl sulfoxide	6	4.74	5.27
12	**135b**	methyl *o*-tolyl sulfoxide	6	n.d.	n.d.
13	**136b**	methyl *m*-tolyl sulfoxide	6	5.33	5.86
14	**124b**	methyl *p*-tolyl sulfoxide	6	5.37	5.65
15	**137b**	4-fluorophenyl methyl sulfoxide	6	4.29	4.73
16	**138b**	4-chlorophenyl methyl sulfoxide	6	7.71	8.29
17	**139b**	4-methoxyphenyl methyl sulfoxide	6	10.54	11.32
18	**140b**	4-(methylsulfinyl)benzonitrile	6	n.d.	n.d.
19	**141b**	styrene oxide	7	11.00	11.25
20	**88b**	(1S,2S)-(−)-1-phenylpropylene oxide	7	n.d.	n.d.

n.d. = not determined

10.1.7. Synthesis

10.1.7.1. Sulfoxidations 135a – 138a

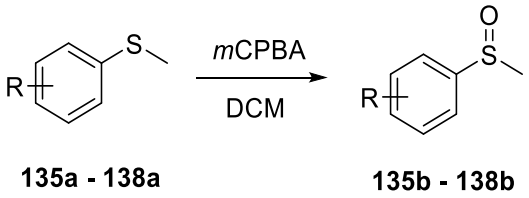


Figure 10.5 Sulfoxidation of thioethers **135a** – **138a** using *m*CPBA to their respective sulfoxides **135b** – **138b**

Compounds **135a** – **138a** (5 mmol) were added to a solution of *m*CPBA (6 mmol.) in dry DCM (15 mL) and the reactions were stirred overnight at r.t. The reactions were washed with saturated $NaHCO_3$ (30 mL × 2) and the organic layer was dried with $MgSO_4$. This was filtered *via* vacuum filtration and the filtrate was collected. The filtrate was then concentrated *in vacuo* and purified by flash column chromatography (SiO_2, ethyl acetate:hexane (7:3) with 2% acetic acid) to afford the target sulfoxides.

10.1.7.2. Ketone reduction to form 133b

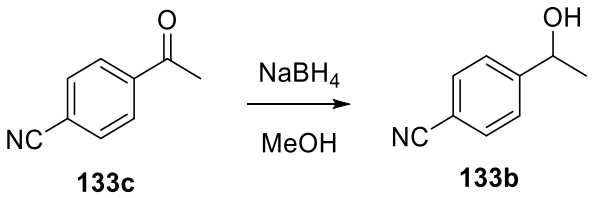
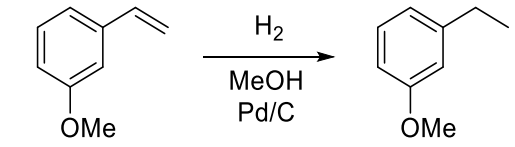


Figure 10.6 Reduction of **133c** to **133b**

4-acetylbenzonitrile (1.00 g, 6.89 mmol) was added to a solution of sodium borohydride (1.04 g, 27.5 mmol) in dry methanol (15 mL) and was left overnight at r.t.. The reaction was quenched with water, washed with 2 × 30 mL EtOAc and dried with $MgSO_4$. The mixture was filtered and the solvent was removed *in vacuo* to afford the target compound **133b**. NMR data was compared to literature. [191]

^{1}H NMR: (400 MHz, chloroform-*d*): δ_H 7.49 – 7.63 (m, 4H). 4.96 (br d, 1H). 2.16 (s, 1H), 1.49 (s, 3H)

10.1.7.3. Hydrogenation of 3-vinylanisole



3-vinylanisole **132a**

Figure 10.7 Hydrogenation of 3-vinylanisole to **132a**

To a solution of 3-vinylanisole (1.00 g, 7.45 mmol) in methanol (37 mL), Pd/C (745 mg, 10% loading) was added, and then the mixture was purged with argon. The reaction vessel was then evacuated and fitted with a H_2 balloon and stirred at r.t. for 24 h. The hydrogen was then evacuated and the reaction mixture filtered, rinsing with ethyl acetate, and concentrated *in vacuo* to afford the target compound **132a**. NMR data was compared to literature. [192]

^{1}H NMR: (400 MHz, chloroform-*d*): δ_H 6.75 (m, 4 H). 3.82 (s, 3 H). 2.65 (q, *J* = 7.63 Hz, 2 H). 1.25 (t, *J* = 7.63, 3 H).

10.1.7.4. Methylation of *o*-toluenethiol

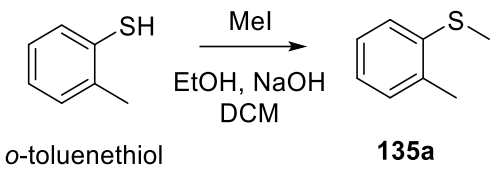


Figure 10.8 Methylation of *o*-toluenethiol to compound **135a**

O-toluenethiol (1.00 g, 8.1 mmol) was added to methyl iodide (1.70 mL, 12.1 mmol) and sodium hydroxide (4.84 mg, 12.1 mmol) in 9.7 mL of ethanol. The reaction was stirred overnight and was extracted with 2 × 10 mL DCM. The resulting sulfide, **135a**, was then used for biotransformations or oxidised following the sulfoxidation protocol in Section 10.1.7.1. NMR data was compared to literature. [193]

^{1}H NMR: (400 MHz, chloroform-*d*): δ_H 7.08 – 7.17 (m, 4 H), 2.48 (s, 3 H). 2.35 (s, 3 H).

10.2.Results and Discussion

10.2.1. Initial biotransformations

To investigate the oxygenation ability of artUPO$_{yeast}$, a small initial substrate screen was carried out using compounds **67a**, **122a**, **123a** and **124a** in order to obtain a vague idea of the reaction capabilities of artUPO$_{yeast}$. artUPO$_{yeast}$ supernatant from the *P. Pastoris* fermentation culture was used which had a specific activity of 1.5 U mL^{-1} NBD. The reactions were stirred at 300 rpm at r.t. and were monitored for 24 h using GC. The conversion was determined based on relative peak areas of the starting material and the products. The results from this initial screen are shown in Figures 10.9, 10.10 and 10.11. artUPO$_{yeast}$ was able to catalyse the oxygenation reactions of **67a**, **123a** and **124a** but not **122a**. The fact that artUPO$_{yeast}$ did not convert any **122a** was surprising because *Mro*UPO, the enzyme artUPO is derived from, was able to catalyse the oxidation to both the alcohol and ketone product successfully in another study carried out by Peter *et al*. 2014. [194] This result may suggest that the mutations in the substrate tunnel of artUPO$_{yeast}$ compared to *Mro*UPO may have affected the ability of the enzyme to accept **122a** as a substrate. In Figure 10.9A, the alcohol **67b** is formed as the major product and the respective ketone **67c** is formed in lesser quantities. This is also seen in Figure 10.9B where the major product is the first oxidation product, **124b**. artUPO$_{yeast}$ was also able to catalyse the formation of the sulfone **124c** which differs from sulfoxidations carried out by *Aae*UPO because *Aae*UPO has not been shown to form sulfones from sulfides. [138] In Figure 10.9B the major product is less clear as compounds **123b** and **123c** are formed in similar amounts. This may suggest that artUPO$_{yeast}$ has a similar affinity towards **123a** and **123b** in these conditions.

Interestingly, no traces of benzoic acid were observed which implies that artUPO$_{yeast}$ cannot oxidise the aldehyde further.

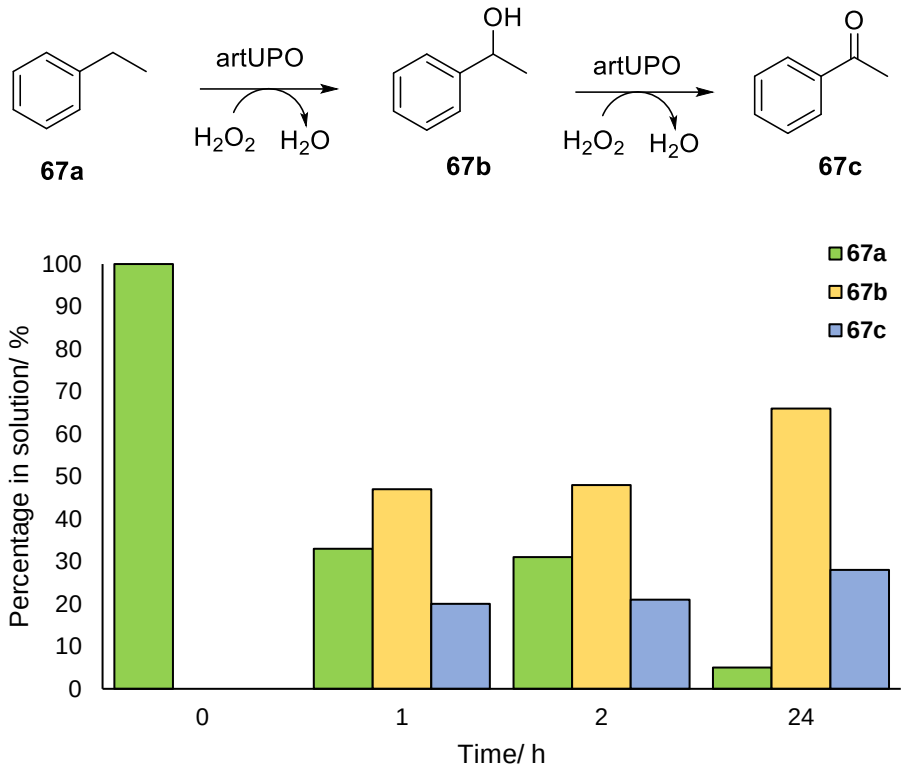


Figure 10.9 2 mL biotransformation of ethylbenzene **67a** with artUPO$_{yeast}$ analysed by GC with **67a** in green, phenylethanol **67b** in yellow and acetophenone **67c** in blue. Reaction conditions were as follows: 50 mM potassium phosphate buffer pH 7.0, 1.5 U mL^{-1} artUPO$_{yeast}$, 10 mM substrate with either 10% ACN or 1% EtOH final concentration, 10 mM H_2O_2 at r.t., 300 rpm

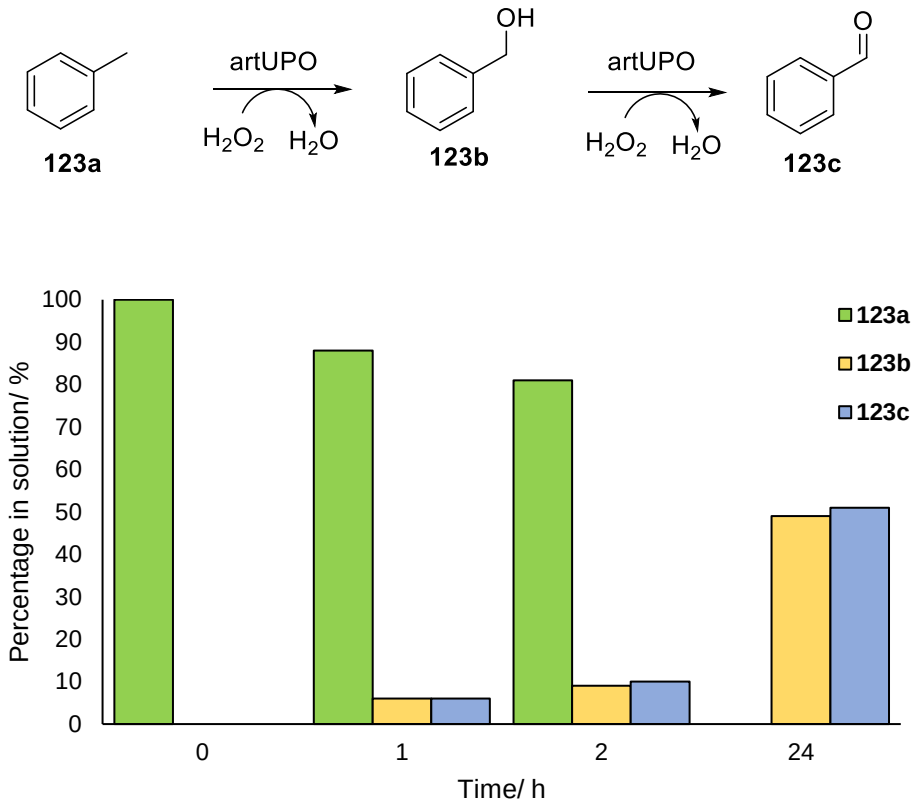


Figure 10.10 2 mL biotransformation of toluene **123a** with artUPO$_{yeast}$ analysed by GC with **123a** in green, benzyl alcohol **123b** in yellow and benzaldehyde **123c** in blue. Reaction conditions were as follows: 50 mM potassium phosphate buffer pH 7.0, 1.5 U mL^{-1} artUPO$_{yeast}$, 10 mM substrate with either 10% ACN or 1% EtOH final concentration, 10 mM H_2O_2 at r.t., 300 rpm

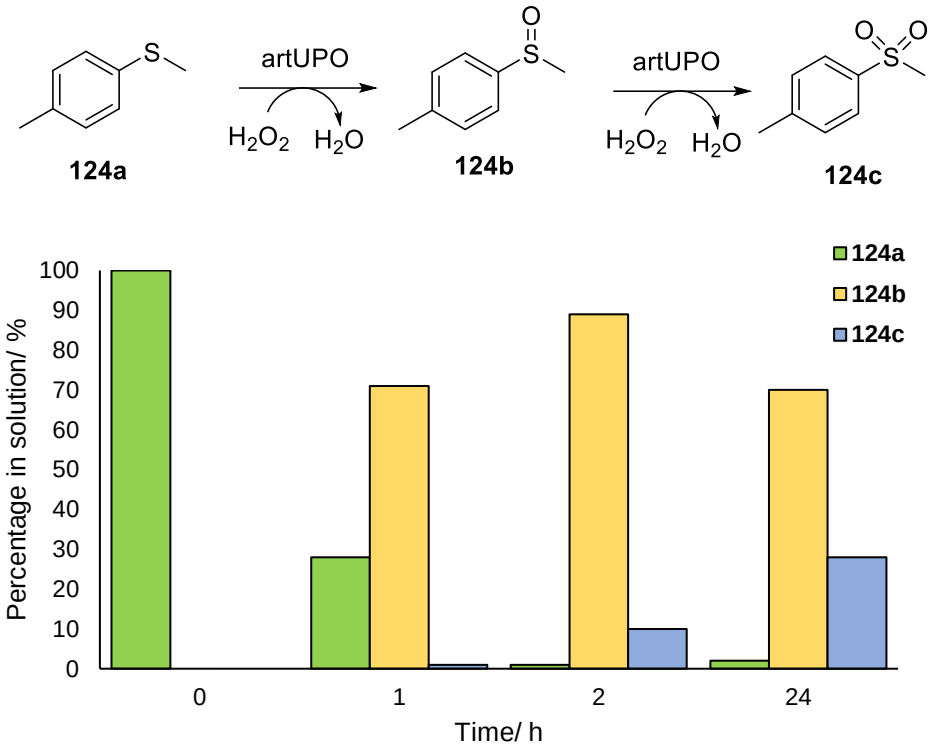


Figure 10.11 2 mL biotransformation of methyl *p*-tolyl sulide **124a** with artUPO$_{yeast}$ analysed by GC with **124a** in green, methyl *p*-tolyl sulfoxide **124b** in yellow and methyl *p*-tolyl sulfone **124c** in blue. Reaction conditions were as follows: 50 mM potassium phosphate buffer pH 7.0, 1.5 U mL^{-1} artUPO$_{yeast}$, 10 mM substrate with either 10% ACN or 1% EtOH final concentration, 10 mM H_2O_2 at r.t., 300 rpm

In the hydroxylation reaction of **67a** and the sulfoxidation reaction of **124a**, the reactions appear to be complete after 1 h or 2 h and so the next set of experiments were then carried out for only 2 h.

Before further screening of reactions was carried out, the reaction conditions were first optimised. The initial screen was caried out at r.t. which was ~20°C but this temperature fluctuated. To attempt to control “room temperature”, experiments were carried out at 20 °C and 25 °C. The method of mixing the reactions was also investigated because the stirrer bars used in the reactions did not appear to be mixing the reaction evenly. The effects of these two parameters can be seen in Figure 10.12

where two biotransformations, hydroxylation of ethylbenzene **67a** and sulfoxidation of methyl *p*-tolyl sulfide **124a**, were carried out for 1 h.

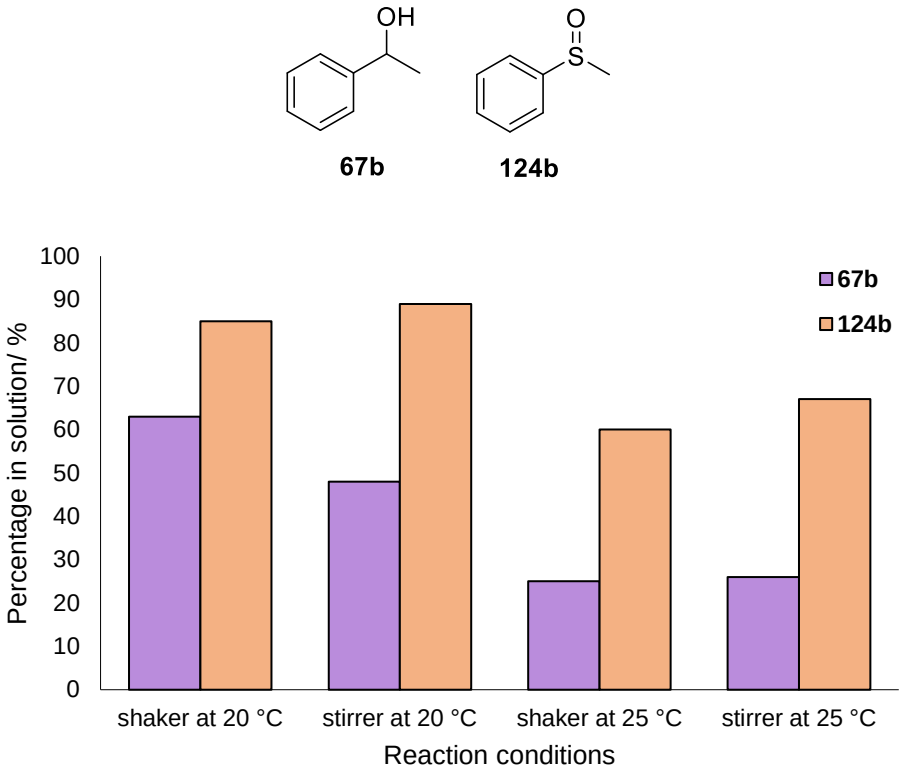


Figure 10.12 Investigation of the effects of temperature and method of mixing on the formation of phenylethanol **67b** (in purple) and methyl *p*-tolyl sulfoixide **124b** (in orange). Reaction conditions were as follows: 50 mM potassium phosphate buffer pH 7.0, 1.5 U mL^{-1} $artUPO_{yeast}$, 10 mM substrate with either 10% ACN or 1% EtOH final concentration, 2 mM H_2O_2, temperature and mixing given in the figure.

For both hydroxylation (purple) and sulfoxidation (orange) the conversion did not seem to be affected by the different methods of mixing but was affected by the difference in temperature as both reactions had higher rates at 20°C. This may be due to higher enzyme stability at 20°C compared to 25°C. This may raise questions about how other studies have carried out UPO biotransformations when specific temperatures are not

stated. From these experiments, it was decided that further reactions would be carried out at 20 °C and mixed using a benchtop shaker at 700 rpm.

10.2.2. Preliminary biotransformation screen on 2 mL scale

To compare the catalytic ability of artUPO$_{yeast}$ with r*Aae*UPO (made previously by Mielke *et al*.), preliminary reactions were performed using a screen of compounds based on ethylbenzne **67a**, methyl *p*-tolyl sulfide **124a** and styrene **141a**.[95] r*Aae*UPO shares the same sequence as PaDa-I which has nine point mutations, four in the signalling peptide and five in the main sequence as described by Molina-Espeja *et al*. 2014.[96] The hydroxylation and sulfoxidation screens were designed to investigate the effect of substituents on the phenyl ring of both **67a** and **124a** where the position of the substituents and the electronic character was varied. Controls were also run where no UPO was added to the reactions and was monitored for 2 h (see appendix). In most of the controls for the hydroxylations, sulfoxidations and epoxidations, only traces of the oxidation products were observed. However, some background oxidation is observed in the sulfoxidation screen when the compound is monitored for 5 h with no UPO. An example of this is shown in Figure 10.13 with compound **124a** in the reaction. The percentage of **124a** starts to fall after 2 h and the percentage of the respective sulfoxide **124b** increases up to 21% after 5 h. The sulfone **124c** was not detected. The background oxidation is likely to be caused by the hydrogen peroxide in solution. Although in the enzymatic reactions much of the hydrogen peroxide is consumed by the UPO, some residual hydrogen peroxide will be free to oxidise the sulfides. It is therefore important to take this into consideration when analysing sulfoxidation

conversions. Despite this, comparisons of sulfoxidations between UPOs under the same conditions should still be valid.

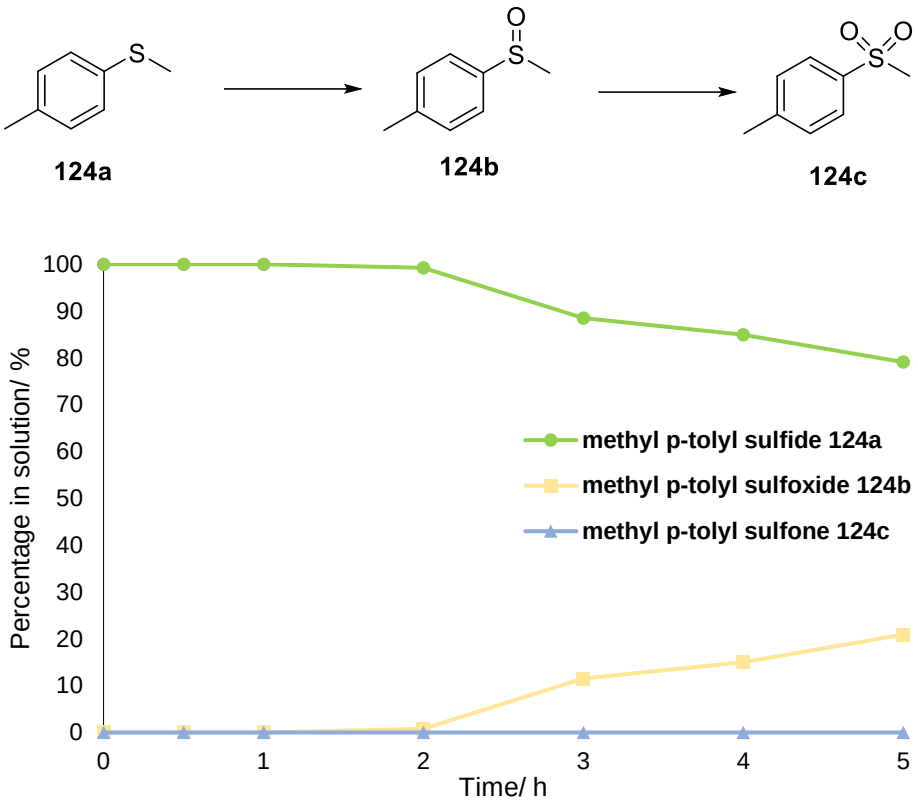


Figure 10.13 Control reaction of 10 mM methyl *p*-tolyl sulfide **124a**, 2 mM H_2O_2, with no UPO added to the reaction at 700 rpm, 20 °C

The results of the preliminary biotransformation screen are given below in Figures 10.14, 10.15 and 10.16. Transformations carried out by artUPO$_{yeast}$ are given in blue and transformations carried out by r*Aae*UPO are given in red. Given that the reactions for ethylbenzene **67a** and methyl *p*-tolyl sulfide **124a** appeared to be complete in Figures 10.9 and 10.11, the biotransformations were carried out for 2 h. To prevent deactivation of the heme, only 2 mM of hydrogen peroxide was added to the reaction.

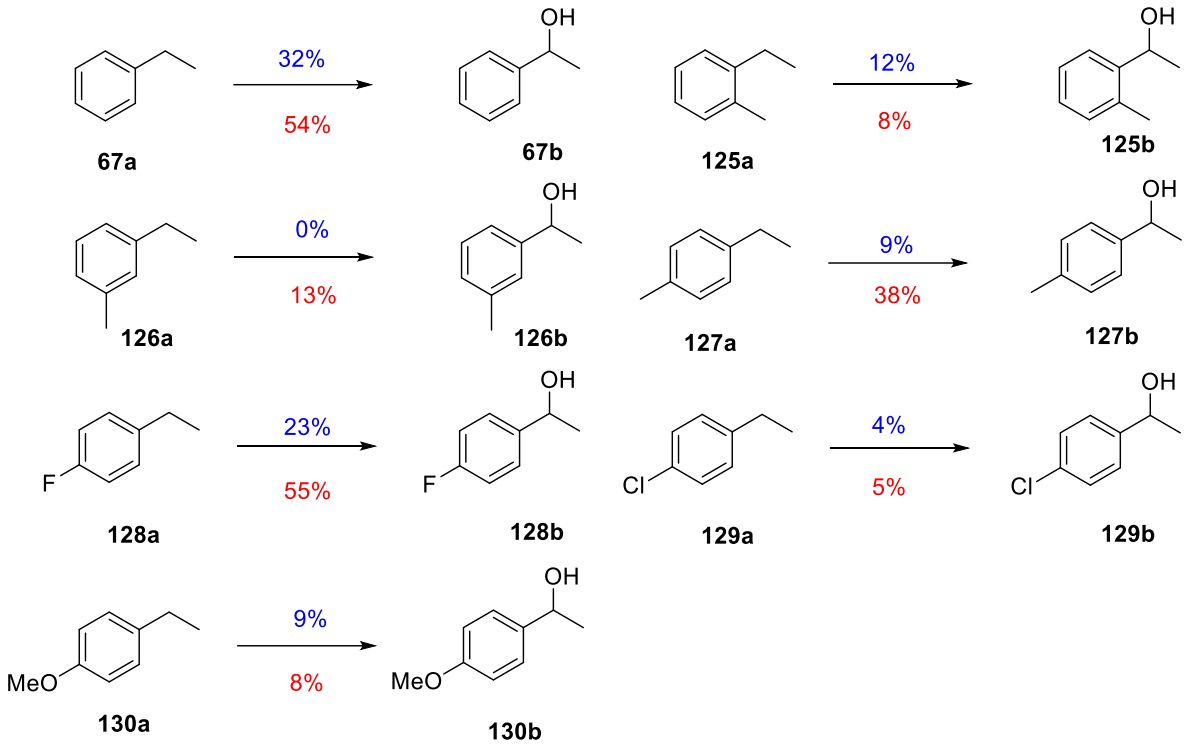


Figure 10.14 Hydroxylation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 2 mL scale with conversions as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 10 mM substrate with 10% ACN final concentration, 2 mM H_2O_2, 700 rpm, 20 °C, 2 h

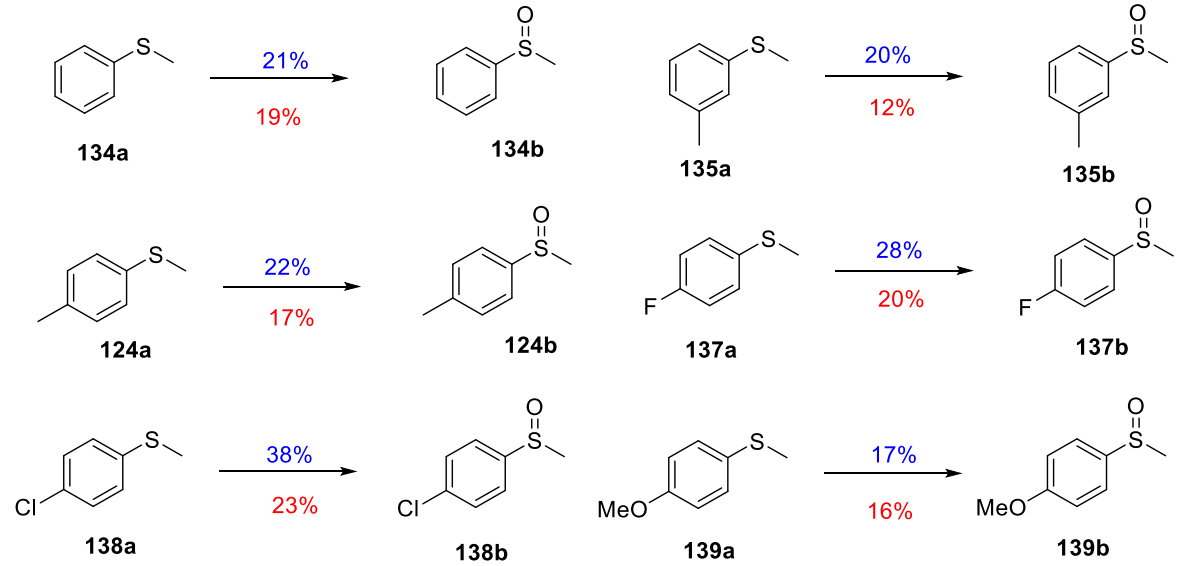


Figure 10.15 Sulfoxidation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 2 mL scale with conversions as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 10 mM substrate with 1% EtOH final concentration, 2 mM H_2O_2, 700 rpm, 20 °C, 2 h

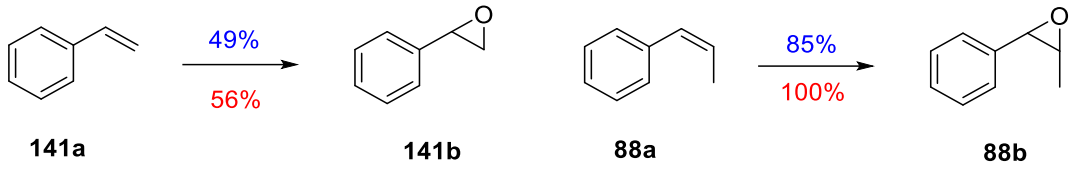


Figure 10.16 Epoxidation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 2 mL scale with conversions as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 1 mM substrate with 10% ACN final concentration, 1 mM H_2O_2, 700 rpm, 20 °C, 2 h

Overall in all of the screens, artUPO$_{yeast}$ and r*Aae*UPO were able to regioselectively oxidise the benzylic carbon or sulfur. In the hydroxylation, r*Aae*UPO was generally able to generate higher yields of the alcohol than artUPO$_{yeast}$. A reason for this difference in conversion could potentially be due to the difference in the structures of the substrate channel in each UPO. The substrate channel of *Aae*UPO has five key phenylalanine residues, making it highly aromatic compared to the substrate tunnel of artUPO which has a more aliphatic character (4 Ile, 4 Leu, 2 Val, 2 Ala, 2 Thr, Lys, Ser and Phe). [89, 91] The aromatic nature of the substrates in the screens are similar to the phenylalanine residues found in the channel of *Aae*UPO and so the enzyme may be able to accept these substrates more readily than artUPO. In the hydroxylation screen, the highest conversions by both UPOs are seen with the substrates ethylbenzene **67a** and 1-ethyl-4-fluorobenzene **128a**. Interestingly, artUPO$_{yeast}$ was unable to accept 3-ethyltoluene **126a** as a substrate but r*Aae*UPO was able to oxidise this compound. The reason for this was unclear and so further investigation was carried out and discussed in Section 10.3.1. r*Aae*UPO was also able to catalyse the epoxidation of both styrene derivatives more efficiently than artUPO$_{yeast}$ which may be for the same reasons as described for the benzylic hydroxylations.

In contrast to the previous screens, In the sulfoxidation screen artUPO$_{yeast}$ shows slightly higher conversions than r*Aae*UPO. In this case, the difference may be reasoned again due to the nature of the residues in the substrate channels of each UPO. Although the artUPO substrate channel is mostly aliphatic and non-polar, the residues are slightly less hydrophobic than the phenylalanine residues found in the substrate channel in *Aae*UPO which may allow for the more hydrophilic sulfide substrates to be more readily accepted by artUPO compared to r*Aae*UPO. The difference in conversion in the sulfoxidation screen between the two UPOs however is relatively minor except for 4-chlorothioanisole **138a**, where artUPO$_{yeast}$ is able to convert more of the sulfide than r*Aae*UPO (38% by artUPO$_{yeast}$ compared to 23% r*Aae*UPO) .

Not enough material could be collected from the 2 mL biotransformations for chiral GC analysis as the GC could not detect the products. Because of this, the reactions in the next section were scaled up to 5 mL. The conversions of **67a** and **124a** to the primary oxidation product by artUPO$_{yeast}$ in these reactions (Figures 10.14 and 10.15) were lower than expected when compared to the results in Figure 10.12. However, less UPO was added to the reactions, from 1.5 U mL^{-1} to 0.5 U mL^{-1}, to preserve the stock of artUPO$_{yeast}$ for future use. For the next set of biotranformations, 10 mM H_2O_2 was added to the reaction and the transformations were carried out for 6 h.

10.2.3. Biotransformation screens and time courses on 5 mL scale

10.2.3.1. Conversions of akylbenzenes, thioethers and styrene derivatives by artUPO$_{yeast}$ and r*Aae*UPO

Biotransformations including a hydroxylation screen, a sulfoxidation screen and an epoxidation screen were carried out with artUPO$_{yeast}$ and r*Aae*UPO (previously made by Mielke *et al.*) on a 5 mL scale and samples were taken after 6 h.[95] Time courses were also carried out for the hydroxylation and sulfoxidation screen on this scale and were monitored for 6 h. Due to limiting resources, the reactions were shaken at 250 rpm in an orbital shaker at 20 °C. The hydroxylation screen was expanded so that more substituents on the phenyl ring of **67a** were included, such as 2-OMe **131a**, 3-OMe **132a** and 4-CN **133a**. The compounds **131a** and **132a** were added to the screen to investigate potential reasons why artUPO$_{yeast}$ was unable to accept the *meta*-substituted methyl derivative **126a** of ethylbenzene. The *para*-nitrile derivative **133a** was added to the screen to investigate further the effects of electron withdrawing groups on the hydroxylation of the benzylic carbon by artUPO$_{yeast}$. In the sulfoxidation screen, a 4-CN derivative **140a** was added to also investigate electronic effects on the oxygenation of the benzylic sulfur. The *ortho*-methyl derivative **135a** of thioanisole **134a** was added to the screen and was synthesised as described in Section 10.1.7.4.

The results of the hydroxylation, sulfoxidation and epoxidation screens are given in Figures 10.17, 10.18 and 10.19 respectively. The reactions with artUPO$_{yeast}$ in the hydroxylation and sulfoxidations screens were run in duplicates and the error bars show the standard deviation. The enantiomeric excess (*ee*) values of the biotransformations are also given. To determine absolute configuration of the oxidised

compounds, literature of *Aae*UPO biotranformations with either the same or very similar substrates have stated that these reactions occur with a preference to form the *R* enantiomer of both alcohols and sulfoxides. [95,127,144,187,188] Other analysis of similar compounds in the literature by chiral GC was also examined to assign the configuration of each enantiomer in the screens. [189,190]

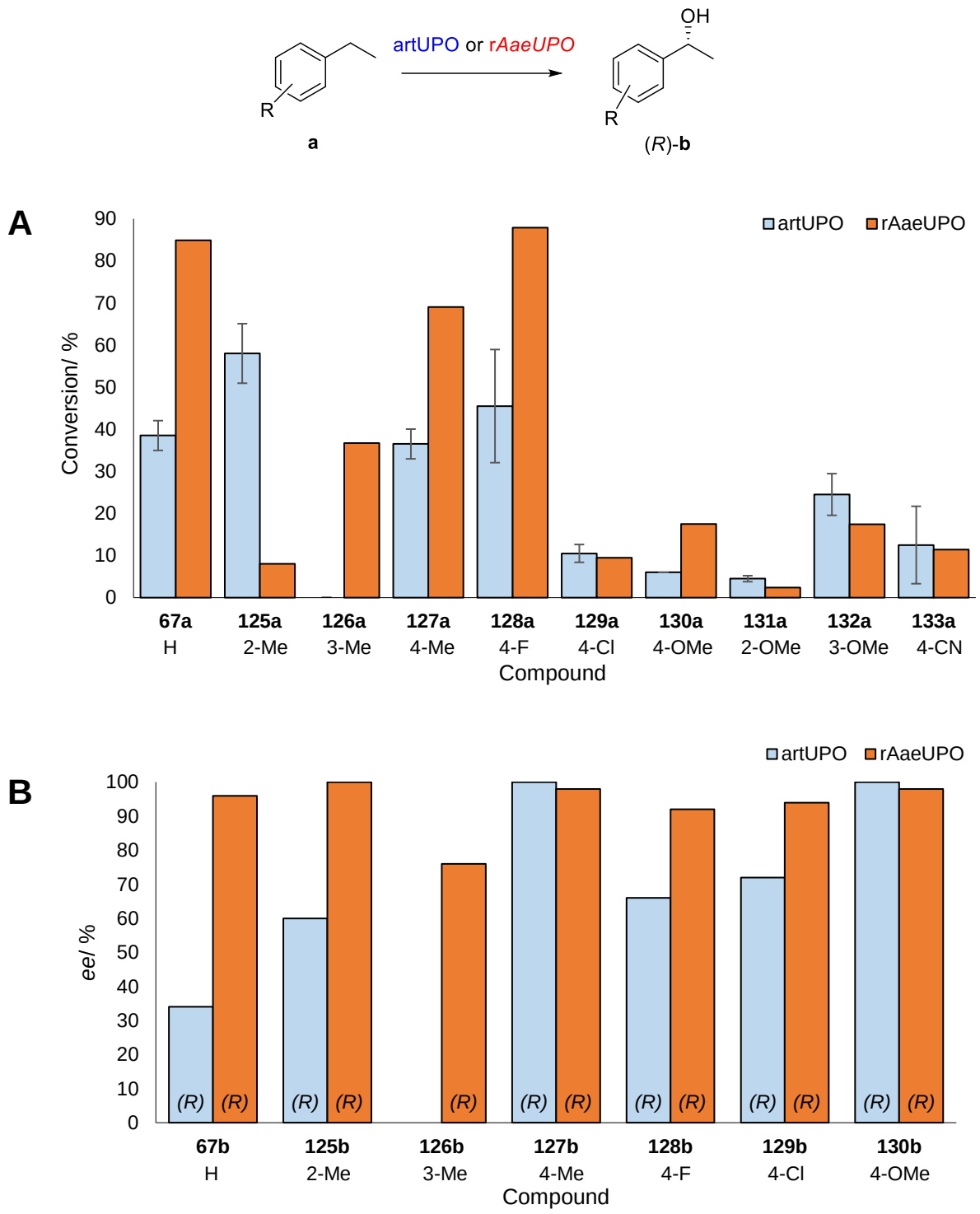


Figure 10.17 Hydroxylation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 5 mL scale with **A** conversions to the alcohol and **B** *ees* as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 10 mM substrate with 10% ACN final concentration, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

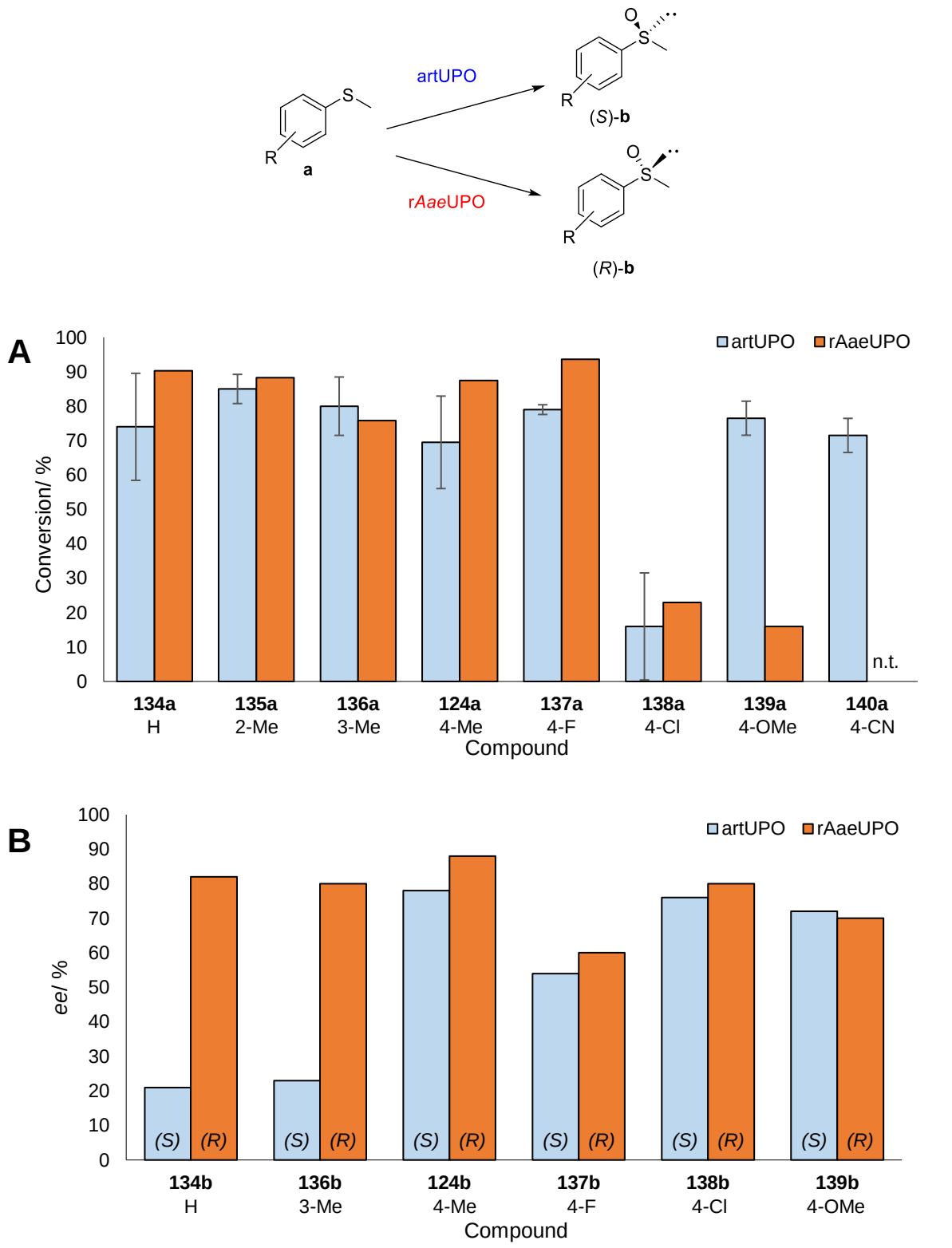


Figure 10.18 Sulfoxidation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 5 mL scale with **A** conversions to the sulfoxide and **B** *ees* as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 10 mM substrate with either 10% ACN or 1% EtOH final concentration, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h. n.t. = not tested

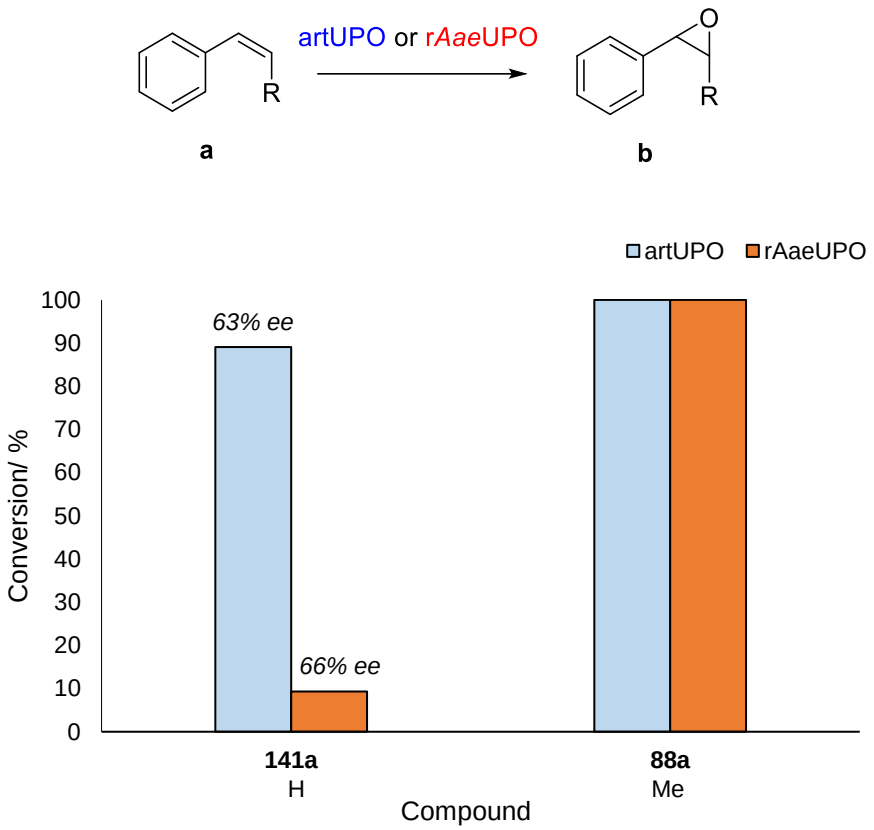


Figure 10.19 Epoxidation screen with artUPO$_{yeast}$ (blue) and r*Aae*UPO (red) on 5 mL scale with conversions to the epoxide and *ees* of reaction with **141a** as percentages (determined by GC). Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} UPO, 10 mM substrate with 10% ACN final concentration, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

In the hydroxylation screen (Figure 10.17), moderate to low conversions to the alcohols are observed when artUPO$_{yeast}$ is used as a catalyst. r*Aae*UPO shows much higher conversions than artUPO$_{yeast}$ for ethylbenzene **67a**, 3-ethyltoluene **126a**, 4-ethyltoluene **127a** and 1-ethyl-4-fluorobenzene **128a**. However, artUPO$_{yeast}$ appears to convert **125a** more readily than r*Aae*UPO. artUPO$_{yeast}$ was able to transform the *meta*-methoxy derivative **132a** which was surprising given artUPO$_{yeast}$ was not reactive to **126a** (3-ethyltoluene). The ability of artUPO$_{yeast}$ to transform **132a** but not **126a** seems to eliminate the possibility of a steric issue in the enzyme. This could potentially be due to electrostatic interactions instead.

In the sulfoxidation screen (Figure 10.18), generally good conversions of ≥ 70% to the sulfoxide are carried out by artUPO$_{yeast}$. The exception to this is seen in sulfoxidation of the *para*-chloro derivative **138a**. The *para*-chloro derivative of ethylbenzene, **129a**, is also poorly converted by artUPO$_{yeast}$ which may suggest that the *para*-chloro-substituent is unfavourable in the active site of artUPO. High conversions are also observed for the epoxidation of styrene **141a** and *cis*-β-methylstyrene **88a** carried out by artUPO$_{yeast}$ (Figure 10.19), indicating that artUPO$_{yeast}$ may have an affinity towards unsaturated bonds. Surprisingly, the conversion of styrene **141a** to styrene oxide **141b** is very low for r*Aae*UPO, especially compared to earlier conversions seen in Figure 10.16.

In the hydroxylation screen both UPOs generate the (*R*)-alcohol, where r*Aae*UPO shows consistently high enantioselectivity (Figure 10.17) but the selectivity with artUPO$_{yeast}$ is more variable, with particularly high *ee*s for compounds **127b**, **130b** and **131b**. However in the sulfoxidation screen, the opposite enantioselectivity is observed with artUPO$_{yeast}$ compared to r*Aae*UPO. An example of this is seen in Figure 10.20 with compound **124b**, where r*Aae*UPO enantioselectively catalyses the formation of (*R*)-**124b** whereas artUPO$_{yeast}$ catalyses the formation of (*S*)-**124b**. This complementary enantioselectivity, with r*Aae*UPO preferably forming the (*R*)-sulfoxide and artUPO$_{yeast}$ forming the (*S*)-sulfoxide is observed across the whole screen (see Appendix for all chromatograms) .

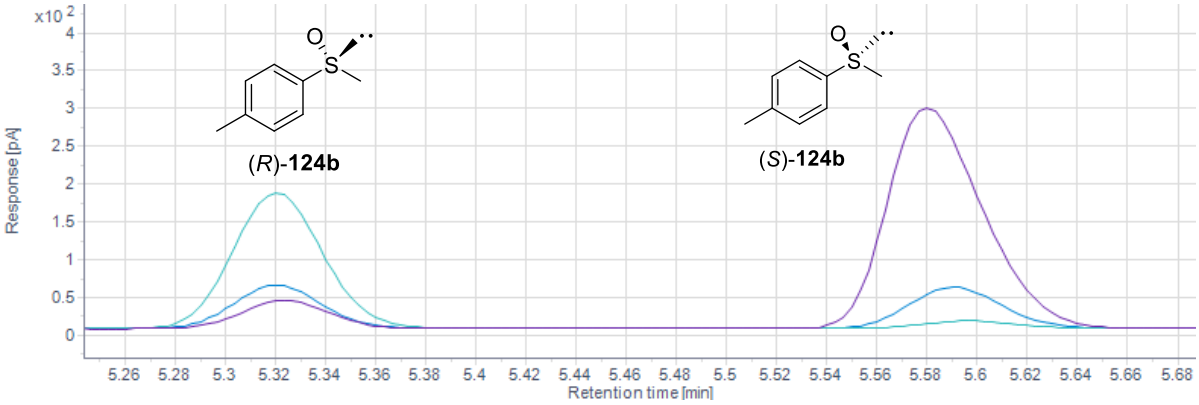


Figure 10.20 Chiral GC trace of methyl *p*-tolyl sulide **124b** on BGB-175 column. **Dark blue** represents rac-**124b**, **teal** is the reaction catalysed by r*Aae*UPO and **purple** is the reaction catalysed by artUPO$_{yeast}$

The reason for the differences in enantioselectivity could be due to the differences in how the residues in the heme channel of r*Aae*UPO and artUPO$_{yeast}$ orientate the sulfide in their respective active site. However, attempted modelling experiments using the UPO structures and AutoDock Vina did not yield any reasonable orientations of the substrate. [195, 196] Other experiments could be carried out to try and elucidate the structure of the enzyme with bound substrate such as crystallography or NMR experiments.

10.2.3.2. Time courses of biotransformation screens by artUPO$_{yeast}$

For the hydroxylation, sulfoxidation and epoxidation screens, a time course was taken for the duration of 6 h. Samples were taken at 0.0 h, 0.5 h, 1.0 h, 2.0 h, 4.0 h and 6.0 h and these reactions were carried out with both 10 mM and 2 mM substrate loading. Hydroxylation activity was again seen with all substrates other than **126b** (Figure 10.21).

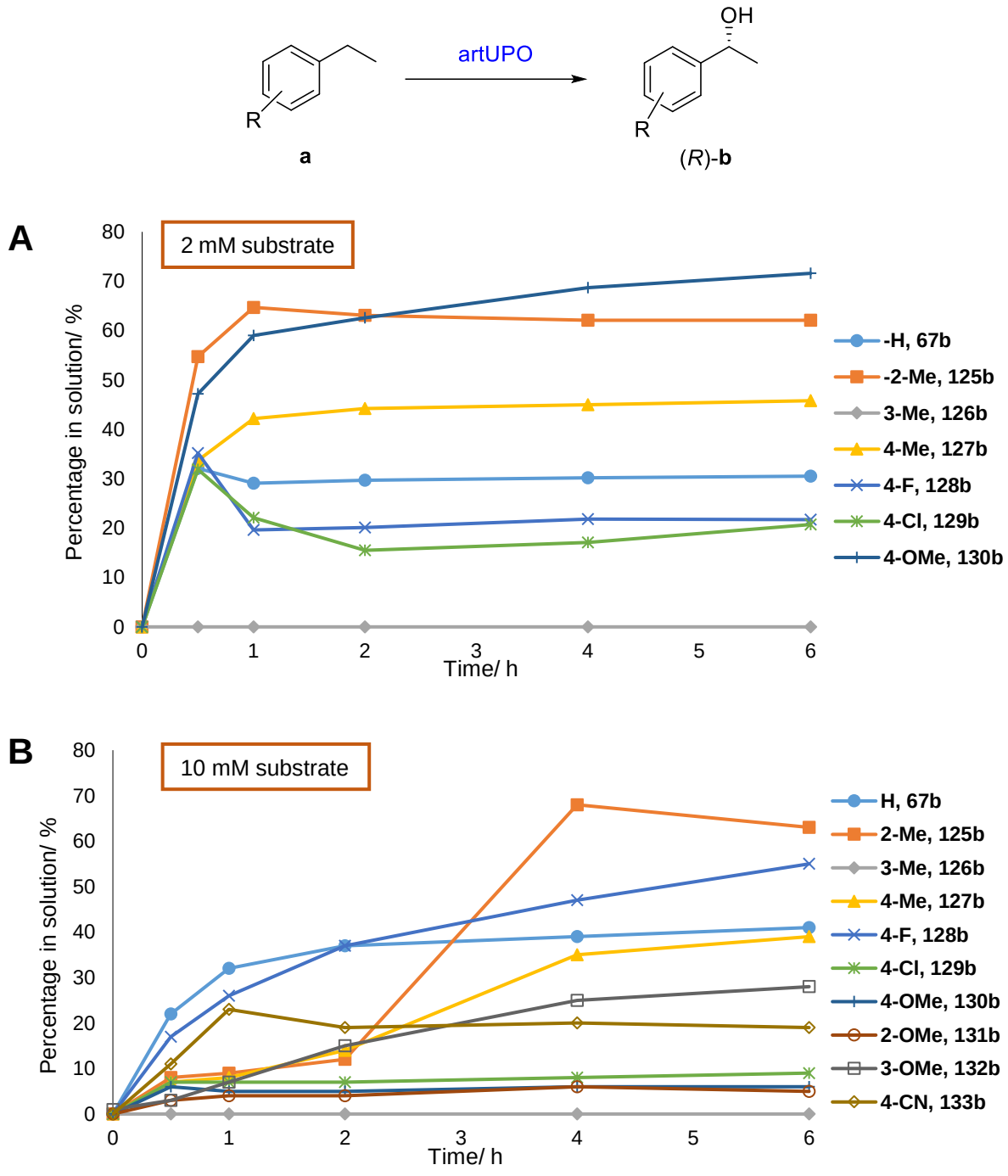


Figure 10.21 Time courses of the hydroxylation screen with **A** 2 mM loading of substrate and **B** 10 mM substrate with 10% ACN final concentration. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} $artUPO_{yeast}$, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

For both substrate loadings, a general trend is observed where percentage of the alcohol formed in the reaction increases over time with the exception of compounds

128b (4-F) and **129b** (4-Cl) when 2 mM substrate loading is used. In these cases, a maximum in product percentage is reached within an hour and rapidly falls again. The fall in the amount of alcohols **128b** and **129b** present is likely due to the oxidation of these alcohols to form the ketone derivatives **128c** and **129c**. Figure 10.22 shows the initial increase and then decrease of alcohols **128b** and **129b** whereas the respective ketones continue to increase as more time passes.

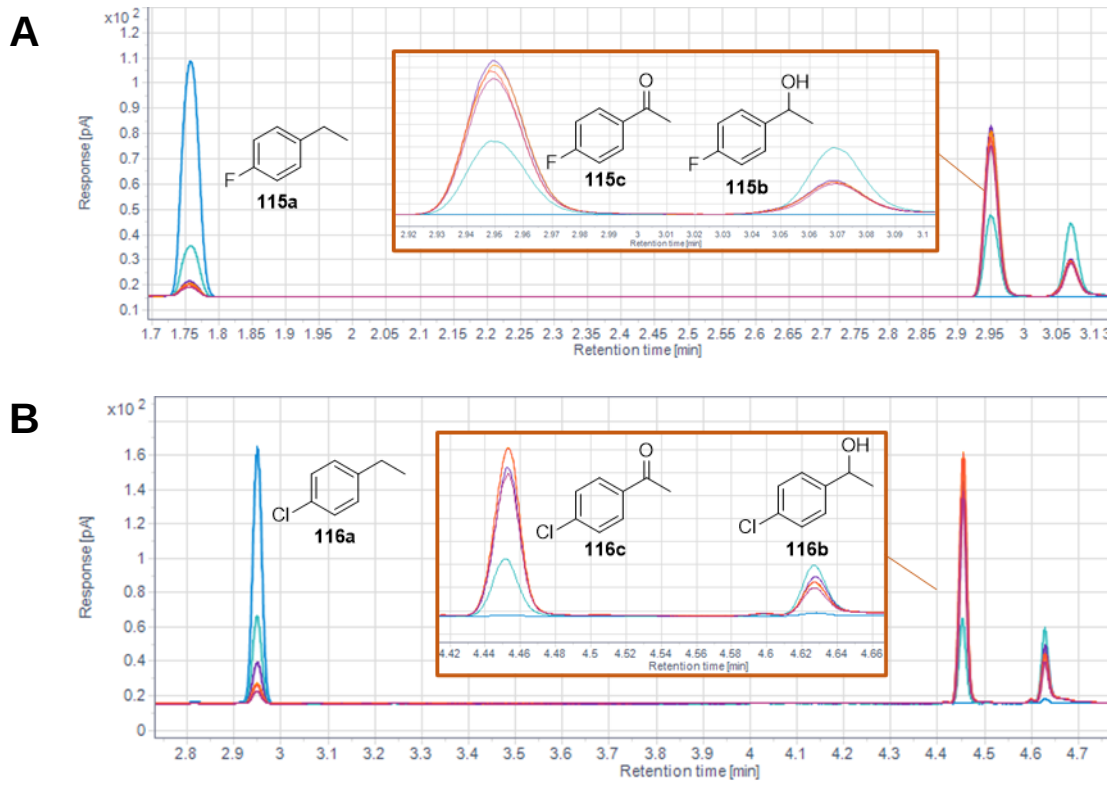


Figure 10.22 10.23 GC trace biotransformation of **A**: **128a** and **B**: **129a** with artUPO$_{yeast}$ over 6 h with 2 mM substrate loading. **Blue**: 0 h, **teal**: 0.5 h, **indigo**: 1.0 h, **orange**: 2.0 h, **red**: 4.0 h, **maroon**: 6.0 h. Close ups of the product formation included

Despite these two exceptions, higher conversions of hydroxylation are observed when the substrate loading is 2 mM. Product maxima are also reached within roughly 2 h in these conditions.

For a substrate loading of 2 mM in the sulfoxidation time courses, all reactions reach a product peak and have rapidly fallen within 2 h, and the amount of sulfide has decreased to nearly 0% by the end of the experiment (Figure 10.24).

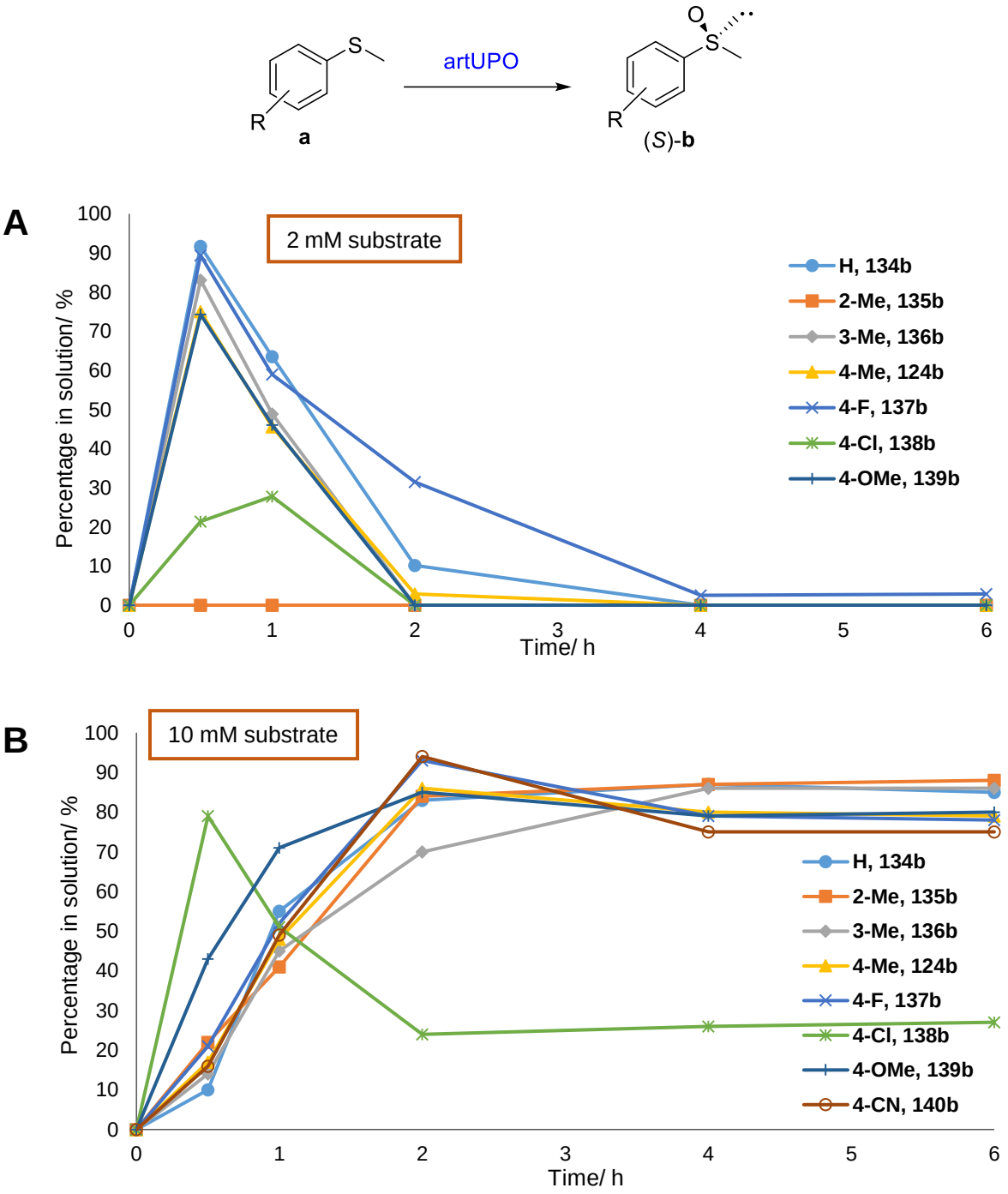


Figure 10.24 Time courses of the sulfoxidation screen with **A** 2 mM loading of substrate and **B** 10 mM substrate with either 10% ACN or 1% EtOH final concentration. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} $artUPO_{yeast}$, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

In contrast, with 10 mM substrate loading the formation of sulfoxide reaches a maximum at 2 h and falls slightly towards 6 h, except for **138b** (Figure 10.24). The reason for these differences may be the effect of the excess hydrogen peroxide in the system. The control reaction in Figure 10.13 has demonstrated that the hydrogen peroxide is able to oxidise the sulfide without the presence of a catalyst. It is possible that with a high percentage of sulfoxides in solution in a short period of time, an excess of hydrogen peroxide may drive the oxidation to form the sulfone. The time course of the formation of sulfones at 2 mM substrate loading is shown in Figure 10.25A sharp increase in the percentage of sulfones in solution is seen after 0.5 h, which corresponds to the sharp decrease in the percentage of the respective sulfoxides in Figure 10.24A. This affect does not seem to appear as dramatically with the oxidation of alcohols to ketones.

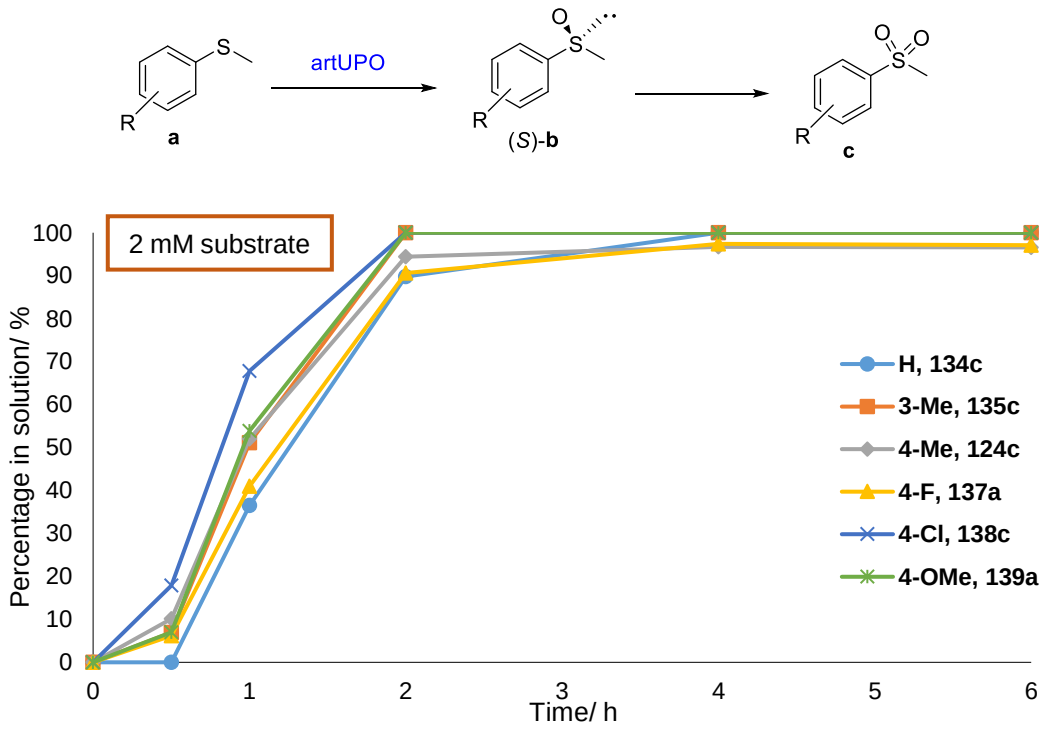


Figure 10.25 Time course of the formation of sulfone in the sulfoxidation screen. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 2 mM substrate with 1% EtOH final concentration, 0.5 U mL^{-1} $artUPO_{yeast}$, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

Although true initial rates have not been determined, meaning true Hammett-analysis cannot be carried out, a crude Hammett plot can be tentatively made for artUPO$_{yeast}$ using the first 30 min of the time courses to attempt to find a pattern between the effects of the substituents on the phenyl ring of the substrates and the rate of oxidation by artUPO$_{yeast}$. In the mechanism, a radical is formed *via* hydrogen atom abstraction. Stabilising the radical intermediate formed may aid the rate of reaction, hence the electronic effects of the substituents on the phenyl ring may play a role. The resulting plots are shown in Figure 10.26.

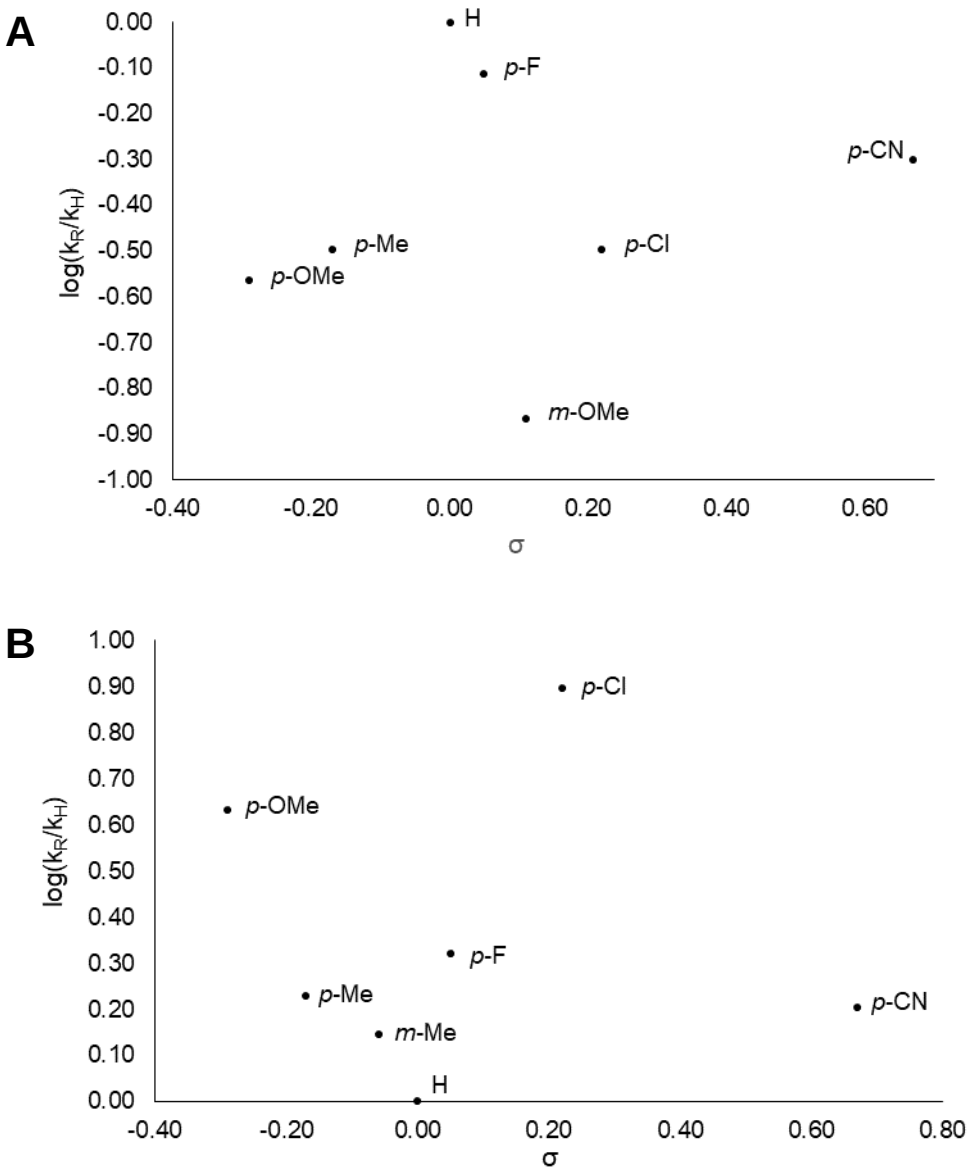


Figure 10.26 Hammett plots of **A** hydroxylation screen and **B** sulfoxidation screen. The observed rates were calculated from the amount of primary oxidation product in the reaction after 30 min determined by GC. Sigma values found in literature by Hansch et al. [197]

The Hammett plots in Figure 10.26 show that no clear trend can be made from the electronic effects of the substituents on the phenyl rings of both ethylbenzene derivatives and thioanisole derivatives from the data in the time courses, and therefore ρ parameters were not calculated. To improve the quality of these results, more dedicated experiments could be carried out with shorter time courses to determine the

initial rates with more repeats for more accurate data. Despite this, the plots are still useful in generating a vague idea of the lack of obvious Hammett-relationship of the substituent effects on the substrates and the rate of the reaction carried out by artUPO.

artUPO$_{yeast}$ is able to selectively form the epoxide with styrene **141a** and *cis*-β-methylstyrene **88a**, and the product maxima for both reactions is reached after 2 h (Figure 10.27).

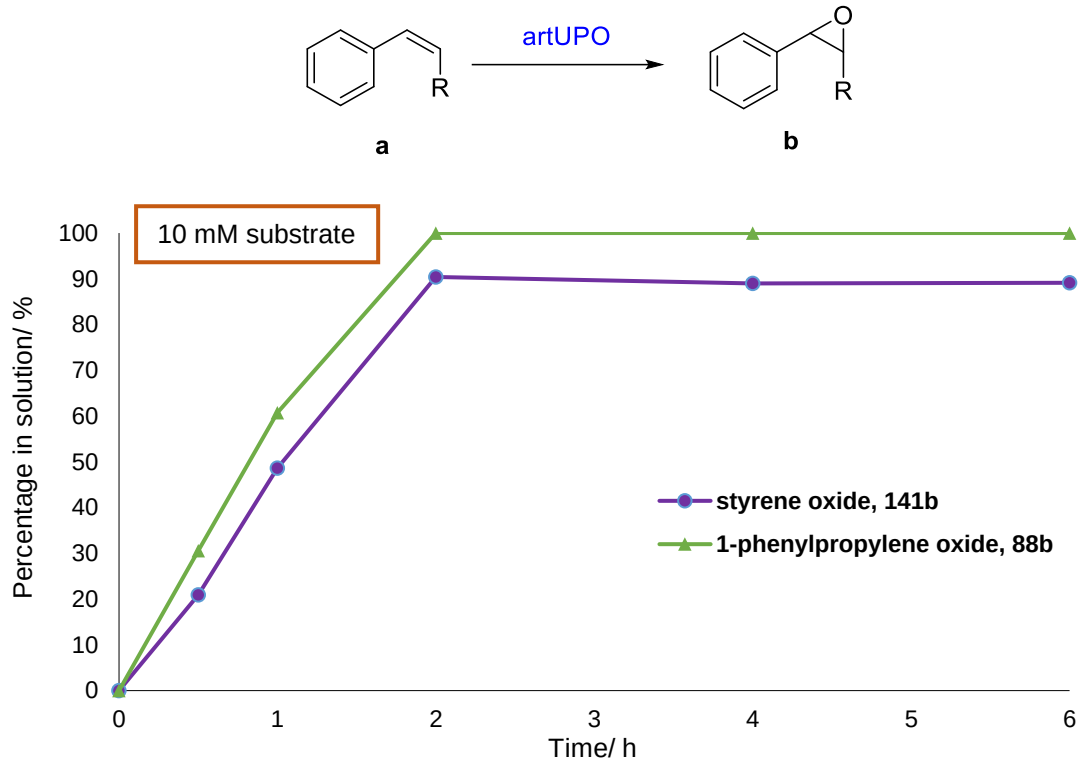


Figure 10.27 Time courses of the epoxidation screen with 10 mM substrate with 10% ACN final concentration. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.5 U mL^{-1} artUPO$_{yeast}$, 10 mM H_2O_2, 250 rpm, 20 °C, 6 h

The epoxidation of **88a** catalysed by artUPO$_{yeast}$ reaches 100% conversion unlike the reaction with **141a**. A potential reason for this could be due to a stabilising effect of the presence of a methyl group in **141a** which may improve the conversion by stabilising the intermediate state.

10.2.4. Biotransformations using artUPO$_{bact}$

To investigate the catalytic properties of artUPO$_{bact}$, alkylbenznes **67a** and **127a**, thioethers **134a** and **140a**, and styrene **141a** were selected for biotransformations and carried out on a 5 mL scale. The results are shown in Figure 10.28.

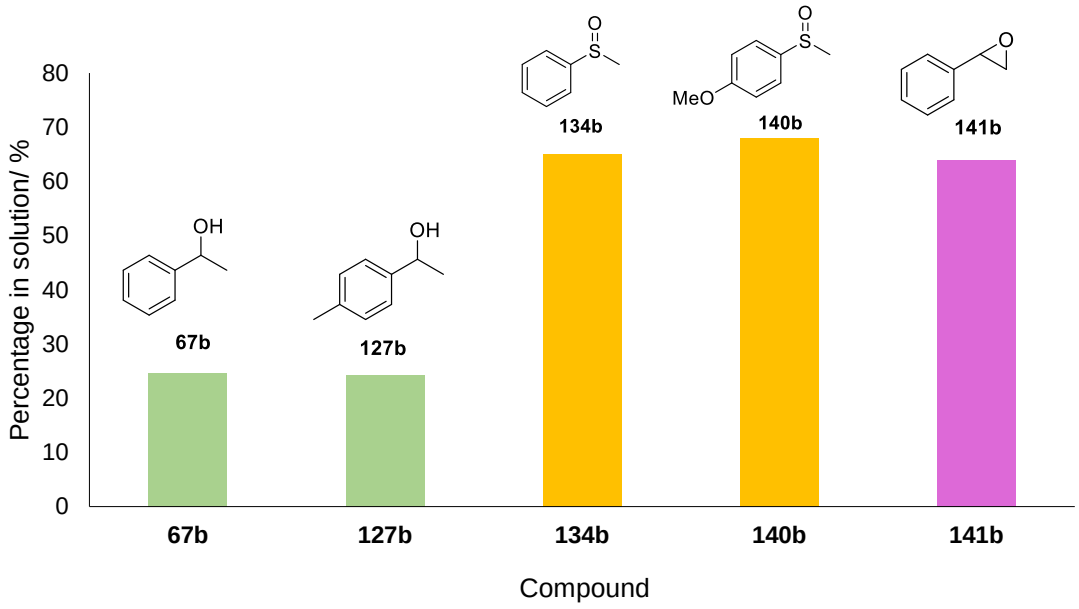


Figure 10.28 Biotransformations carried out by artUPO$_{bact}$ with hydroxylations in green, sulfoxidations in orange and epoxidations in purple. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 0.1 U mL^{-1} UPO, 2 mM substrate with either 10% ACN or 1% EtOH final concentration, 2 mM H_2O_2, 250 rpm, 20 °C, 2 h

In these reactions, a lower activity of UPO towards NBD was added to the reactions as limited amounts of artUPO$_{bact}$ could be made. Despite the lower activity, the conversions of the compounds to their oxidation products are relatively similar to the conversions carried out by artUPO$_{yeast}$. artUPO$_{bact}$ appears to carry out sulfoxidations and epoxidations more efficiently than hydroxylations; this pattern is also observed for

artUPO$_{yeast}$. The biotransformations catalysed by artUPO$_{bact}$ also contained either 10% ACN or 1% EtOH in the reactions, but artUPO$_{bact}$ did not seem to be hampered significantly by the presence of solvent.

10.2.5. Evaluation of artUPO as a biocatalyst

To date, *Aae*UPO is the most studied UPO and its catalytic capabilities have been thoroughly investigated. It is therefore useful to compare the catalytic potential of artUPO with *Aae*UPO in order to assess its potential in biocatalysis.

Both forms of artUPO has been shown to be able to catalyse hydroxylations, sulfoxidations and epoxidations of alkylbenzenes, phenyl thioethers and styrene derivatives respectively. In the hydroxylation screens, r*Aae*UPO transformed the alkylbenzenes **67a** (H), **126a** (3-Me), **127a** (4-Me) and **128a** (4-F) with much higher conversions than artUPO$_{yeast}$ with 10 mM substrate loading. These conversions were accompanied with *ee* values of over 90% to form the (*R*)-alcohol. If one desired to form (*R*)-**67b**, -**126b**, -**127b** or -**128b** then r*Aae*UPO would be preferable to artUPO$_{yeast}$. artUPO$_{yeast}$ may be a more suitable biocatalyst for the production of the other alcohols in the screen but these conversions were still low, ranging from 5 – 25% to form the respective alcohol. However, as shown in Figure 10.22, high yields of the benzylic ketone can be achieved with artUPO$_{yeast}$ if the substrate loading is lowered to 2 mM. r*Aae*UPO used in this study and wild type *Aae*UPO do not appear to be able to oxidise alcohols in high quantities, [127,198] and so artUPO could be a favourable biocatalyst for the formation of benzylic ketones.

Conversions of sulfides to sulfoxides between r*Aae*UPO and artUPO$_{yeast}$ are relatively similar, except for 4-methoxythioanisle **139a** where artUPO$_{yeast}$ transforms this substrate roughly five times more than r*Aae*UPO. Of particular interest, artUPO$_{yeast}$ has complementary enantioselectivity to r*Aae*UPO where the (*S*)-sulfoxide is preferentially formed instead of the (*R*)-sulfoxide with r*Aae*UPO. The difference in the enantioselectivity of the UPOs shown in these experiments may have applications for the synthesis of pharmaceutical compounds. For example, antiulcer drug esomeprazole has a chiral sulfoxide in the (*S*)-configuration and has greater activity than the (*R*)-configuration, omeprazole (Figure 10.29). [199] Another pharmaceutical which has a chiral sulfoxide is the wakefulness medication armodafinil (Figure 10.29). The (*R*)-enantiomer is believed to be more effective as it remains in the plasma for longer throughout the day in patients compared to the (*S*)-enantiomer, modafinil. [200] These examples demonstrate the need to produce both (*R*)- and (*S*)-sulfoxides. The complementary enantioselectivities of artUPO and r*Aae*UPO potentially allow for biocatalytic production of both enantiomers. This represents the beginning of an enzymatic toolbox for chiral sulfoxide production where artUPO and *Aae*UPO (and its variants) can be used as starting points for evolution towards the desired compounds.

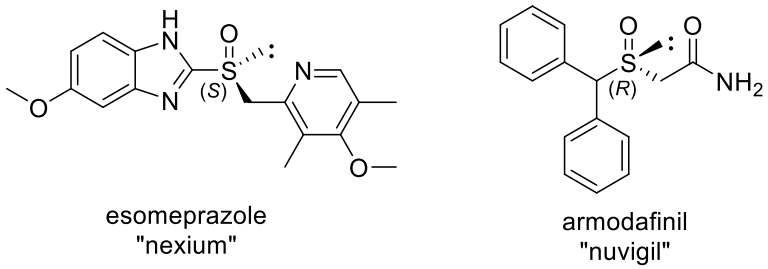


Figure 10.29 **Left** esomeprazole (brand name nexium) with the chiral sulfur centre in the (*S*)-configuration and **Right** armodanfinil (brand name nuvigil) with the chiral sulfur centre in the (*R*)-configuration

artUPO$_{yeast}$ was able to transform styrene **141a** and *cis*-β-methylstyrene **88a** with high conversions of ≥ 90%. Similar *ee* values are seen for both artUPO$_{yeast}$ and r*Aae*UPO in the epoxidation of styrene **141a** where (*R*)-**141b** is selectively formed. r*Aae*UPO poorly transformed styrene **141a** compared to *cis*-β-methylstyrene **88a**. This result was surprising as high conversions of **141a** with wild type *Aae*UPO but low *ee* values below 10% have been observed in other studies. [127,198] Further investigation could be carried out to explain the differences between the epoxidation of styrene with r*Aae*UPO (a.k.a Pada-I) and wild type *Aae*UPO.

The activity of the artUPO$_{bact}$ appears to be relatively similar compared to artUPO$_{yeast}$, despite the differences in glycosylation. This suggests that glycosylation in artUPO$_{yeast}$ may not affect the catalytic activity or stability of the enzyme under the reaction conditions. These findings indicate that differences in activity may not be a concern when selecting expression hosts for artUPO. Despite overall higher conversions and *ee* values seen with r*Aae*UPO than artUPO, the ease of bacterial expression and evolution may mean that artUPO is a more desirable biocatalyst than r*Aae*UPO, which currently requires eukaryotic expression systems. Bacterial expression of artUPO allows for rapid mutagenesis screening for directed evolution approaches, which may allow tailoring of the enzyme towards targeted synthesis. Although directed evolution is possible in eurkaryotic systems, this process is usually slower and more labour intensive. [94]

10.3. Summary of the catalytic characteristics of artUPO derived from yeast and bacteria

To investigate the catalytic activity of artUPO$_{yeast}$, initial screens containing ethylbenzene **67a**, cyclohexane **122a**, toluene **123a** and methyl *p*-tolyl sulfide **124a** were carried out on a 2 mL scale with 10 mM substrate, 10 mM hydrogen peroxide and analysed after 24 h of the beginning of the reaction by GC. artUPO$_{yeast}$ was able to transform ethylbenzene **67a**, toluene **122a** and methyl *p*-tolyl sulfide **124a** into the respective first and second oxidation products, but did not accept cyclohexane **122a** as a substrate. Optimisation experiments were carried out which resulted in the next set of biotransformations being carried out at 20 °C, shaken at 700 rpm and analysed after 2 h with 2 mM hydrogen peroxide.

Preliminary biotransformations including screens of ethylbenzene **67a** derivatives, thioanisole **134a** derivatives and styrene **141a** derivatives were carried out by artUPO$_{yeast}$ and r*Aae*UPO with 10 mM substrate, 2 mM hydrogen peroxide and analysed after 2 h by GC. Both artUPO and r*Aae*UPO could oxyfunctionalise most of the compounds with poor to moderate yields. Because the yields were lower than expected, the concentration of hydrogen peroxide for the next set of transformations was increased to 10 mM and the reaction time was increased to 6 h to allow for steady peroxide addition and avoid heme inactivation. The scale of the biotransformations were also increased because chiral GC analysis could not detect the small amounts of compounds that were extracted from the reaction.

In the 5 mL scale biotransformations, both artUPO$_{yeast}$ and r*Aae*UPO could transform alkylbenzenes into alcohols, thioethers into sulfoxides, and styrene derivatives into epoxides. 10 mM substrate loading was used and 10 mM H_2O_2 was added to the reaction over a period of 2 h. In the hydroxylation screen, artUPO$_{yeast}$ could hydroxylate most ethylbenzene **67a** derivatives with poor to moderate yields and the *ee* values ranged from 34% to 100%, forming the respective (*R*)-alcohol. r*Aae*UPO performed similarly to artUPO$_{yeast}$, however high conversions were observed with compounds ethylbenzene **67a** and 1-ethyl-4-fluorobenzene **128a** of 85% and 88% respectively. In the sulfoxidation screen, high conversions to the sulfoxide overall were observed by both UPOs. Most interestingly, artUPO$_{yeast}$ appeared to have opposite enantioselectivity to r*Aae*UPO, where artUPO$_{yeast}$ selectively formed the (*S*)-sulfoxide whereas r*Aae*UPO exclusively formed the (*R*)-sulfoxide. Epoxidations of styrene **141a** and *cis*-β-methylstyrene **88a** were achieved with high yields with artUPO$_{yeast}$. r*Aae*UPO demonstrated good conversion of *cis*-β-methylstyrene **88a** but had poor conversion (9%) for the epoxidation of styrene **141a**. Time courses of the biotransformations were carried out with either 2 mM substrate or 10 mM substrate. The results revealed that the substrate loading can greatly affect the products formed in the reaction, where lower substrate loadings could give rise to an increase in the formation of the second oxidations products i.e. ketones and sulfones. This effect was observed particularly in the sulfoxidation screen.

Finally a screen of five compounds was carried out using artUPO$_{bact}$ in order to probe its ability to oxyfunctionalise ethylbenzene derivates, thioethers and styrene **141a**. artUPO$_{bact}$ was able to demonstrate similar conversions compared to artUPO$_{yeast}$ and so both forms of artUPO could be considered equally for biotransformations.

11. Final Summary, Conclusions, and Future Work

Unspecific peroxygenases have seen great interest in the last 20 years as potential industrial biocatalysts because of their ability to activate inert C-H bonds using only hydrogen peroxide as a cosubstrate. They are able to carry out a range of oxidations, such as hydroxylation, sulfoxidation, epoxidation, halogenation and heteroatom dealkylation. [78] Various studies of different homologs and mutants of UPOs have revealed their potential to transform a broad range of molecules and has also shown how mutagenesis work can enhance their reaction capabilities even further. [85,88] However, application of UPOs for wider use has been limited for several reasons including ease of heterologous expression, limited activity and stability, low substrate loadings and sensitivity to hydrogen peroxide. This study aims to broaden our understanding of UPOs by investigating the industrial potential of an artificial UPO (artUPO).

The expression of long UPOs has been limited to eurkaryotic hosts such as fungi and yeast, whereas it has been shown more recently that some short UPOs can also be expressed in bacterial hosts such as *E. coli*. [85] UPOs are extracellular enzymes which usually require post-translational modifications such as glycosylation and disulfide bond formation, therefore it would be valuable to understand how the expression of a given UPO is affected by different hosts that have different post-translational modification abilities.

This work focussed on two UPOs, an artificial UPO created by NovoZymes based on *Mro*UPO, [154] and *Cci*UPO, which has shown potential in its reaction scope in previous

studies. [107, 130, 134] The targets were cloned in *P. pastoris* and expression was attempted, but r*Cci*UPO was not successfully expressed. artUPO was successfully expressed in both shake flasks and fermentations resulting in artUPO$_{yeast}$. The gene for artUPO expression in *E. coli* was also expressed in shake flasks which yielded artUPO$_{bact}$. The expression of artUPO in *P. pastoris* and in *E. coli* has demonstrated the ability of artUPO to be used with two different, industrially relevant, host organisms from different kingdoms. The choice of a preferred expression host is a nuanced decision as both yeast and bacterial hosts are advantageous for different scenarios. For example, generating high yields of UPO are easily accomplished in *P. pastoris*, however rapid mutagenesis work can be carried out in *E. coli* due to a shorter doubling time, and is relatively more simple to grow and express recombinant protein. [94,98] To better compare the heterologous systems, fermentations with *E. coli* should be carried out and yields over time could be compared. If these yields are similar, *E. coli* would be an advantageous host due to the simplicity and speed of creating mutants compared to the work required to generate mutants in *P. pastoris*. However, the choice of expression host not only depends on yield but also depends on the activity of the resulting UPO.

UV-visible kinetic assays has revealed that there is little difference in peroxygenase activity between artUPO$_{bact}$ and artUPO$_{yeast}$ in reactions containing either solvent or buffer alone, despite differences in glycosylation. Because glycosylation was believed to improve the stability of UPOs, [89,104] this result was surprising. However to truly investigate the durability of both artUPOs in solvents, a solvent screen which varies the type of solvent, such as acetone, DMSO or methanol, and the percentage of solvent should be carried out before reaching any conclusions. Compatibility of

artUPO with increasing concentrations of solvents would be desirable as it would allow the use of artUPO with compounds that are not fully soluble in water. Further investigation of the effect of physical properties on the rate of reaction of both artUPOs, such as temperature and pH, would also be beneficial to our understanding of UPO stability and substrate compatibility. In addition to the substrates used in the kinetic assay experiments, more substrates could be screened to investigate more specific peroxygenase activity, such as naphthalene for aromatic hydroxylation which has been used in previous studies for *Aae*UPO, *Mro*UPO and the two bacterially expressed r*Cvi*UPO and r*Dca*UPO.[77, 158, 110] Another substrate that could be added to the screen is H_2O_2 to investigate the tolerance of both artUPOs to H_2O_2. This is because the current use of UPOs are limited due to their sensitivity to H_2O_2, therefore characterising these catalytic parameters may aid our understanding to improve H_2O_2 resistance. Although the characterisation of both artUPO's activity is important, analysing the protein structures and comparing with other UPO structures is equally vital. Bringing both together improves our knowledge about the structure-activity relationship of UPOs.

To determine the structure of artUPO, both forms were successfully crystallised and in total, three structures were solved using x-ray protein crystallography. Combined with analytical SEC, it was determined that both $artUPO_{bact}$ and $artUPO_{yeast}$ were dimers with the sub-units connected by a disulfide bond, formed from Cys232 on each monomer. The $artUPO_{yeast}$ structures showed *N*-glycosylation, where *N*-acetylglucosamine was covalently bonded to certain asparagine residues on the surface of the protein. $artUPO_{bact}$ had no such glycosylation as *E. coli* do not generally glycosylate recombinant proteins. All artUPO structures show features key to short

UPOs such as a heme group, a His-Glu stabilising pair seen in short UPOs and a structural Mg^{2+} ion. [88] The presence of mutations distal from the active site suggests that the modifications to artUPO may not directly affect the activity of the enzyme but were perhaps made to increase the stability. These modifications include changes from non-polar residues to polar residues on the surface of the protein, single mutations to increase hydrophobic interactions or introduce π stacking within the protein, and shortening flexible loops. A noticeable mutation is located near the opening of the substrate channel where Leu149 in *Mro*UPO is substituted to Lys156 in artUPO. This change may alter the substrate specificity of the enzyme, but further mutagenesis at this site would aid understanding in the importance of the chemical environment around residue 156 in artUPO. Solving the structure of artUPO allows for further study to be carried out to investigate the structure-activity relationship of the enzyme and hence other UPOs. For example, the structure of artUPO allows identification of the amino acids that point towards the active site, which can then be targeted for mutagenesis experiments such as iterative saturation mutagenesis (ISM) and combinatorial active-site tests (CASTing) which may improve the substrate specificity and enantioselectivity. [201, 202] The crystallographic data of artUPO could be used to improve the stability using the B-factor iterative test (B-FIT). B-FIT aims to improve protein stability by identifying flexible residues for mutagenesis based upon the B-factors from an x-ray crystal structure. [203, 204]

To explore the catalytic abilities of artUPO, both forms of artUPO were screened with ethylbenzene, thioanisole and styrene derivatives for hydroxylations, sulfoxidations and epoxidations respectively. These biotransformations were also carried out with r*Aae*UPO to compare artUPO's capabilities with a well characterised UPO. $artUPO_{yeast}$

was able to transform most of the compounds in the screens but to varying degrees. Hydroxylations of alkylbenzenes into the respective benzylic alcohols by artUPO$_{yeast}$ and r*Aae*UPO showed poor to moderate conversions,but artUPO$_{yeast}$ and r*Aae*UPO demonstrated enantioselectivity for the formation of the (*R*)-alcohol. artUPO$_{yeast}$ could readily accept styrene and *cis*-β-methylstyrene to form the respective epoxide in high yields, likewise artUPO$_{bact}$ was also very capable in epoxidising styrene to styrene oxide. Sulfoxidations of phenyl thioethers to the respective sulfoxides by the UPOs were generally converted with moderate to high percentages. A major finding in this work was that artUPO$_{yeast}$ was able to selectively form the (*S*)-sulfoxide whereas r*Aae*UPO selectively formed the (*R*)-sulfoxide in sulfoxidation reactions. This complementary enantioselectivity only broadens the potential for use of UPOs in industry, where different UPOs could be selected to carry out reactions with a favoured selectivity or could be engineered towards the synthesis of a particular target compound, such as pharmaceuticals, bioactive molecules or fine chemicals.

With the screens carried out in this work, artUPO$_{bact}$ was also able to oxyfunctionalise ethylbenzene derivatives, thioanisole derivatives and styrene with similar conversions to artUPO$_{yeast}$. These screens have demonstrated the ability of both artUPOs to transform a range of small molecules and have also shown that either form of artUPO may be suitable for these biotransformations. However the full screen that was carried out with artUPO$_{yeast}$ should also be carried out for artUPO$_{bact}$ for a better comparison, including the stereospecificity of each reaction.

The choice of substrates in the hydroxylation and sulfoxidation screens were designed to probe the steric and electronic effects of different substituents around the phenyl

ring on the rate of reaction, however when generating a preliminary Hammett plot, these effects on the rate were not very clear. To fully investigate these effects, reactions could be monitored using stopped-flow kinetics on a short time scale to obtain initial rates. With more accurate data dedicated to investigate substituent effects, a Hammett plot could be made to further the understanding of the structure-activity relationship between artUPO and its substrate scope.

The screens carried out were based upon three different substrates, ethylbenzene, thioanisole and styrene. Although these chosen substrates were useful in investigating Hammett-relationships, a broader substrate screen should be carried out to include different groups of molecules, such as fatty acids, terpenes, and pharmaeceutical compounds, in order to elucidate the substrate scope of artUPO$_{yeast}$ and artUPO$_{bact}$. In addition, a comparison can then be made to the substrate scope of *Mro*UPO which may aid our understanding in the importance of the L156K mutation on the effect of substrate selectivity. To determine the viability of the use of artUPO in industry, the reactions should be scaled up to reflect a more preparative scale.

Although not investigated in this project, the delivery of hydrogen peroxide is a major issue for the use of UPOs as biocatalysts. A recent study demonstrated a tethered system for *in situ* hydrogen peroxide generation. [148] This system could potentially be applied to artUPO in order to produce a self-sufficient variant of artUPO. artUPO could also be coupled to other *in situ* hydrogen peroxide generators as mentioned previously to control H_2O_2 delivery and hence increase the lifetime of artUPO.

This study has provided the groundwork in understanding the use of artUPO in industrial biotransformations by investigating potential expression systems, elucidating the structure of each artUPO variant, and characterised the activity and catalytic capabilities of artUPO. It is clear however that further investigation is still needed before artUPO can be considered ready and viable for industrial purposes. Despite this, artUPO has shown great potential to develop into a practical biocatalyst for future use in research and in industry.

12. Appendix

12.1. AlphaFold 2 results

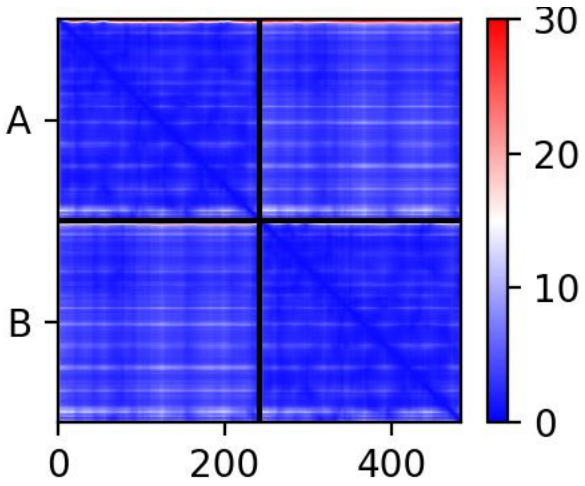


Figure 12.1 PAE plot for AlphaFold model artUPO. The residue number is on the x-axis, subunits of the y-axis

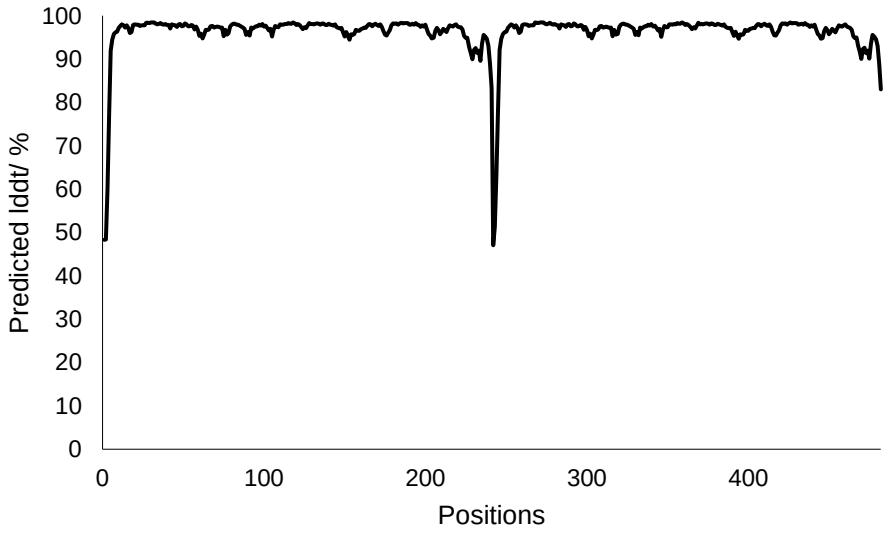


Figure 12.2 Predicted LDDT against residue number in AlphaFold model of artUPO

12.2. Non-enzyme control reactions

Table 12.1 Negative control (no enzyme) of reaction with compound **67a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	67a	**67b**
0	100.0	0.0
2	100.0	0.0

Table 12.2 Negative control (no enzyme) of reaction with compound **125a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	125a	**125b**
0	100.0	0.0
2	96.8	3.2

Table 12.3 Negative control (no enzyme) of reaction with compound **126a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	126a	**126b**
0	100.0	0.0
2	100.0	0.0

Table 12.4 Negative control (no enzyme) of reaction with compound **127a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	127a	**127b**
0	100.0	0.0
2	98.1	1.1

Table 12.5 Negative control (no enzyme) of reaction with compound **127a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	127a	**127b**
0	100.0	0.0
2	98.1	1.1

Table 12.6 Negative control (no enzyme) of reaction with compound **128a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	128a	**128b**
0	100.0	0.0
2	98.1	1.1

Table 12.7 Negative control (no enzyme) of reaction with compound **129a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	129a	**129b**
0	100.0	0.0
2	100.0	100.0

Table 12.8 Negative control (no enzyme) of reaction with compound **130a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 10% ACN final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	130a	**130b**
0	100.0	0.0
2	100.0	100.0

Table 12.9 Negative control (no enzyme) of reaction with compound **134a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	134a	**134b**
0	100.0	0.0
2	96.4	3.6

Table 12.10 Negative control (no enzyme) of reaction with compound **135a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	135a	**135b**
0	100.0	0.0
2	95.1	4.1

Table 12.11 Negative control (no enzyme) of reaction with compound **124a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	124a	**124b**
0	100.0	0.0
2	99.2	0.8

Table 12.12 Negative control (no enzyme) of reaction with compound **137a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	137a	**137b**
0	100.0	0.0
2	100.0	0.0

Table 12.13 Negative control (no enzyme) of reaction with compound **138a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	138a	**138b**
0	100.0	0.0
2	100.0	0.0

Table 12.14 Negative control (no enzyme) of reaction with compound **139a**. Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 10 mM substrate with 1% EtOH final concentration, 2 mM H2O2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	139a	**139b**
0	100.0	0.0
2	98.7	1.3

Table 12.15 Negative control (no enzyme) of reaction with compound **141a.** Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 1 mM substrate with 10% ACN final concentration, 1 mM H_2O_2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	141a	**141b**
0	100.0	0.0
2	95.0	5.0

Table 12.16 Negative control (no enzyme) of reaction with compound **88a.** Reaction conditions are as follows: 50 mM potassium phosphate buffer pH 7.0, 1 mM substrate with 10% ACN final concentration, 1 mM H_2O_2, 700 rpm, 20 °C, 2 h

time/ h	% in solution	
	88a	**88b**
0	100.0	0.0
2	93.1	6.9

12.3. Examples of achiral GC analysis of biotransformations with 10 mM substrate loading

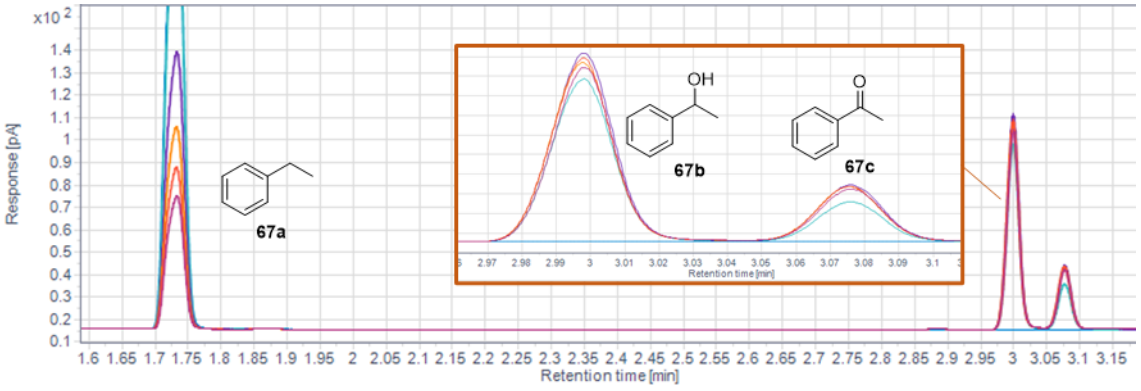


Figure 12.3 GC trace biotransformation of **67a** with artUPO over 6 h. **Blue**: 0 h, **teal**: 0.5 h, **indigo**: 1.0 h, **orange**: 2.0 h, **red**: 4.0 h, **maroon**: 6.0 h. Close up of the product formation included

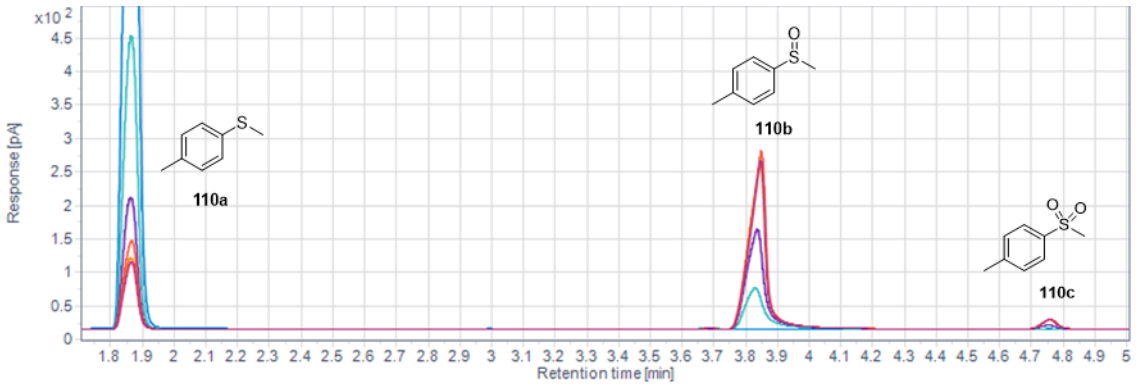


Figure 12.4 GC trace biotransformation of **110a** with artUPO over 6 h. **Blue**: 0 h, **teal**: 0.5 h, **indigo**: 1.0 h, **orange**: 2.0 h, **red**: 4.0 h, **maroon**: 6.0 h

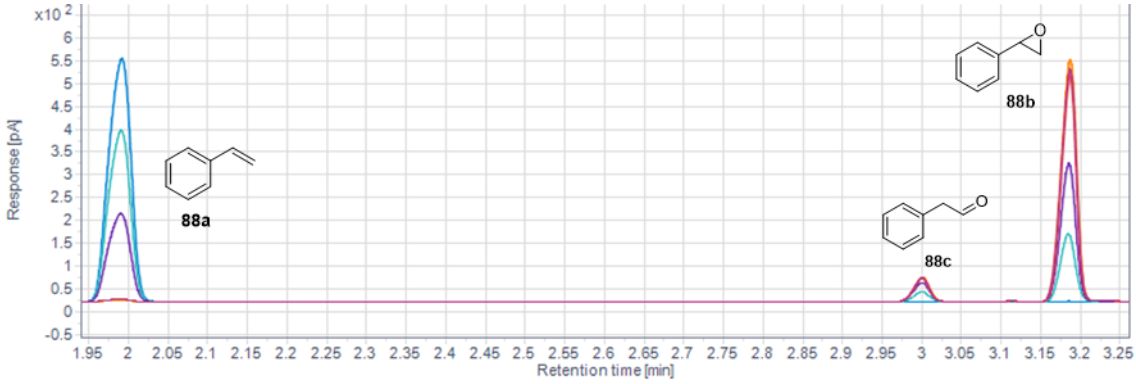


Figure 12.5 GC trace biotransformation of **128a** with artUPO over 6 h. **Blue**: 0 h, **teal**: 0.5 h, **indigo**: 1.0 h, **orange**: 2.0 h, **red**: 4.0 h, **maroon**: 6.0 h

12.4. Chiral GC analysis of oxidation products

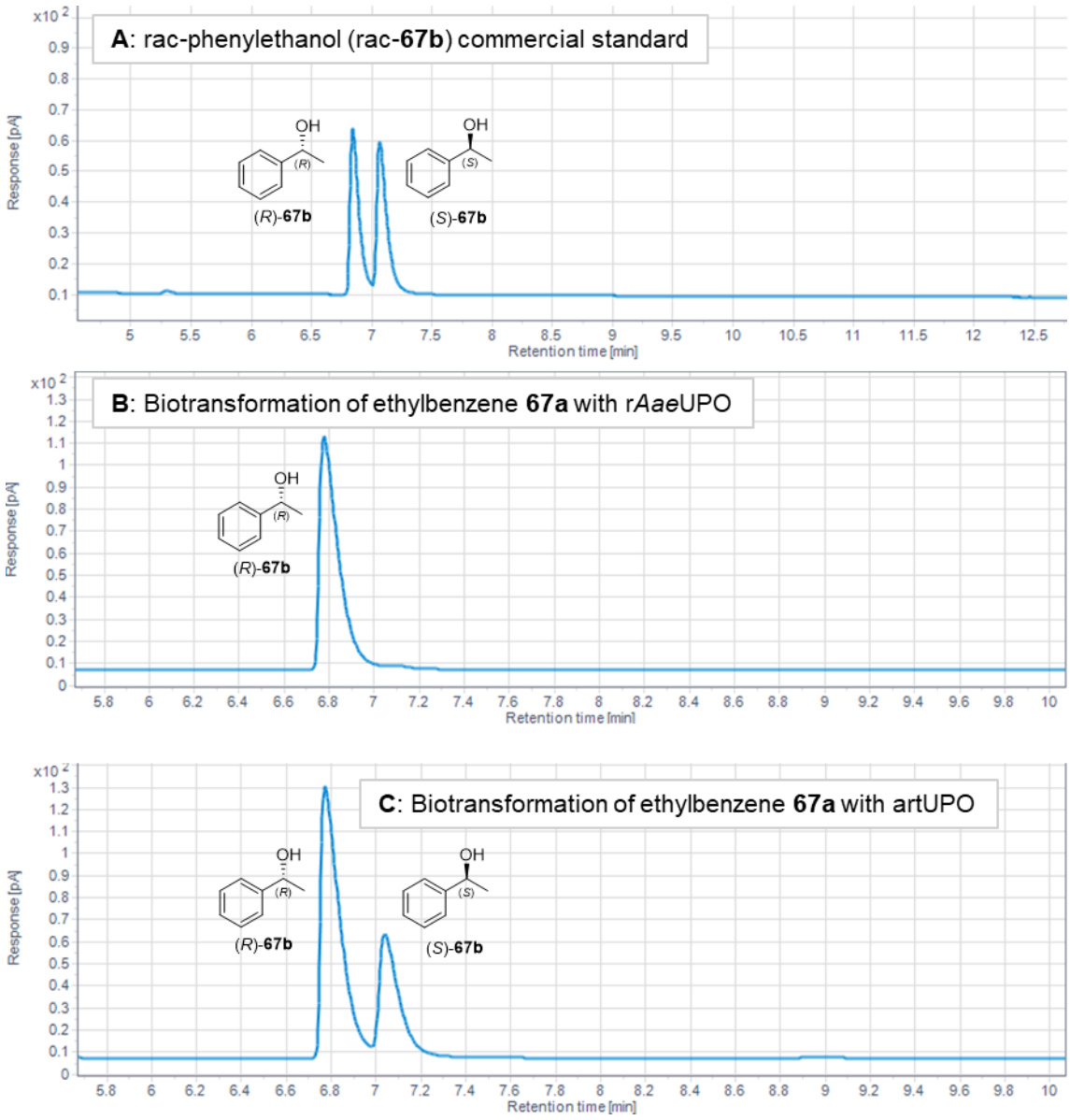


Figure 12.6 Chiral GC trace of **A**: rac-**67b**, **B**: biotransformation of **67a** with r*Aae*UPO and **C**: biotransformation of **67a** with artUPO

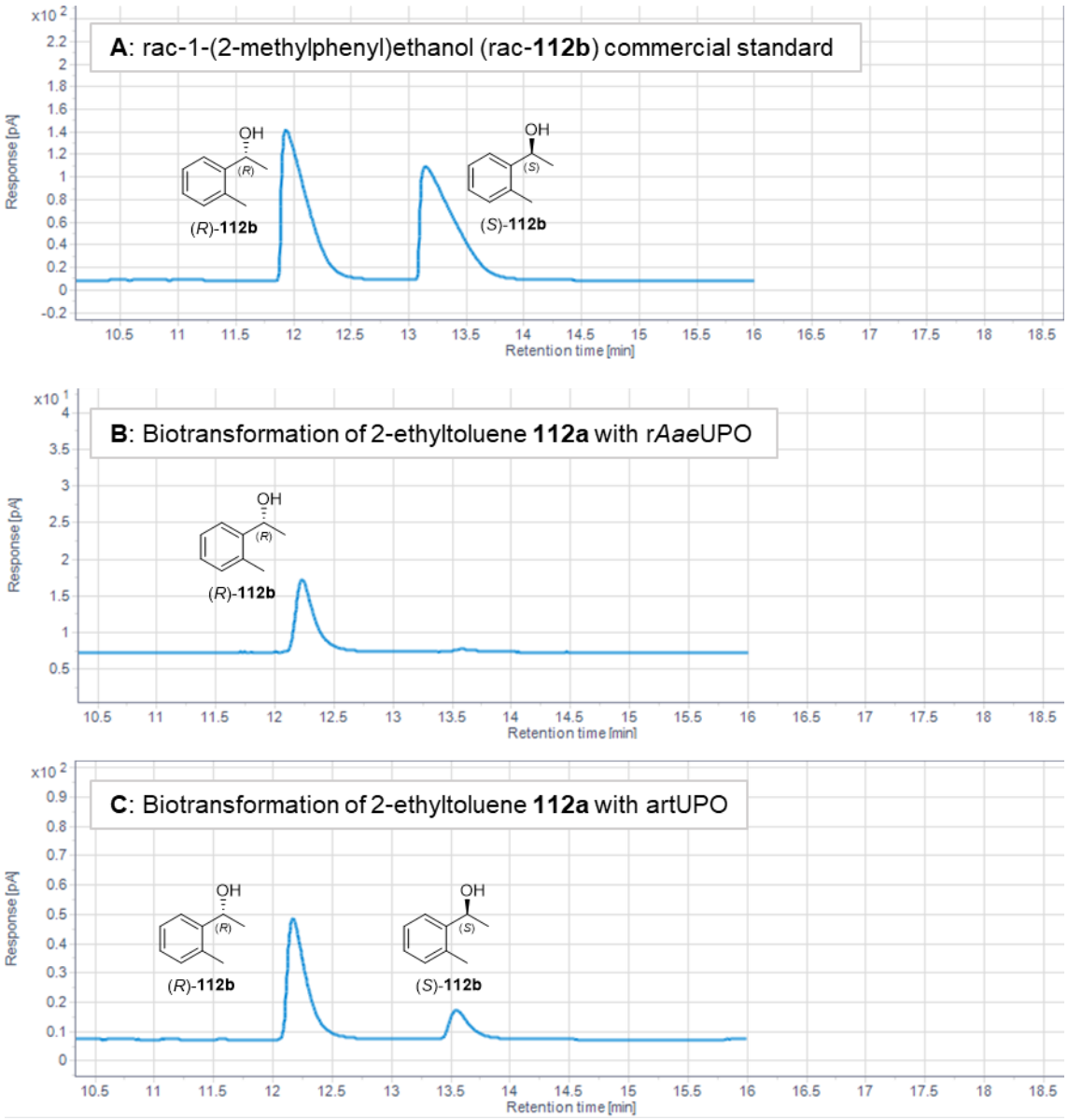


Figure 12.7 Chiral GC trace of **A**: rac-**112b**, **B**: biotransformation of **112a** with r*Aae*UPO and **C**: biotransformation of **112a** with artUPO

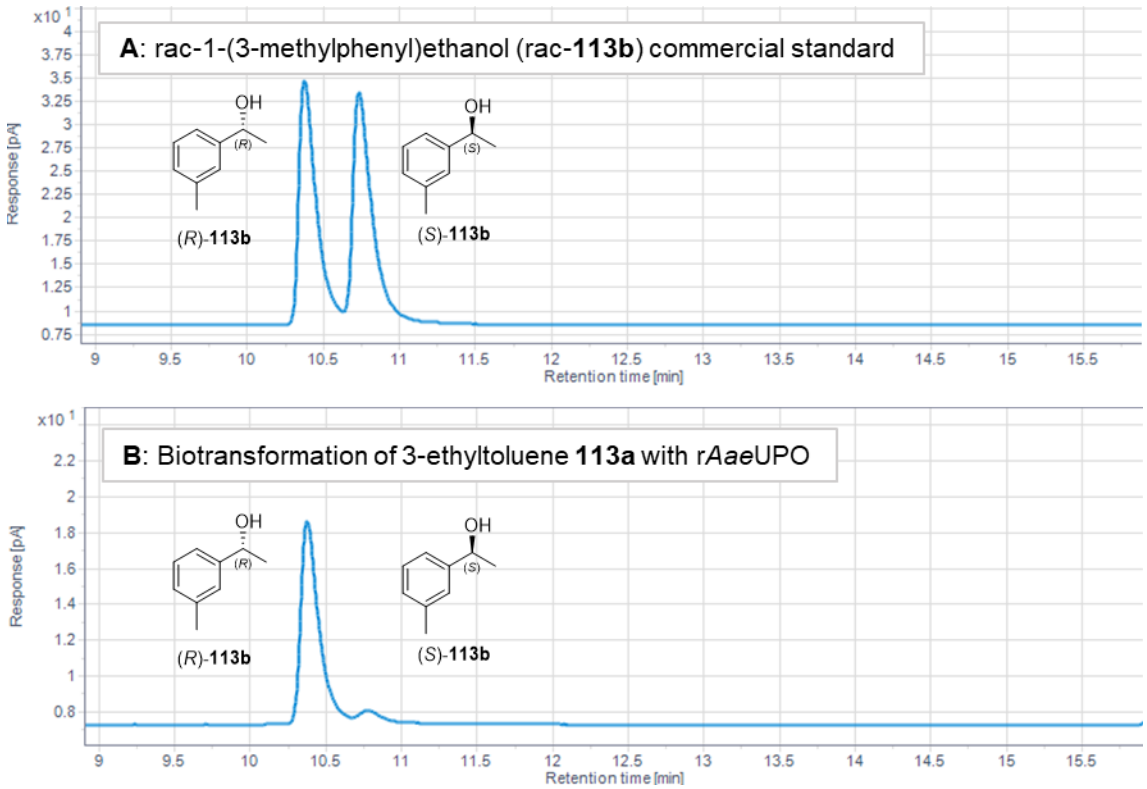


Figure 12.8 Chiral GC trace of **A**: rac-**113b** and **B**: biotransformation of **113a** with r*Aae*UPO

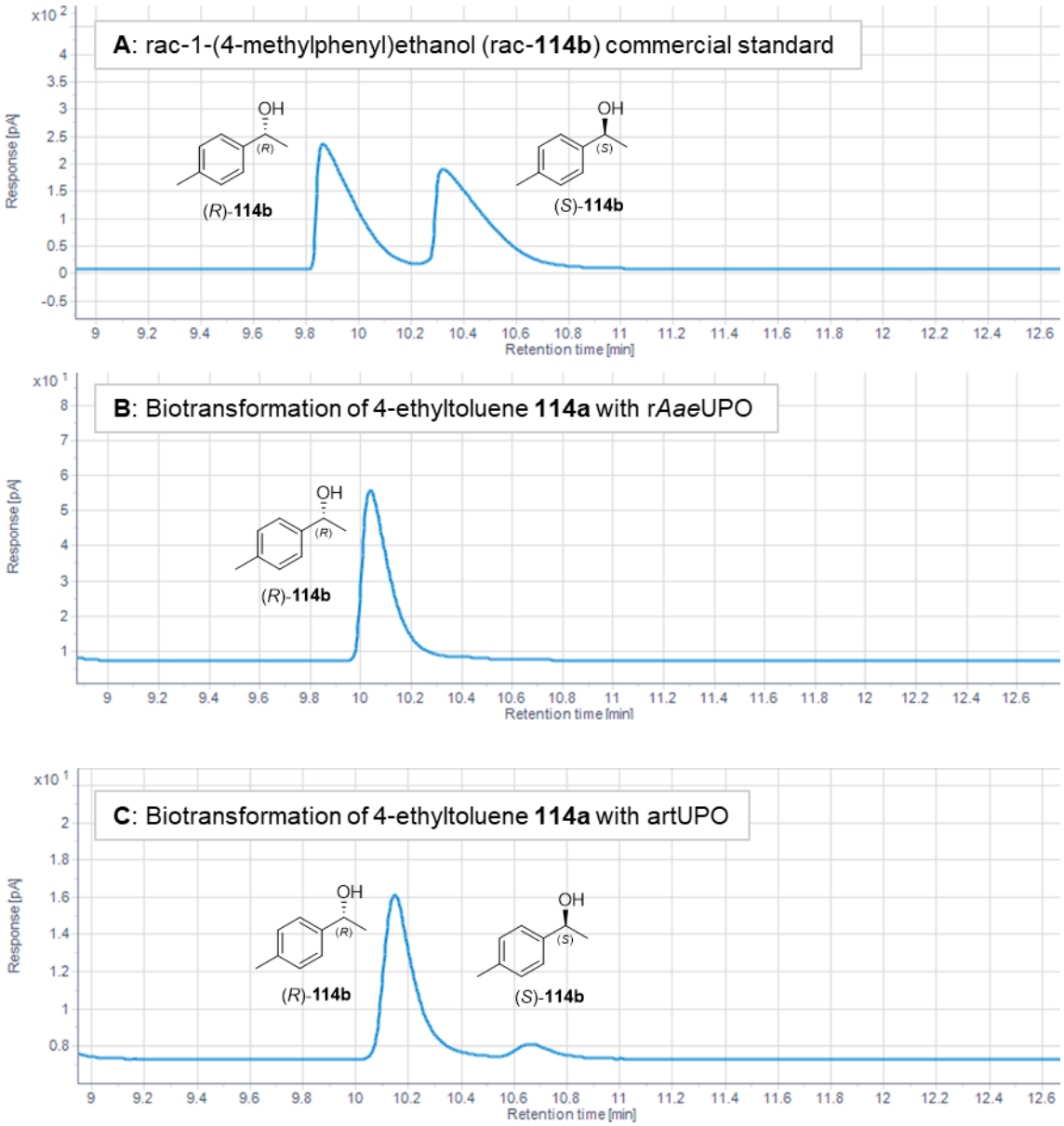


Figure 12.9 Chiral GC trace of **A**: rac-**114b**, **B**: biotransformation of **114a** with r*Aae*UPO and **C**: biotransformation of **114a** with artUPO

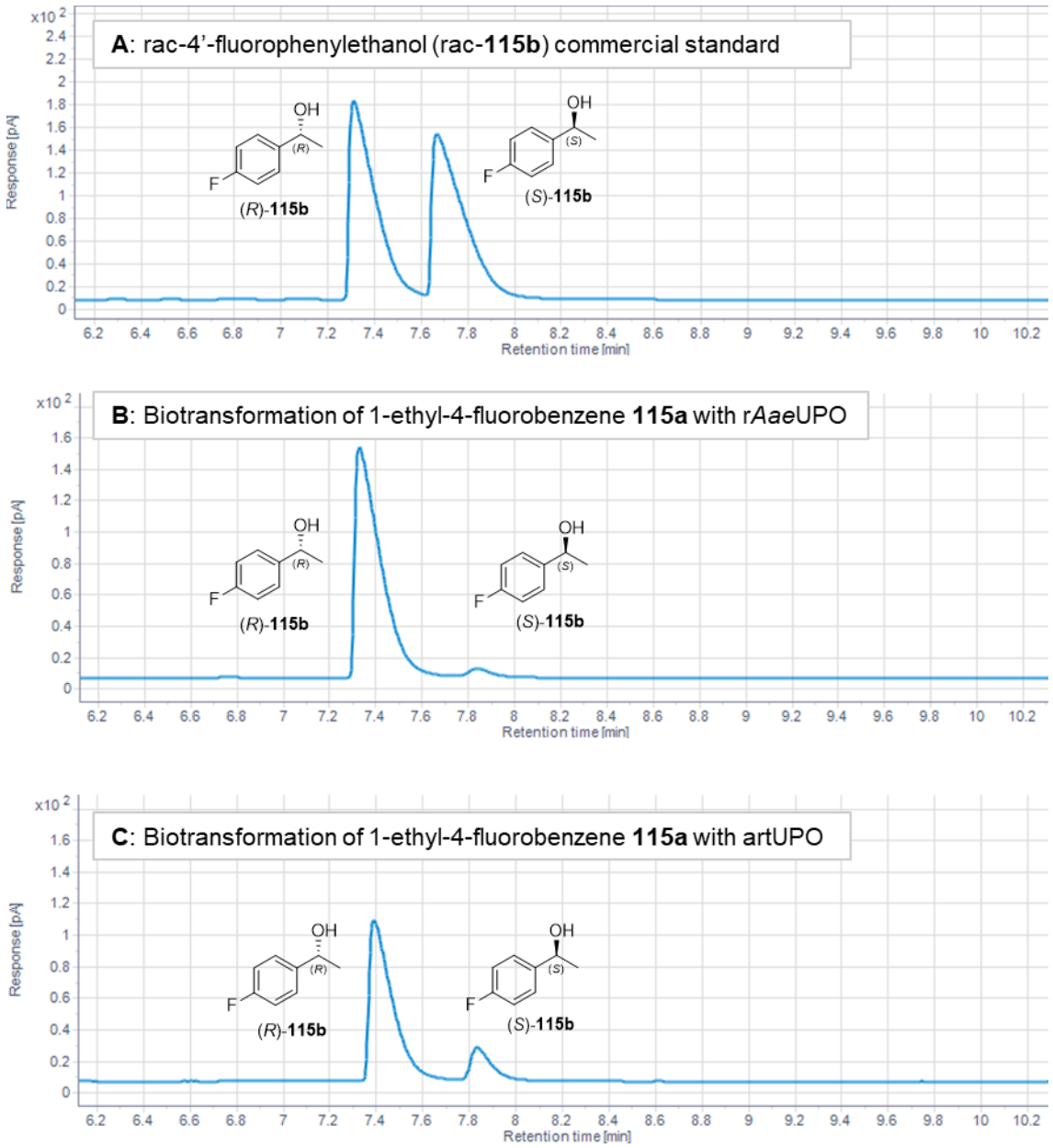


Figure 12.10 Chiral GC trace of **A**: rac-**115b**, **B**: biotransformation of **115a** with r*Aae*UPO and **C**: biotransformation of **115a** with artUPO

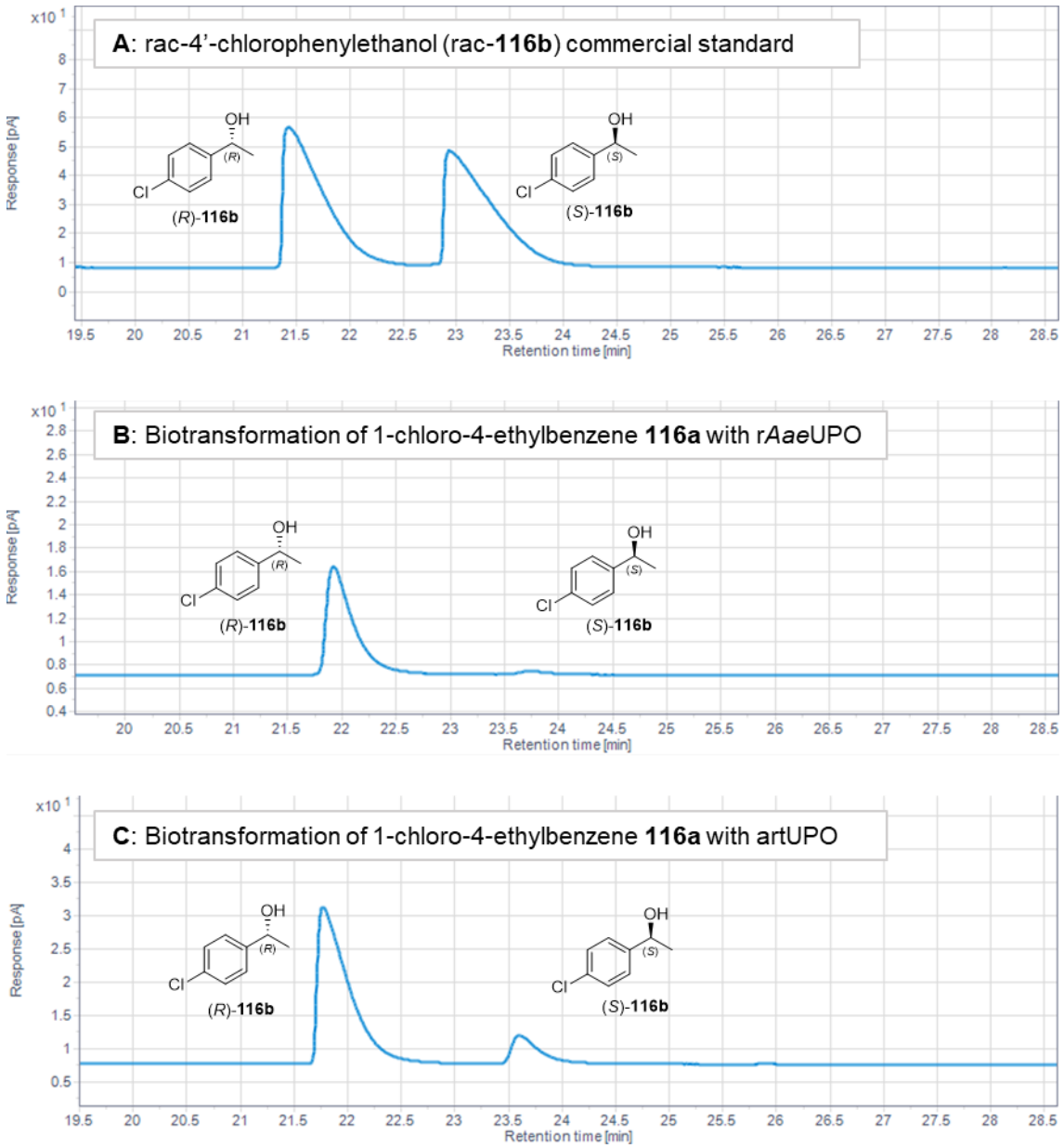


Figure 12.11 Chiral GC trace of **A**: rac-**116b**, **B**: biotransformation of **116a** with r*Aae*UPO and **C**: biotransformation of **116a** with artUPO

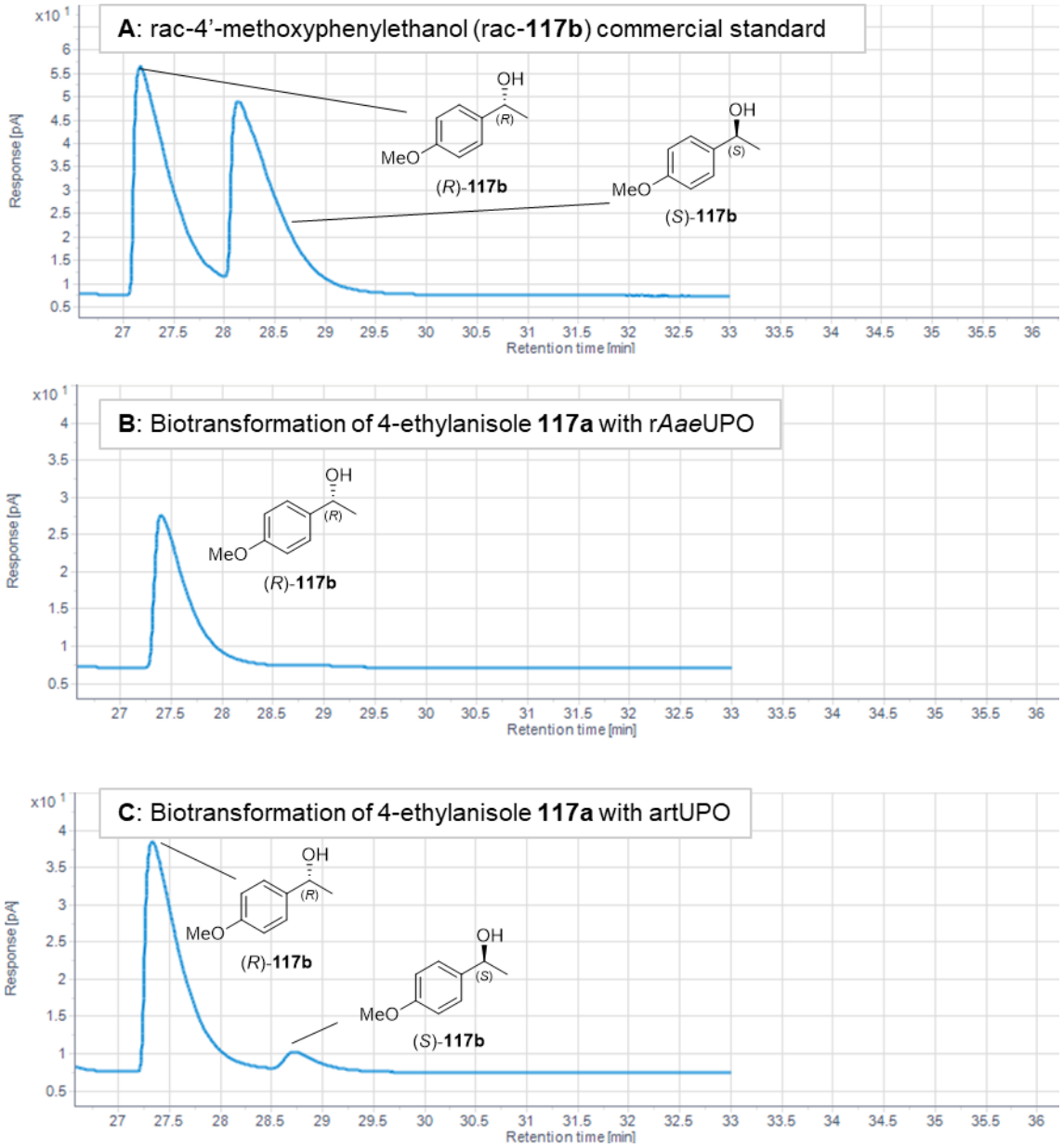


Figure 12.12 Chiral GC trace of **A**: rac-**117b**, **B**: biotransformation of **117a** with r*Aae*UPO and **C**: biotransformation of **117a** with artUPO

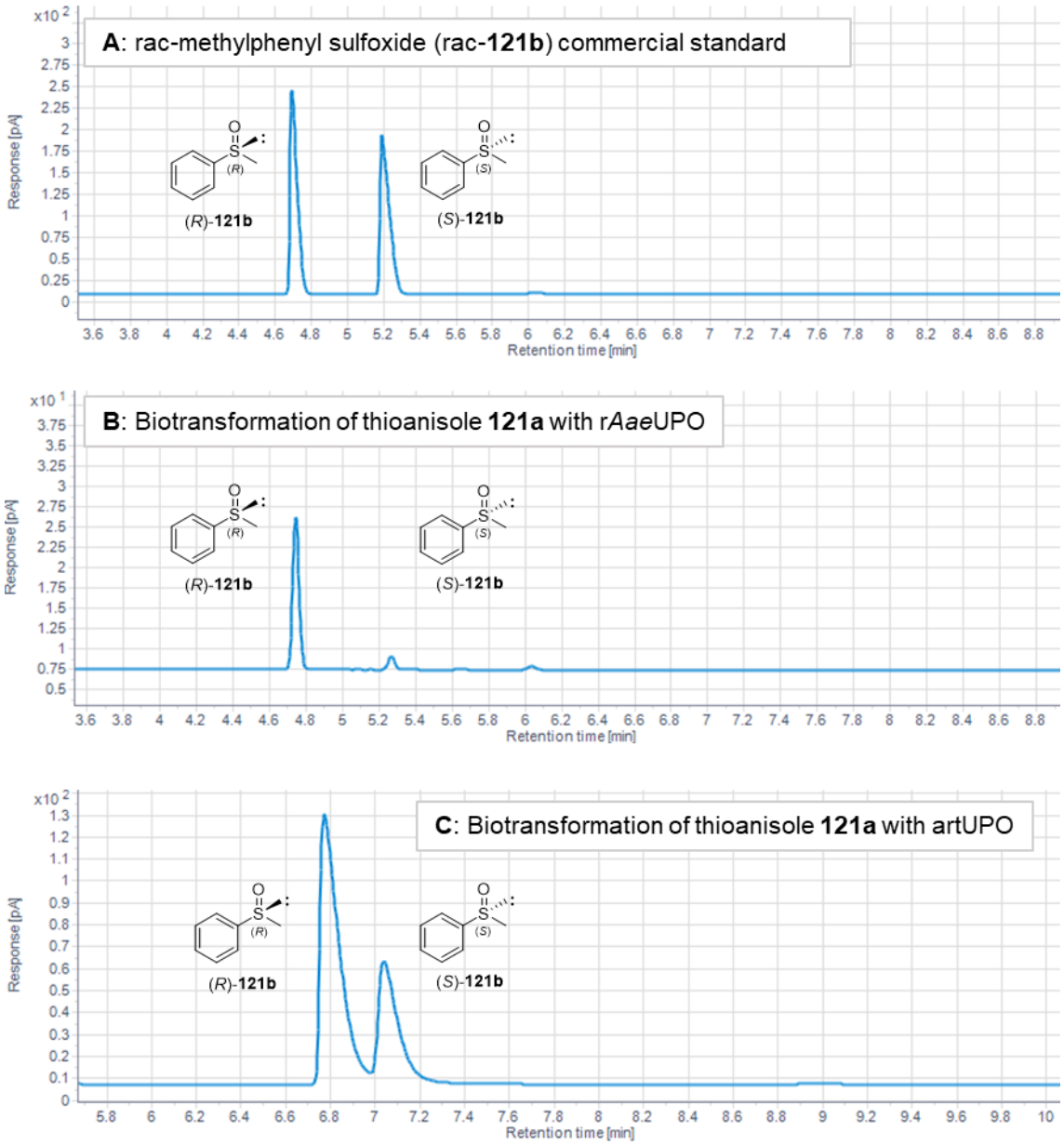


Figure 12.13 Chiral GC trace of **A**: rac-**121b**, **B**: biotransformation of **121a** with r*Aae*UPO and **C**: biotransformation of **121a** with artUPO

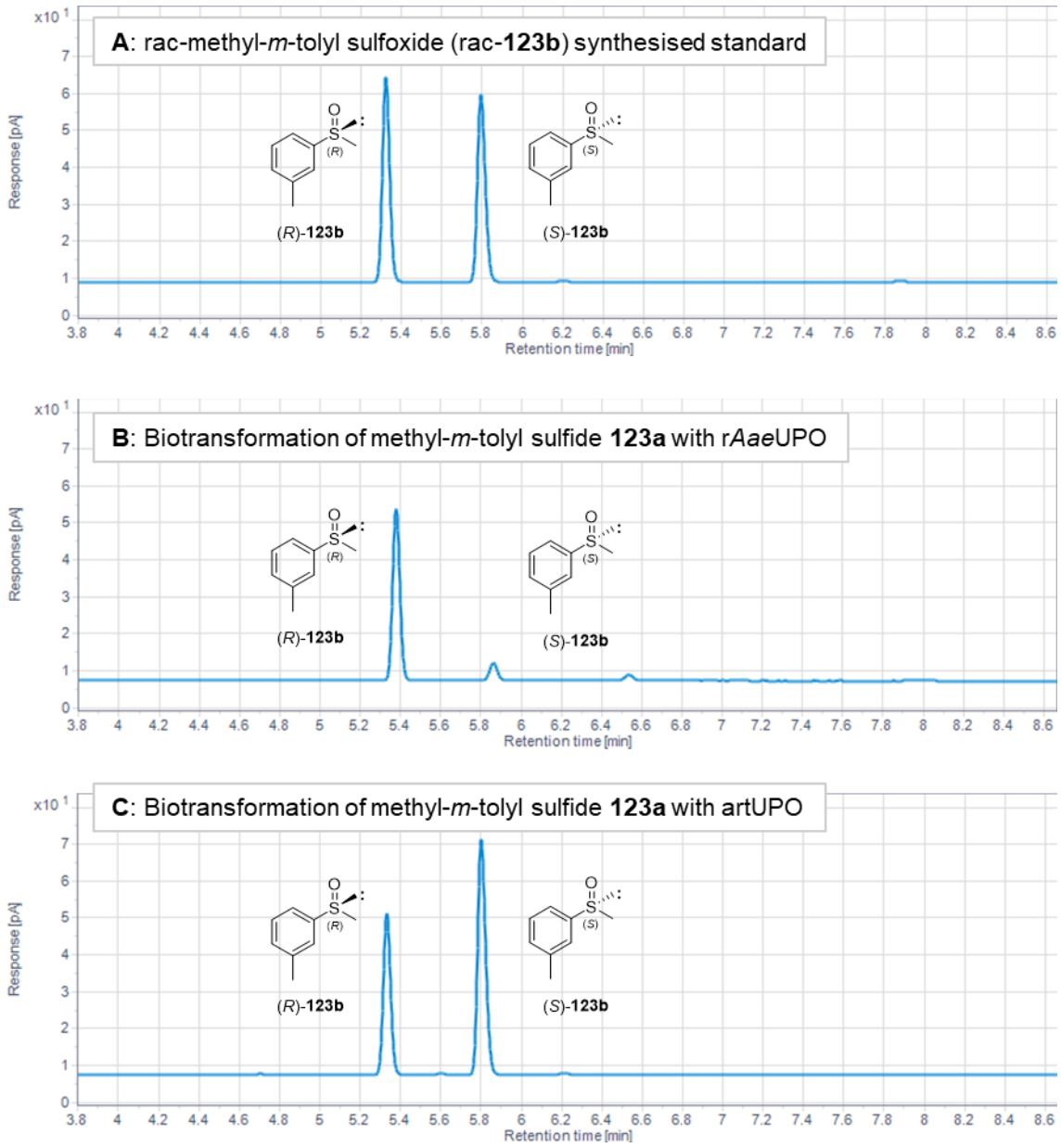


Figure 12.14 Chiral GC trace of **A**: rac-**123b**, **B**: biotransformation of **123a** with r*Aae*UPO and **C**: biotransformation of **123a** with artUPO

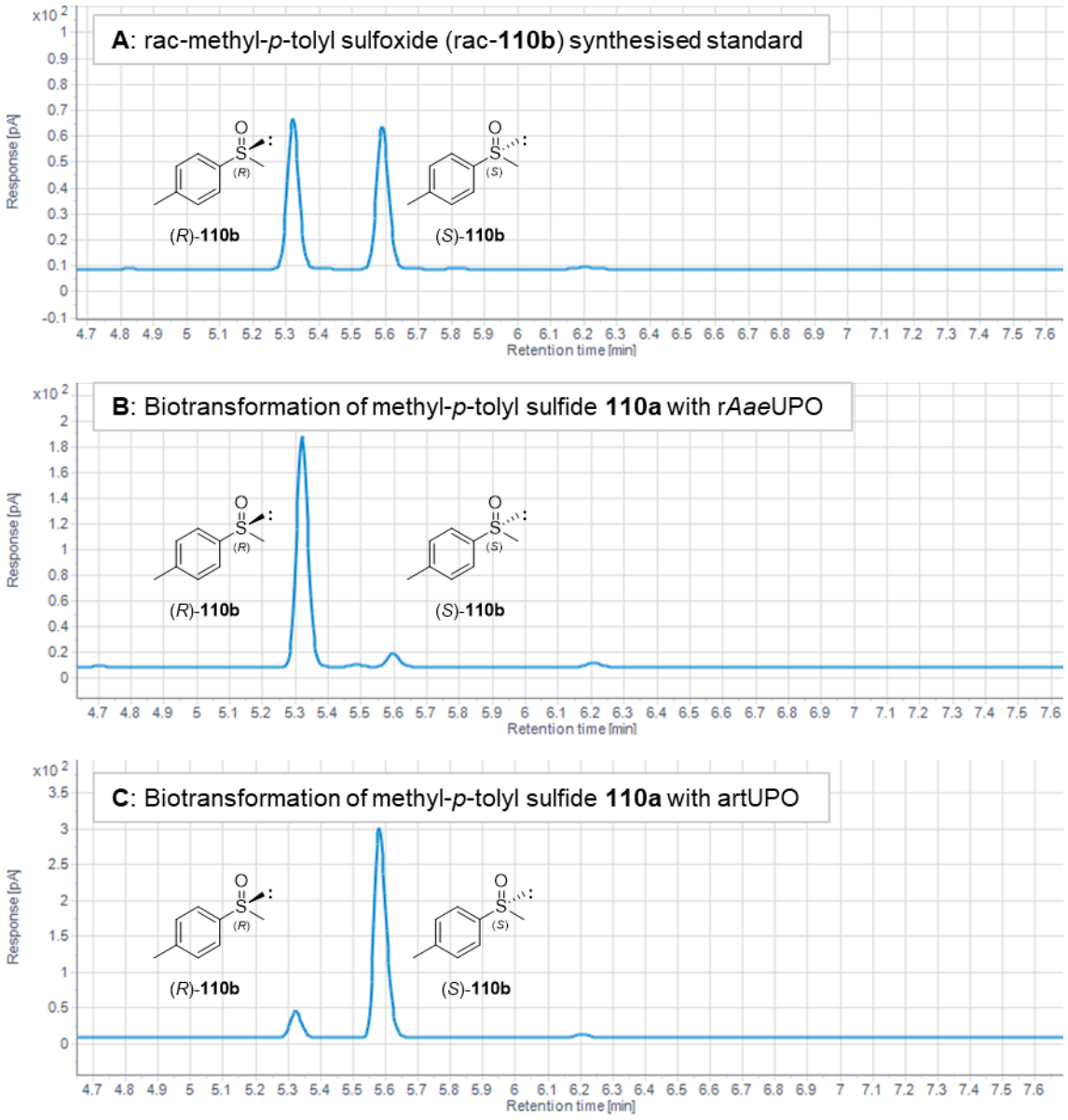


Figure 12.15 Chiral GC trace of **A**: rac-**110b**, **B**: biotransformation of **110a** with r*Aae*UPO and **C**: biotransformation of **110a** with artUPO

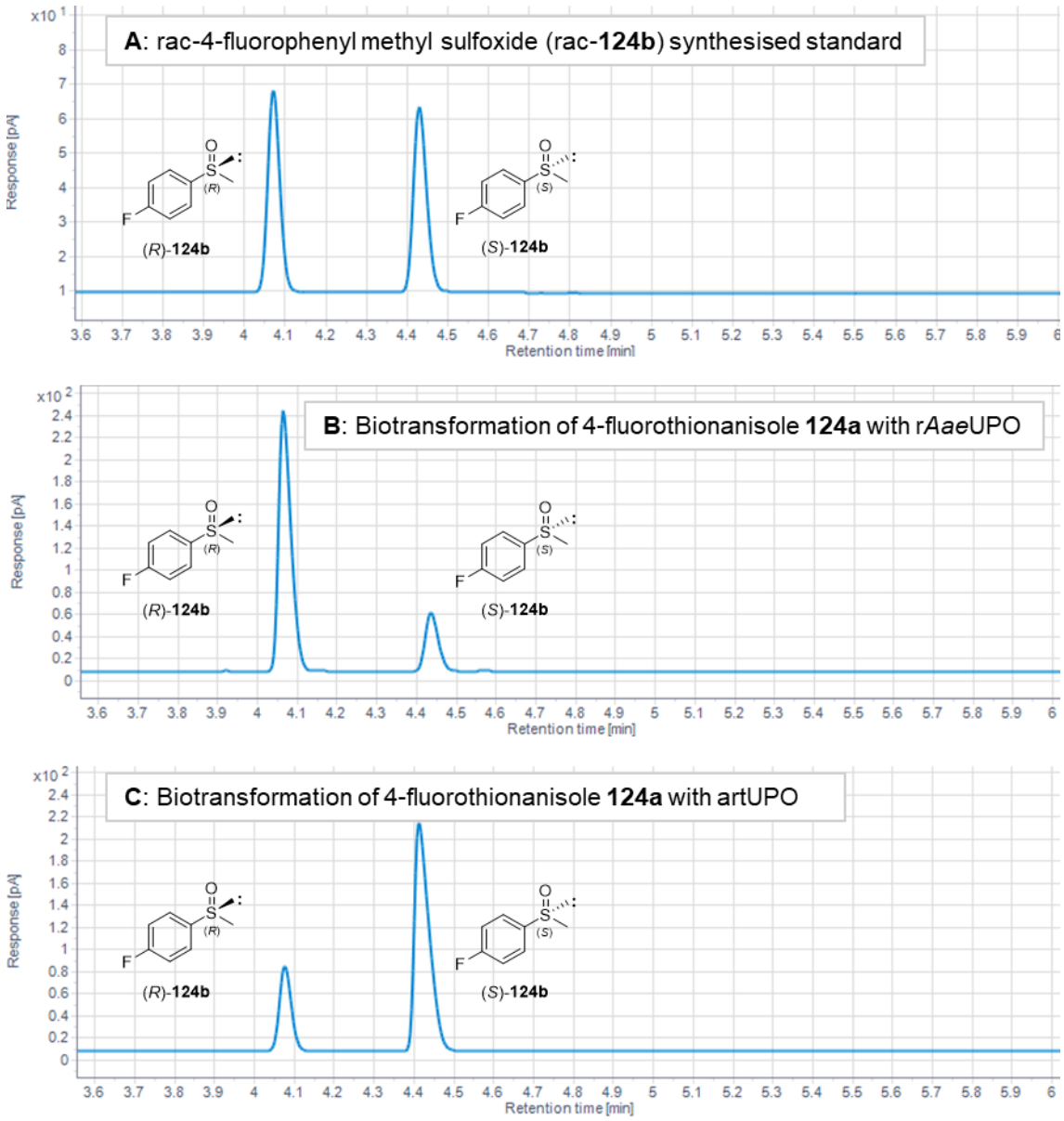


Figure 12.16 Chiral GC trace of **A**: rac-**124b**, **B**: biotransformation of **124a** with r*Aae*UPO and **C**: biotransformation of **124a** with artUPO

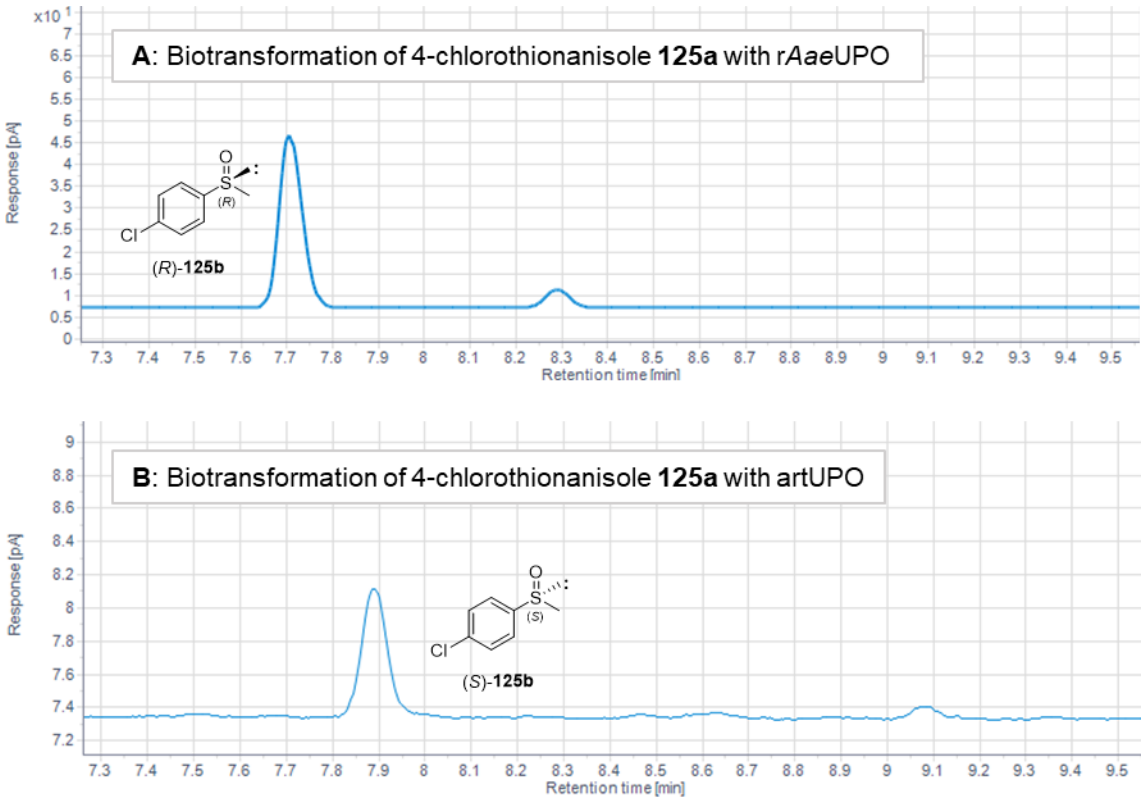


Figure 12.17 Chiral GC trace of **A**: biotransformation of **125a** with r*Aae*UPO and **B**: biotransformation of **125a** with artUPO

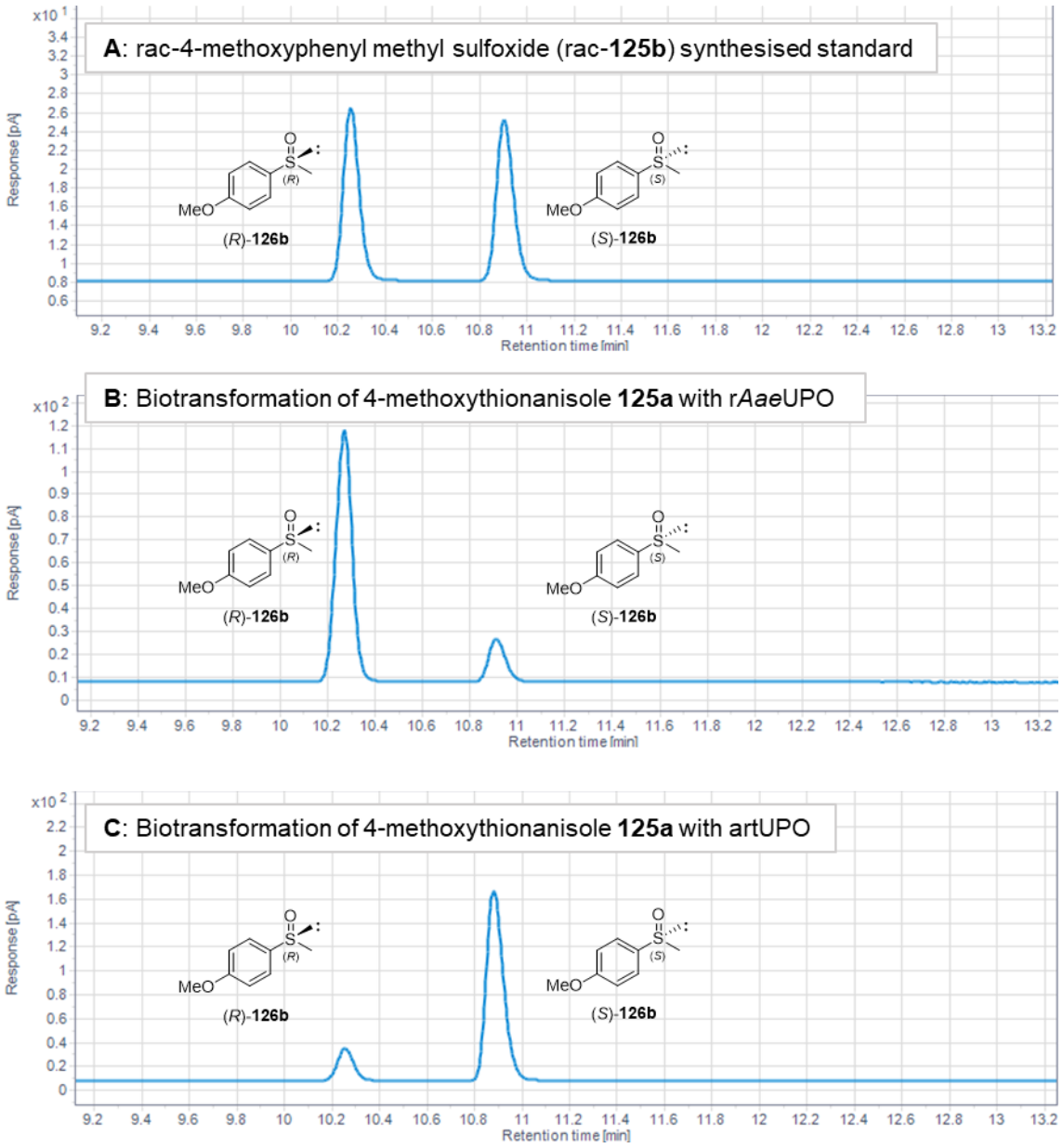


Figure 12.18 Chiral GC trace of **A**: rac-**126b**, **B**: biotransformation of **126a** with r*Aae*UPO and **C**: biotransformation of **126a** with artUPO

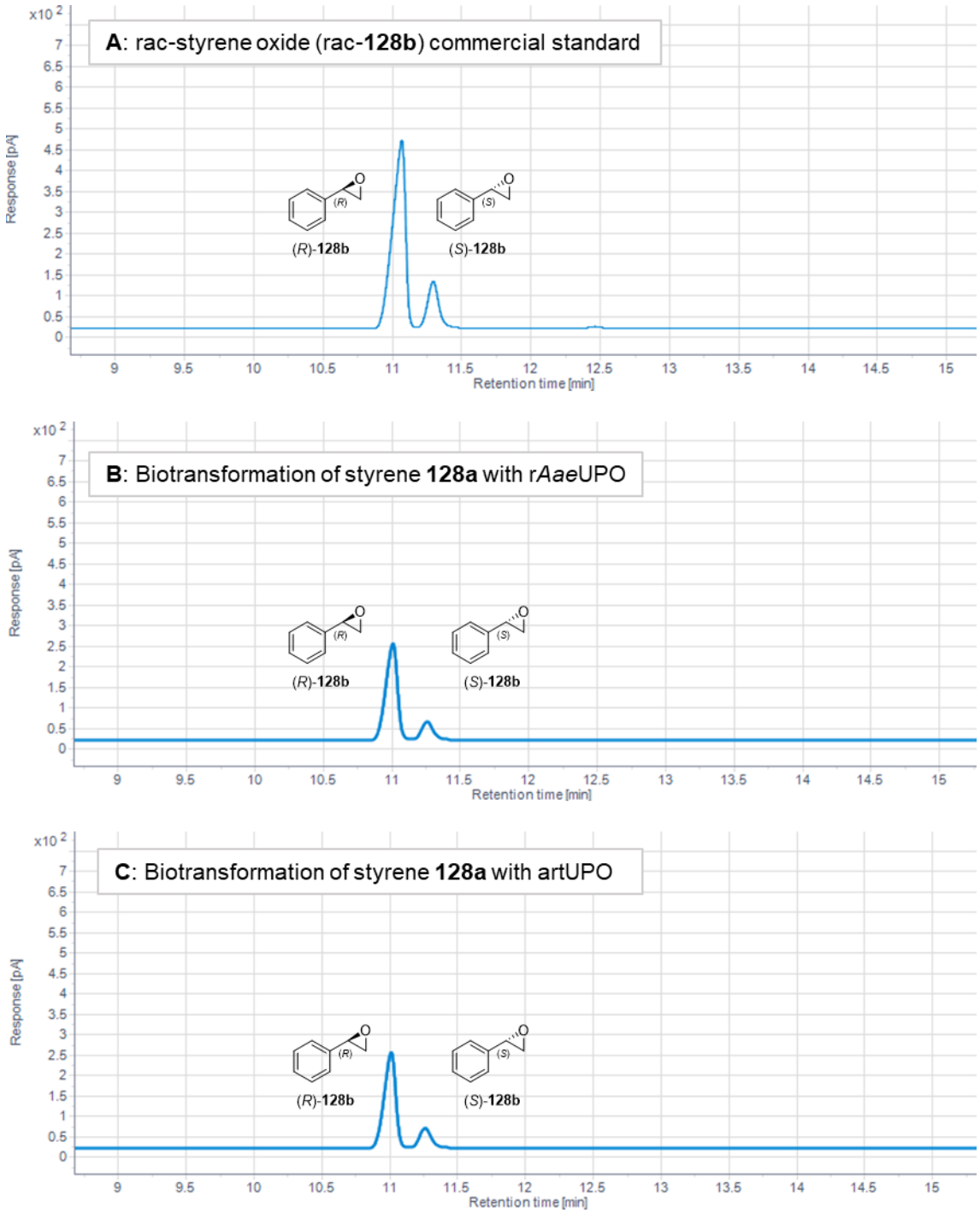


Figure 12.19 Chiral GC trace of **A**: rac-**128b**, **B**: biotransformation of **128a** with r*Aae*UPO and **C**: biotransformation of **128a** with artUPO

12.5. NMR spectra of synthesised products

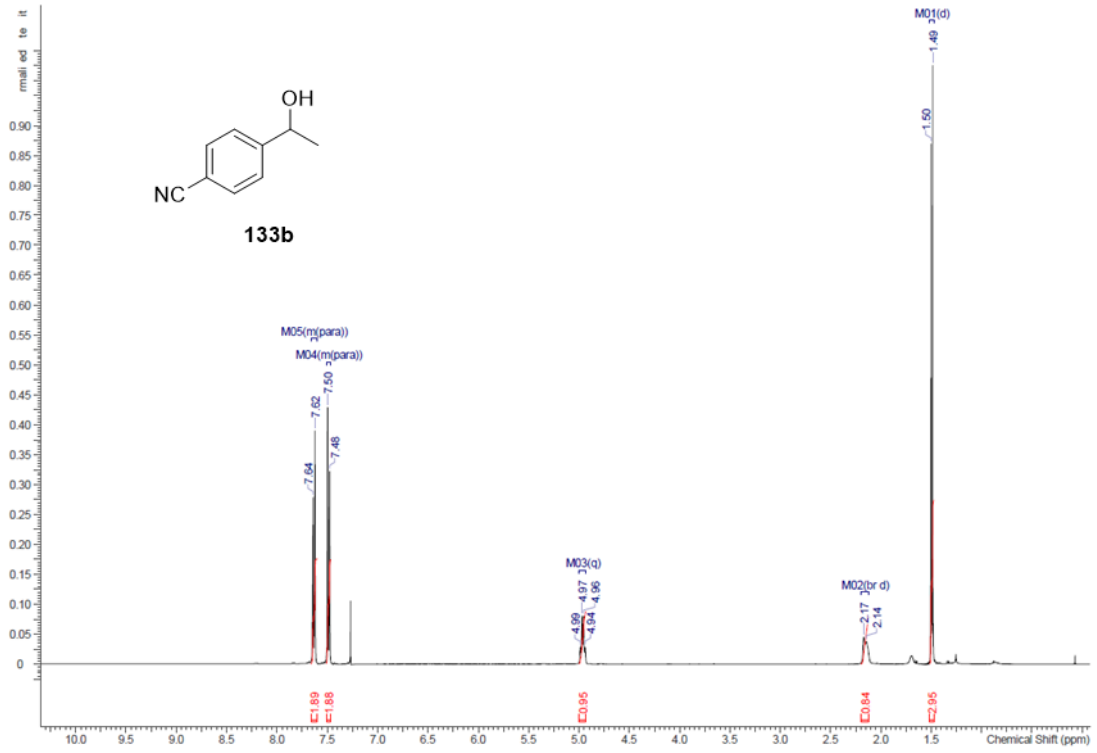


Figure 12.20 ^{1}H NMR spectrum (400 MHz) of 4-(1-hydroxyethyl)benzonitrile **133b** in chloroform-d

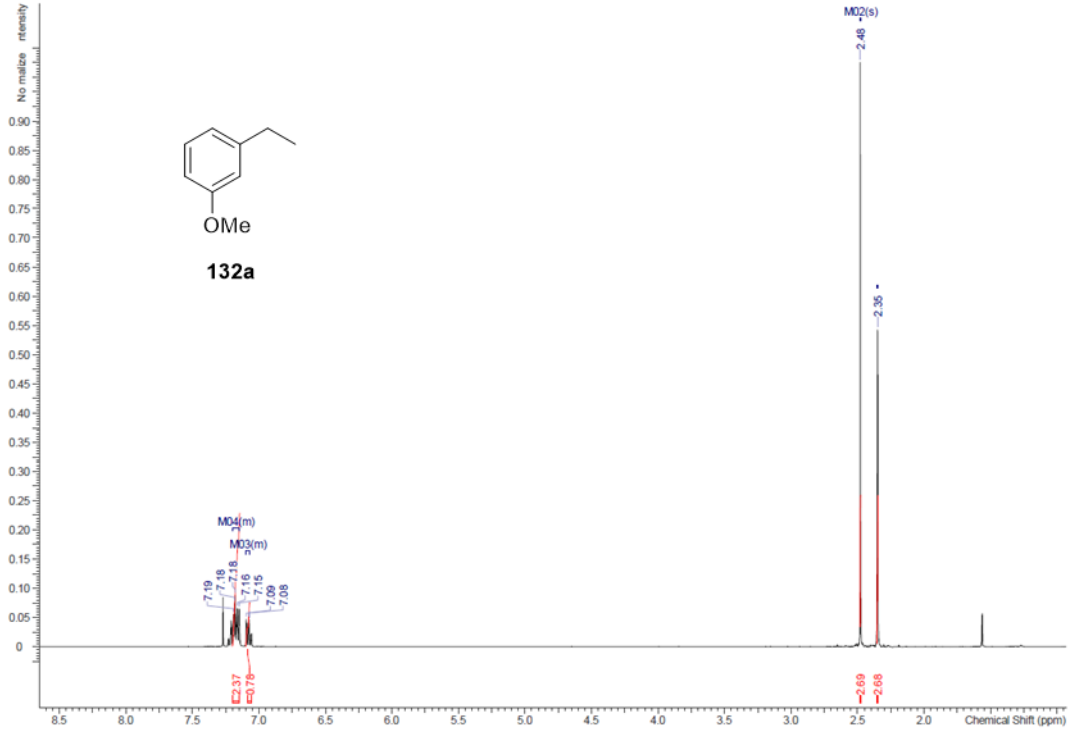


Figure 12.21 ^{1}H NMR spectrum (400 MHz) of 3-ethylanisole **132a** in chloroform-d

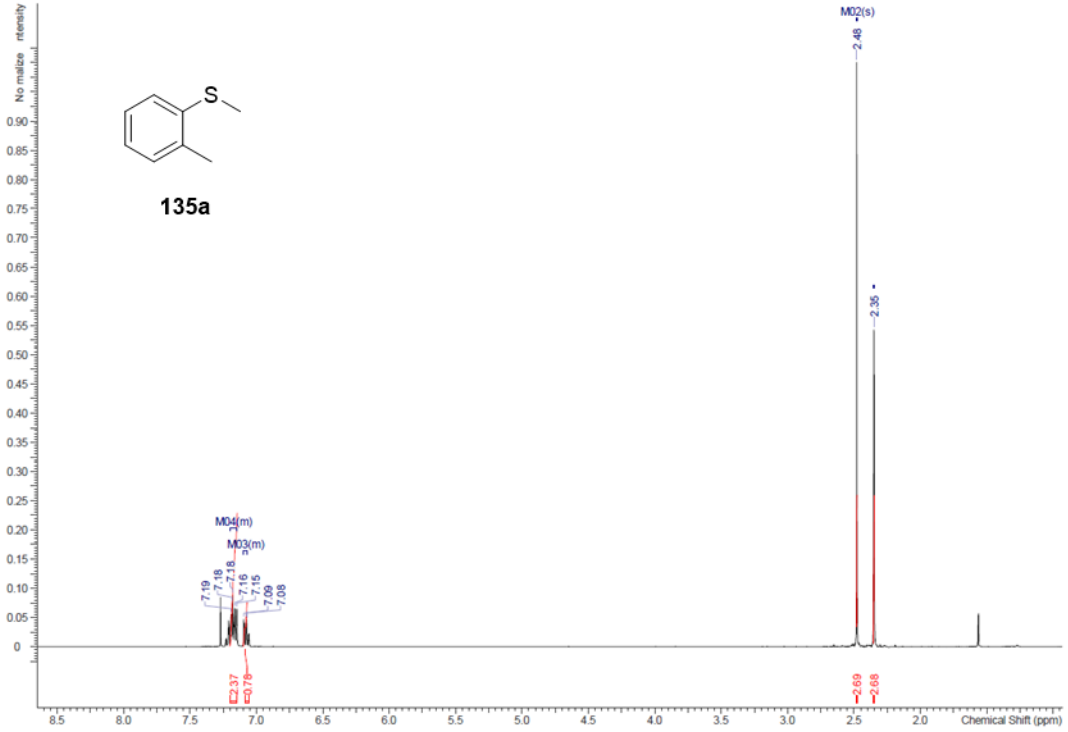


Figure 12.22 ^{1}H NMR spectrum (400 MHz) of methyl *o*-tolyl sulfide **135a** in chloroform-d

13. Abbreviations

(*v*/*v*)	Volume to volume ratio
(*w*/*v*)	Weight to volume ratio
[E_T]	Enzyme concentration
[S]	Substrate concentration
× g	Multiplied by g-force
°C	Degrees centrigrade
µ	Micro (10^{-6})
5-ALA	5-aminolevulinic acid
Å	Ångström (10^{-10})
*Aae*UPO	Unspecific peroxygenase derived from *Agrocybe aegerita* (a.k.a. *Cyclocybe aegerita*)
*Abi*UPO	Unspecific peroxygenase derived from *Agaricus bisporus*
ABTS	2,2-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
ACN	acetonitrile
AOX1	Alcohol oxidase
APS	Ammonium persulfate
artUPO	An artificial peroxygenase derived from *Mro*UPO and developed by NovoZymes
$artUPO_{bact}$	artUPO expressed in *E. coli*
$artUPO_{yeast}$	artUPO expressed in *P. pastoris*
B. megaterium	*Bacillus megaterium*
BMGY	Buffered glycerol complex medium
BMMY	Buffered methanol complex medium
bp	Base pair
BSA	Bovine serum albumin
*Cci*UPO	Unspecific peroxygenase derived from *Coprinius cinerea*
*Cfu*CPO	Chloroperoxidase derived from the fungus *Caldariomyces fumago* (a.k.a. *Leptoxyphium fumago*)
*Cgl*UPO	Unspecific peroxygenase derived from *Chaetomium globosum*
COD	cyclooctadiene

CPO	chloroperoxidase
CPR	Cytochrome P450 reductase
*Cra*UPO	Unspecific peroxygenase derived from *Coprinellus radians*
CSM	Combinatorial saturation mutagenesis
CV	Column volumes
*Cvi*UPO	Unspecific peroxygenase derived from *Collariella virescens*
CYP	Cytochrome P450 monooxygenase
Da	Dalton
*Dca*UPO	Unspecific peroxygenase derived from *Daldinia caldariorum*
DE3	Strain contains λDE3 lysogen
DFSMs	Dual-functional small molecules
DMF	dimethylformamide
DMP	2,6-dimethoxyphenol
DMPK	Drug metabolism pharmokinetic studies
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
dNTP	deoxyribonucleotide triphosphate
DO	Dissolved oxygen
DpnI	Restriction enzyme that digests methylated DNA derived from *Diplococcus pneumoniae*
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ee	Enantiomeric excess
EPA	Environmental Protection Agency
eq	equivalent
ESI-MS	Electrospray ionisation-mass spectrometry
eV	Electron volt
F	Farad
FAD	Flavin adenine dinucleotide reductase
FeS	ferredoxin
FF	Fast flow
FMN	Flavin mononucleotide dependent flavodoxin

FPLC	Fast protein liquid chromatography
g	grams
GC	Gas chromatography
GlcNAc	*N*-acetylglucosamine
h	hour
HDDA	12-hydroxydodecanoic acid
HP	High performance
HRV 3C	recombinant 3C protease derived from human *Rhinovirus type 14*
*Hsp*UPO	Unspecific peroxygenase derived from *Hypoxylon* strain
HTP	Heme-thiolate protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k	Kilo (10^3)
k_{cat}	Turnover number
K_M	Michaelis-Menten constant
L	litre
lac	lactose
LB	Lysogeny broth or Luria
ln	Natural logarithm
LRET	Long-ranged electron transfer
m	Milli (10^{-3})
M	Molar
m	metres
m	*meta*
M. oryzae	*Magnaporthe oryzae*
m/*z*	Mass to charge ratio
MCD	Monochlorodimedone (1,1-dimethyl-4-chloro-3,5-cyclohexanedione)
*m*CPBA	*Meta*-chloroperoxybenzoic acid
Me_4Phen	3,4,7,8-tetramethyl-1,10 phenanthroline
min	minute
MME	Monomethyl ether
MOPS	(3-(N-morpholino)propanesulfonic acid)

MORPHING	Mutagenic Organised Recombination Process by Homologous *in vivo* Grouping
*Mro*UPO	Unspecific peroxygenase derived from *Marasmius rotula*
MW	Molecular weight
*Mwe*UPO	Unspecific peroxygenase derived from *Marasmius wettstenii*
n	Nano (10^{-9})
n.d.	Not determined
n.t.	Not tested
NADH	Nicotinamide adenine dinulceotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NanoDSF	Nano differential scanning fluorimetry
NBD	1,2-(methylenedioxy-4-nitrobenzene)
NCS	*N*-chlorosuccinimide
NiNTA	Nickel-nitrilotriacetic acid
NMO	*N*-methyl morpholine-*N*-oxide
Nu	nucleophile
o	*ortho*
OD_{600}	Optical density at 600 nm
p	*para*
P. pastoris	*Pichia pastoris* (a.k.a. *Komagataella phaffii*)
P450	Cytochrome P450 monooxygenase
P450 BM3	Cytochrome P450 enzyme derived from *Bacillus megaterium*
P450 BS_{β}	Cytochrome P450 derived from *Bacillus subtilis*
P450 CAM	Cytochrome P450 derived from *Pseudomonas putida*
P450 CLA	Cytochrome P450 derived from *Clostridium acetobutylicum*
P450 $OleT_{JE}$	Cytochrome P450 derived from *Jeotgalicoccus* strain
P450 Rhf	Cytochrome P450 derived from *Rhodococcus* strain
P450 SP_{α}	Cytochrome P450 derived from *Sphingomonas paucimobilis*
PAE	Predicted aligned error
PAHs	Polycyclic aromatic hydrocarbons
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PELE	Protein energy landscape exploration

pET-28a(+)	*E. coli* expression vector
pLDDT	predicted local distance difference test
PPG	Polypropylene glycol
pPICZ	*P. pastoris* expression vector developed by Invitrogen
PTM	Predicted template modelling score
PTM_1 salts	Fermentation trace salts
Q	Anion exchange
QM/MM	Quantum mechanics/molecular mechanics
r.t.	Room temperature
r*Aae*UPO	Recombinantly expressed *Aae*UPO
r*Cci*UPO	Recombinantly expressed *Cci*UPO
r*Hin*UPO	Recombinantly expressed UPO derived from *Humicola insolens*
r*Mro*UPO	Recombinantly expressed *Mro*UPO
RMSD	Root mean square deviation
r*NOVO*	Recombinantly expressed UPO developed by NovoZymes
rpm	Rotations per minute
Rz	Reinheitzahl
s	second
S. carbophilus	*Streptomyces carbophilus*
S. cerevisiae	*Saccharomyces cerevisiae*
*Sac*I	Restriction enzyme derived from *Streptomyces achromogenes*
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SSM	Secondary structure matching
StEP	Staggered extension process
TAE	Tris, acetic acid, EDTA buffer
TBAF	tetrabutylammonium fluoride
TBHP	*tert*-butyl hydroperoxide
TBST	Tris-buffered saline solution with Tween-20
TEMED	Tetramethylethylenediamine
TTN	Total turnover number
UPO	Unspecific peroxygenase

UV	Ultraviolet radiation
UV/Vis	Ultraviolet/ visible radiation
V	Volts
V	Velocity
VA	Veratryl alcohol (3,4-dimethoxybenzyl alcohol)
V_{max}	Maximum velocity
*Xho*I	Restriction enzyme derived from *Xanthomonas holcicola*
YEP	Yeast extract peptone
YNB	Yeast nitrogen bases
YPD	Yeast peptone dextrose
YPDS	Yeast peptone dextrose sorbitol
Ω	Ohms

14. Bibliography

1 Nesterov, D. S., Nesterova, O. V. & Pombeiro, A. J. L. Homo- and heterometallic polynuclear transition metal catalysts for alkane C-H bonds oxidative functionalization: Recent advances. *Coord. Chem. Rev.* (2018), **355**, 199-222.

2 Ottenbacher, R. V., Talsi, E. P. & Bryliakov, K. P. Catalytic asymmetric oxidations using molecular oxygen. *Russ. Chem. Rev.* (2018), **87**, 821-830.

3 Molander, G. A. & Figueroa, R. cis-dihydroxylation of unsaturated potassium alkyl- and aryltrifluoroborates. *Org. Lett.* (2006), **8**, 75-78.

4 Hughes, D. L., Smith, G. B., Liu, J., Dezeny, G. C., Senanayake, C. H., Larsen, R. D., Verhoeven, T. R. & Reider, P. J. Mechanistic study of the Jacobsen asymmetric epoxidation of indene. *J. Org. Chem.* (1997), **62**, 2222-2229.

5 Chen, M. S. & White, M. C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* (2007), **318**, 783-787.

6 Li, B. J., Driess, M. & Hartwig, J. F. Iridium-Catalyzed Regioselective Silylation of Secondary Alkyl C-H Bonds for the Synthesis of 1,3-Diols. *J. Am. Chem. Soc.* (2014), **136**, 6586-6589.

7 Gribble, G. W. Biological Activity of Recently Discovered Halogenated Marine Natural Products. *Mar. Drugs* (2015), **13**, 4044-4136.

8 Latham, J., Brandenburger, E., Shepherd, S. A., Menon, B. R. K. & Micklefield, J. Development of Halogenase Enzymes for Use in Synthesis. *Chem. Rev.* (2018), **118**, 232-269.

9 Prakash, G. K. S., Mathew, T., Hoole, D., Esteves, P. M., Wang, Q., Rasul, G. & Olah, G. A. N-Halosuccinimide/BF3-H2O, efficient electrophilic halogenating systems for aromatics. *J. Am. Chem. Soc.* (2004), **126**, 15770-15776.

10 Petrone, D. A., Ye, J. T. & Lautens, M. Modern Transition-Metal-Catalyzed Carbon-Halogen Bond Formation. *Chem. Rev.* (2016), **116**, 8003-8104.

11 Alonso, F., Beletskaya, I. P. & Yus, M. Metal-mediated reductive hydrodehalogenation of organic halides. *Chem. Rev.* (2002), **102**, 4009-4091.

12 Dinh, A. N., Maddox, S. M., Vaidya, S. D., Saputra, M. A., Nalbandian, C. J. & Gustafson, J. L. Catalyst-Controlled Regioselective Chlorination of Phenols and Anilines through a Lewis Basic Selenoether Catalyst. *J. Org. Chem.* (2020), **85**, 13895-13905.

13 Schmid, A., Dordick, J. S., Hauer, B., Kiener, A., Wubbolts, M. & Witholt, B. Industrial biocatalysis today and tomorrow. (2001), **409**, 258-268.

14 Martinez, A. T., Ruiz-Duenas, F. J., Camarero, S., Serrano, A., Linde, D., Lund, H., Vind, J., Tovborg, M., Herold-Majumdar, O. M., Hofrichter, M., Liers, C., Ullrich, R., Scheibner, K., Sannia, G., Piscitelli, A., Pezzella, C., Sener, M. E., Kilic, S., van Berkel, W. J. H., Guallar, V., Lucas, M. F., Zuhse, R., Ludwig, R., Hollmann, F., Fernandez-Fueyo, E., Record, E., Faulds, C. B., Tortajada, M., Winckelmann, I., Rasmussen, J. A., Gelo-Pujic, M., Gutierre, A., del Rio, J. C., Rencoret, J. & Alcalde, M. Oxidoreductases on their way to industrial biotransformations. *Biotechnol. Adv.* (2017), **35**, 815-831.

15 Grogan, G. Hemoprotein Catalyzed Oxygenations: P450s, UPOs, and Progress toward Scalable Reactions. *JACS Au* (2021), **1**, 1312-1329.

16 Huang, X. Y. & Groves, J. T. Oxygen Activation and Radical Transformations in Heme Proteins and Metalloporphyrins. *Chem. Rev.* (2018), **118**, 2491-2553.

17 Frey, R., Hayashi, T. & Buller, R. M. Directed evolution of carbon-hydrogen bond activating enzymes. *Curr. Opin. Biotechnol.* (2019), **60**, 29-38.

18 Morris, D. R. & Hager, L. P. Chloroperoxidase .I. Isolation and properties of crystalline glycoprotein. *J. Biol. Chem.* (1966), **241**, 1763.

19 Hager, L. P., Morris, D. R., Brown, F. S. & Eberwein, H. Chloroperoxidase .II. Ultilization of halogen anions. *J. Biol. Chem.* (1966), **241**, 1769-&.

20 Bormann, S., Baraibar, A. G., Ni, Y., Holtmann, D. & Hollmann, F. Specific oxyfunctionalisations catalysed by peroxygenases: opportunities, challenges and solutions. *Catal. Sci. Technol.* (2015), **5**, 2038-2052.

21 Sundaramoorthy, M., Terner, J. & Poulos, T. L. The crystal structure of chloroperoxidase: A heme peroxidase-cytochrome P450 functional hybrid. *Structure* (1995), **3**, 1367-1377.

22 Dunford, H. B., Lambeir, A. M., Kashem, M. A. & Pickard, M. On the mechanism of chlorination by chloroperoxidase. *Arch. Biochem. Biophys.* (1987), **252**, 292-302.

23 Hofler, G. T., But, A. & Hollmann, F. Haloperoxidases as catalysts in organic synthesis. *Org. Biomol. Chem.* (2019), **17**, 9267-9274.

24 Doerge, D. R. & Corbett, M. D. Peroxygenation mechanism for chloroperoxidase-catalyzed N-oxidation of arylamines. *Chem. Res. Toxicol.* (1991), **4**, 556-560.

25 Zaks, A. & Dodds, D. R. Chloroperoxidase-catalyzed asymmetric oxidations - Substrate-specificity and mechanistic study. *J. Am. Chem. Soc.* (1995), **117**, 10419-10424.

26 Wang, Z. Y., Jian, Y. P., Han, Y. L., Fu, Z. W., Lu, D. N., Wu, J. Z. & Liu, Z. Recent progress in enzymatic functionalization of carbon-hydrogen bonds for the green synthesis of chemicals. *Chin. J. Chem. Eng.* (2020), **28**, 2499-2506.

27 Getrey, L., Krieg, T., Hollmann, F., Schrader, J. & Holtmann, D. Enzymatic halogenation of the phenolic monoterpenes thymol and carvacrol with chloroperoxidase. *Green Chem.* (2014), **16**, 1104-1108.

28 Yamada, H., Itoh, N. & Izumi, Y. Chloroperoxidase-catalyzed halogenation of trans-cinnamic acid and its derivatives. *J. Biol. Chem.* (1985), **260**, 1962-1969.

29 vanDeurzen, M. P. J., Remkes, I. J., vanRantwijk, F. & Sheldon, R. A. Chloroperoxidase catalyzed oxidations in t-butyl alcohol/water mixtures. *J. Mol. Catal. A-Chem.* (1997), **117**, 329-337.

30 Colonna, S., Gaggero, N., Carrea, G. & Pasta, P. A new enzymatic enantioselective synthesis of dialkyl sulfoxides catalysed by monooxygenases. *Chem. Commun.* (1997), 439-440.

31 Allenmark, S. G. & Andersson, M. A. Chloroperoxidase-induced asymmetric sulfoxidation of some conformationally restricted sulfides. *Chirality* (1998), **10**, 246-252.

32 Vargas, R. R., Bechara, E. J. H., Marzorati, L. & Wladislaw, B. Asymmetric sulfoxidation of a beta-carbonyl sulfide series by chloroperoxidase. *Tetrahedron-Asymmetry* (1999), **10**, 3219-3227.

33 Dexter, A. F., Lakner, F. J., Campbell, R. A. & Hager, L. P. Highly enantioselective epoxidation of 1,1-disubstituted alkenes catalyzed by chloroperoxidase. *J. Am. Chem. Soc.* (1995), **117**, 6412-6413.

34 Lakner, F. J., Cain, K. P. & Hager, L. P. Enantioselective epoxidation of omega-bromo-2-methyl-1-alkenes catalyzed by chloroperoxidase. Effect of chain length on selectivity and efficiency. *J. Am. Chem. Soc.* (1997), **119**, 443-444.

35 Park, J. B. & Clark, D. S. New reaction system for hydrocarbon oxidation by chloroperoxidase. *Biotechnol. Bioeng.* (2006), **94**, 189-192.

36 Hu, S. H. & Hager, L. P. Highly enantioselective propargylic hydroxylations catalyzed by chloroperoxidase. *J. Am. Chem. Soc.* (1999), **121**, 872-873.

37 vanDeurzen, M. P. J., vanRantwijk, F. & Sheldon, R. A. Chloroperoxidase-catalyzed oxidation of 5-hydroxymethylfurfural. *J. Carbohydr. Chem.* (1997), **16**, 299-309.

38 Seelbach, K., vanDeurzen, M. P. J., van Rantwijk, F., Sheldon, R. A. & Kragl, U. Improvement of the total turnover number and space-time yield for chloroperoxidase catalyzed oxidation. *Biotechnol. Bioeng.* (1997), **55**, 283-288.

39 Thiel, D., Doknic, D. & Deska, J. Enzymatic aerobic ring rearrangement of optically active furylcarbinols. *Nat. Commun.* (2014), **5**, 7.

40 Nelson, D. R. Cytochrome P450 diversity in the tree of life. *BBA-Proteins Proteomics* (2018), **1866**, 141-154.

41 Hanson, L. K., Sligar, S. G. & Gunsalus, I. C. Electronic-structure of cytochrome-p450. *Croat. Chem. Acta* (1977), **49**, 237-250.

42 Urlacher, V. B. & Girhard, M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol.* (2012), **30**, 26-36.

43 Fasan, R. Tuning P450 Enzymes as Oxidation Catalysts. *ACS Catal.* (2012), **2**, 647-666.

44 Munro, A. W., Girvan, H. M. & McLean, K. J. Variations on a (t)heme - novel mechanisms, redox partners and catalytic functions in the cytochrome P450 superfamily. *Nat. Prod. Rep.* (2007), **24**, 585-609.

45 O'Reilly, E., Kohler, V., Flitsch, S. L. & Turner, N. J. Cytochromes P450 as useful biocatalysts: addressing the limitations. *Chem. Commun.* (2011), **47**, 2490-2501.

46 McLean, K. J., Luciakova, D., Belcher, J., Tee, K. L. & Munro, A. W. in *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450* Vol. 851 *Advances in Experimental Medicine and Biology* (eds E. G. Hrycay & S. M. Bandiera) 299-317 (Springer-Verlag Berlin, 2015).

47 Tsai, R., Yu, C. A., Gunsalus, I. C., Peisach, J., Blumberg, W., Ormejohn.Wh & Beinert, H. Spin-state changes in cytochrome-p-450cam on binding of specific substrates. *Proc. Natl. Acad. Sci. U. S. A.* (1970), **66**, 1157-&.

48 Sligar, S. G. coupling of spin, substrate, and redox equilibria in cytochrome P450. *Biochemistry* (1976), **15**, 5399-5406.

49 Daff, S. N., Chapman, S. K., Turner, K. L., Holt, R. A., Govindaraj, S., Poulos, T. L. & Munro, A. W. Redox control of the catalytic cycle of flavocytochrome P-450 BM3. *Biochemistry* (1997), **36**, 13816-13823.

50 Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* (2005), **105**, 2253-2277.

51 Egawa, T., Shimada, H. & Ishimura, Y. Evidence for compound I formation in the reaction of cytochrome-P450cam with *m*-chloroperbenzoic acid. *Biochem. Biophys. Res. Commun.* (1994), **201**, 1464-1469.

52 Newcomb, M., Zhang, R., Chandrasena, R. E. P., Halgrimson, J. A., Horner, J. H., Makris, T. M. & Sligar, S. G. Cytochrome P450 Compound I. *J. Am. Chem. Soc.* (2006), **128**, 4580-4581.

53 Tajima, K., Edo, T., Ishizu, K., Imaoka, S., Funae, Y., Oka, S. & Sakurai, H. Cytochrome P-450-Butyl Peroxide Complex detected by ESR. (1993), **191**, 157-164.

54 Watanabe, I., Nara, F. & Serizawa, N. Cloning, characterization and expression of the gene encoding cytochrome P-450(SCA-2) from *streptomyces-carbophilus* involved in production of pravastatin, a specific hmg-coa reductase inhibitor. *Gene* (1995), **163**, 81-85.

55 Suzuki, K., Sanga, K., Chikaoka, Y. & Itagaki, E. Purification and properties of cytochrome-P-450 (p-450(lun)) catalyzing steroid 11-beta-hydroxylation in *curvularia-lunata*. (1993), **1203**, 215-223.

56 Sonomoto, K., Hoq, M. M., Tanaka, A. & Sukui, S. 11-beta-hydroxylation of cortexolone (reichstein compound-s) to hydrocortisone by *curvularia-lunata* entrapped in photo-cross-linked resin gels. *Appl. Environ. Microbiol.* (1983), **45**, 436-443.

57 Sowden, R. J., Yasmin, S., Rees, N. H., Bell, S. G. & Wong, L. L. Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450(cam) and P450(BM-3). *Org. Biomol. Chem.* (2005), **3**, 57-64.

58 Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R. & Keasling, J. D. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* (2006), **440**, 940-943.

59 Ryan, J. D., Fish, R. H. & Clark, D. S. Engineering Cytochrome P450 Enzymes for Improved Activity towards Biomimetic 1,4-NADH Cofactors. *ChemBioChem* (2008), **9**, 2579-2582.

60 van der Donk, W. A. & Zhao, H. M. Recent developments in pyridine nucleotide regeneration. (2003), **14**, 421-426.

61 Hollmann, F., Witholt, B. & Schmid, A. Cp*Rh(bpy)(H2O) (2+): a versatile tool for efficient and non-enzymatic regeneration of nicotinamide and flavin coenzymes. *J. Mol. Catal. B-Enzym.* (2002), **19**, 167-176.

62 Narhi, L. O. & Fulco, A. J. characterization of a catalytically self-sufficient 119,000-dalton cytochrome-P-450 monooxygenase induced by barbiturates in *bacillus-megaterium*. *J. Biol. Chem.* (1986), **261**, 7160-7169.

63 Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L. & Turner, N. J. Identification of a new class of cytochrome P450 from a Rhodococcus sp. *J. Bacteriol.* (2002), **184**, 3898-3908.

64 Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.* (2006), **124**, 128-145.

65 Cirino, P. C. & Arnold, F. H. A self-sufficient peroxide-driven hydroxylation biocatalyst. *Angew. Chem.-Int. Edit.* (2003), **42**, 3299-3301.

66 Glieder, A., Farinas, E. T. & Arnold, F. H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat. Biotechnol.* (2002), **20**, 1135-1139.

67 Chen, J., Kong, F. H., Ma, N. N., Zhao, P. X., Liu, C. F., Wang, X. L. & Cong, Z. Q. Peroxide-Driven Hydroxylation of Small Alkanes Catalyzed by an Artificial P450BM3 Peroxygenase System. *ACS Catal.* (2019), **9**, 7350-7355.

68 Nguyen, T. H. H., Yeom, S. J. & Yun, C. H. Production of a Human Metabolite of Atorvastatin by Bacterial CYP102A1 Peroxygenase. *Appl. Sci.-Basel* (2021), **11**, 13.

69 Jiang, Y. H., Wang, C. L., Ma, N. N., Chen, J., Liu, C. F., Wang, F., Xu, J. K. & Cong, Z. Q. Regioselective aromatic O-demethylation with an artificial P450BM3 peroxygenase system. *Catal. Sci. Technol.* (2020), **10**, 1219-1223.

70 Ma, N. N., Chen, Z. F., Chen, J., Chen, J. F., Wang, C., Zhou, H. F., Yao, L. S., Shoji, O., Watanabe, Y. & Cong, Z. Q. Dual-Functional Small Molecules for Generating an Efficient Cytochrome P450BM3 Peroxygenase. *Angew. Chem.-Int. Edit.* (2018), **57**, 7628-7633.

71 Podgorski, M. N., Harbort, J. S., Lee, J. H. Z., Nguyen, G. T. H., Bruning, J. B., Donald, W. A., Bernhardt, P. V., Harmer, J. R. & Bell, S. G. An Altered Heme Environment in an Engineered Cytochrome P450 Enzyme Enables the Switch from Monooxygenase to Peroxygenase Activity. *ACS Catal.* (2022), **12**, 1614-1625.

72 Matsunaga, I., Sumimoto, T., Ueda, A., Kusunose, E. & Ichihara, K. Fatty acid-specific, regiospecific, and stereospecific hydroxylation by cytochrome P450 (CYP152B1) from Sphingomonas paucimobilis: Substrate structure required for alpha-hydroxylation. *Lipids* (2000), **35**, 365-371.

73 Matsunaga, I., Ueda, A., Sumimoto, T., Ichihara, K., Ayata, M. & Ogura, H. Site-directed mutagenesis of the putative distal helix of peroxygenase cytochrome P450. *Arch. Biochem. Biophys.* (2001), **394**, 45-53.

74 Girhard, M., Schuster, S., Dietrich, M., Durre, P. & Urlacher, V. B. Cytochrome P450 monooxygenase from Clostridium acetobutylicum: A new alpha-fatty acid hydroxylase. *Biochem. Biophys. Res. Commun.* (2007), **362**, 114-119.

75 Liu, Y., Wang, C., Yan, J. Y., Zhang, W., Guan, W. N., Lu, X. F. & Li, S. Y. Hydrogen peroxide-independent production of alpha-alkenes by OleT(JE) P450 fatty acid decarboxylase. *Biotechnol. Biofuels* (2014), **7**, 12.

76 Munro, A. W., McLean, K. J., Grant, J. L. & Makris, T. M. Structure and function of the cytochrome P450 peroxygenase enzymes. *Biochem. Soc. Trans.* (2018), **46**, 183-196.

77 Ullrich, R., Nüske , J., Scheibner, K., Spantzel, J. & Hofrichter, M. Novel Haloperoxidase from the Agaric Basidiomycete *Agrocybe aegrita* Oxidises Aryl Alcohols and Aldehydes. (2004), **70**, 4575 - 4581.

78 Hobisch, M., Holtmann, D., de Santos, P. G., Alcalde, M., Hollmann, F. & Kara, S. Recent developments in the use of peroxygenases - Exploring their high potential in selective oxyfunctionalisations. *Biotechnol. Adv.* (2021), **51**, 13.

79 Ullrich, R. & Hofrichter, M. The haloperoxidase of the agaric fungus *Agrocybe aegerita* hydroxylates toluene and napthalene. (2005), **579**, 6247 - 6250.

80 Hofrichter, M., Kellner, H., Pecyna, M. J. & Ullrich, R. in *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450* Vol. 851 *Advances in Experimental Medicine and Biology* (eds E. G. Hrycay & S. M. Bandiera) 341-368 (Springer-Verlag Berlin, 2015).

81 Hofrichter, M. & Ullrich, R. Oxidations catalyzed by fungal peroxygenases. *Curr. Opin. Chem. Biol.* (2014), **19**, 116-125.

82 Wang, X., Peter, S., Kinne, M., Hofrichter, M. & Groves, J. T. Detection and Kinetic Characterisation of a Highly Reactive Heme-Thiolate Peroxygenase Compound I. (2012), **134**, 12897 - 12900.

83 Wang, X. S., Ullrich, R., Hofrichter, M. & Groves, J. T. Heme-thiolate ferryl of aromatic peroxygenase is basic and reactive. *Proc. Natl. Acad. Sci. U. S. A.* (2015), **112**, 3686-3691.

84 Faiza, M., Huang, S. F., Lan, D. M. & Wang, Y. H. New insights on unspecific peroxygenases: superfamily reclassification and evolution. *BMC Evol. Biol.* (2019), **19**, 19.

85 Kinner, A., Rosenthal, K. & Lutz, S. Identification and Expression of New Unspecific Peroxygenases - Recent Advances, Challenges and Opportunities. *Front. Bioeng. Biotechnol.* (2021), **9**, 17.

86 Faiza, M., Lan, D. M., Huang, S. F. & Wang, Y. H. UPObase: an online database of unspecific peroxygenases. *Database* (2019), 15.

87 Anh, D. H. Novel extracellular haloperoxidaseperoxygenases from the coprophilous fungi *Coprinus radians* and *Coprinus verticillatus*: production, purification and biochemical characterization, International Graduate School of Zittau, (2008).

88 Beltrán-Nogal, A., Sánchez-Moreno, I., Méndez-Sánchez, D., Gómez de Santos, P., Hollman, F. & Alcade, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. (2022), **73**.

89 Piontek, K., Strittmatter, E., Ullrich, R., Grobe, G., Pecyna, M. J., Kluge, M., Scheibner, K., Hofrichter, M. & Plattner, D. A. Structural Basis of Substrate Conversion in a New Aromatic Peroxygenase - Cytochrome P450 functionality with benefits. *J. Biol. Chem.* (2013), **288**, 34767-34776.

90 Rotilio, L., Swoboda, A., Ebner, K., Rinnofner, C., Glieder, A., Kroutil, W. & Mattevi, A. Structural and Biochemical Studies Enlighten the Unspecific

Peroxygenase from Hypoxylon sp. EC38 as an Efficient Oxidative Biocatalyst. *ACS Catal.* (2021), **11**, 11511-11525.

91 Ramirez-Escudero, M., Molina-Espeja, P., de Santos, P. G., Hofrichter, M., Sanz-Aparicio, J. & Alcalde, M. Structural Insights into the Substrate Promiscuity of a Laboratory-Evolved Peroxygenase. *ACS Chem. Biol.* (2018), **13**, 3259-3268.

92 Ramirez-Ramirez, J., Martin-Diaz, J., Pastor, N., Alcalde, M. & Ayala, M. Exploring the Role of Phenylalanine Residues in Modulating the Flexibility and Topography of the Active Site in the Peroxygenase Variant PaDa-I. *Int. J. Mol. Sci.* (2020), **21**, 15.

93 Olmedo, A., Aranda, C., del Rio, J. C., Kiebist, J., Scheibner, K., Martinez, A. T. & Gutierrez, A. From Alkanes to Carboxylic Acids: Terminal Oxygenation by a Fungal Peroxygenase. *Angew. Chem.-Int. Edit.* (2016), **55**, 12248-12251.

94 Pourmir, A. & Johannes, T. W. Directed Evolution: Selection of the Host Organism. (2012), **2**.

95 Mielke, T. Heterologous Expression and Characterisation of Unspecific Peroxygenases Ph.D. thesis, The University of York, (2017).

96 Molina-Espeja, P., Garcia-Ruiz, E., Gonzalez-Perez, D., Ullrich, R., Hofrichter, M. & Alcalde, M. Directed Evolution of Unspecific Peroxygenase from Agrocybe aegerita. *Appl. Environ. Microbiol.* (2014), **80**, 3496-3507.

97 Gonzalez-Perez, D., Molina-Espeja, P., Garcia-Ruiz, E. & Alcalde, M. Mutagenic Organized Recombination Process by Homologous In Vivo Grouping (MORPHING) for Directed Enzyme Evolution. *PLoS One* (2014), **9**, 12.

98 Molina-Espeja, P., Ma, S., Mate, D. M., Ludwig, R. & Alcalde, M. Tandem-yeast expression system for engineering and producing unspecific peroxygenase. *Enzyme Microb. Technol.* (2015), **73-74**, 29-33.

99 Knorrscheidt, A., Soler, J., Hunecke, N., Pullmann, P., Garcia-Borras, M. & Weissenborn, M. J. Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase. *ACS Catal.* (2021), **11**, 7327-7338.

100 Pullmann, P., Knorrscheidt, A., Munch, J., Palme, P. R., Hoehenwarter, W., Marillonnet, S., Alcalde, M., Westermann, B. & Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxygenases. *Commun. Biol.* (2021), **4**, 20.

101 Molina-Espeja, P., Canellas, M., Plou, F. J., Hofrichter, M., Lucas, F., Guallar, V. & Alcalde, M. Synthesis of 1-Naphthol by a Natural Peroxygenase Engineered by Directed Evolution. *ChemBioChem* (2016), **17**, 341-349.

102 de Santos, P. G., Canellas, M., Tieves, F., Younes, S. H. H., Molina-Espeja, P., Hofrichter, M., Hollmann, F., Guallar, V. & Alcalde, M. Selective Synthesis of the Human Drug Metabolite 5 '-Hydroxypropranolol by an Evolved Self-Sufficient Peroxygenase. *ACS Catal.* (2018), **8**, 4789-4799.

103 Jacob, S., Bormann, S., Becker, M., Antelo, L., Holtmann, D. & Thines, E. Magnaporthe oryzae as an expression host for the production of the unspecific peroxygenase AaeUPO from the basidiomycete Agrocybe aegerita. *MicrobiologyOpen* (2021), **10**, 6.

104 Pecyna, M. J., Ullrich, R., Bittner, B., Clemens, A., Scheibner, K., Schubert, R. & Hofrichter, M. Molecular characterization of aromatic peroxygenase from Agrocybe aegerita. *Appl. Microbiol. Biotechnol.* (2009), **84**, 885-897.

105 Ullrich, R., Poraj-Kobielska, M., Scholze, S., Halbout, C., Sandvoss, M., Pecyna, M. J., Scheibner, K. & Hofrichter, M. Side chain removal from corticosteroids by unspecific peroxygenase. *J. Inorg. Biochem.* (2018), **183**, 84-93.

106 Kiebist, J., Schmidtke, K. U., Zimmermann, J., Kellner, H., Jehmlich, N., Ullrich, R., Zander, D., Hofrichter, M. & Scheibner, K. A Peroxygenase from Chaetomium globosum Catalyzes the Selective Oxygenation of Testosterone. *ChemBioChem* (2017), **18**, 563-569.

107 Babot, E. D., del Rio, J. C., Kalum, L., Martinez, A. T. & Gutierrez, A. Oxyfunctionalization of Aliphatic Compounds by a Recombinant Peroxygenase From Coprinopsis cinerea. *Biotechnol. Bioeng.* (2013), **110**, 2323-2332.

108 Carro, J., Gonzalez-Benjumea, A., Fernandez-Fueyo, E., Aranda, C., Guallar, V., Gutierrez, A. & Martinez, A. T. Modulating Fatty Acid Epoxidation vs Hydroxylation in a Fungal Peroxygenase. *ACS Catal.* (2019), **9**, 6234-6242.

109 Fernandez, E., Aranda, C., Gutierrez, A. & Martinez, A. Method of Heterologous Expression of Active Fungal Unspecific Peroxygenase in Bacterial Host Cells for Fatty-Acid Epoxidation and other Oxygenation Reaction. (2020).

110 Linde, D., Olmedo, A., Gonzalez-Benjumea, A., Estevez, M., Renau-Minguez, C., Carro, J., Fernandez-Fueyo, E., Gutierrez, A. & Martinez, A. T. Two New Unspecific Peroxygenases from Heterologous Expression of Fungal Genes in Escherichia coli. (2020), **86**.

111 Mate, D. M., Palomino, M. A., Molina-Espeja, P., Martin-Diaz, J. & Alcalde, M. Modification of the peroxygenative: peroxidative activity ratio in the unspecific peroxygenase from Agrocybe aegerita by structure-guided evolution. *Protein Eng. Des. Sel.* (2017), **30**, 191-198.

112 de Santos, P. G., Cervantes, F. V., Tieves, F., Plou, F. J., Hollmann, F. & Alcalde, M. Benchmarking of laboratory evolved unspecific peroxygenases for the synthesis of human drug metabolites. *Tetrahedron* (2019), **75**, 1827-1831.

113 Ullrich, R. & Hofrichter, M. The haloperoxidase of the agaric fungus Agrocybe aegerita hydroxylates toluene and naphthalene. *FEBS Lett.* (2005), **579**, 6247-6250.

114 Wang, X. S., Peter, S., Ullrich, R., Hofrichter, M. & Groves, J. T. Driving Force for Oxygen-Atom Transfer by Heme-Thiolate Enzymes. *Angew. Chem.-Int. Edit.* (2013), **52**, 9238-9241.

115 Anh, D. H., Ullrich, R., Benndorf, D., Svatos, A., Muck, A. & Hofrichter, M. The coprophilous mushroom Coprinus radians secretes a haloperoxidase that catalyzes aromatic peroxygenation. *Appl. Environ. Microbiol.* (2007), **73**, 5477-5485.

116 Kluge, M., Ullrich, R., Dolge, C., Scheibner, K. & Hofrichter, M. Hydroxylation of naphthalene by aromatic peroxygenase from Agrocybe aegerita proceeds via oxygen transfer from H2O2 and intermediary epoxidation. *Appl. Microbiol. Biotechnol.* (2009), **81**, 1071-1076.

117 Barkova, K., Kinne, M., Ullrich, R., Hennig, L., Fuchs, A. & Hofrichter, M. Regioselective hydroxylation of diverse flavonoids by an aromatic peroxygenase. *Tetrahedron* (2011), **67**, 4874-4878.

118 Kinne, M., Ullrich, R., Hammel, K. E., Scheibner, K. & Hofrichter, M. Regioselective preparation of (R)-2-(4-hydroxyphenoxy)propionic acid with a fungal peroxygenase. *Tetrahedron Lett.* (2008), **49**, 5950-5953.

119 Kinne, M., Poraj-Kobielska, M., Aranda, E., Ullrich, R., Hammel, K. E., Scheibner, K. & Hofrichter, M. Regioselective preparation of 5-

hydroxypropranolol and 4 '-hydroxydiclofenac with a fungal peroxygenase. *Bioorg. Med. Chem. Lett.* (2009), **19**, 3085-3087.

120 Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K., Kayser, G., Hammel, K. E. & Hofrichter, M. Preparation of human drug metabolites using fungal peroxygenases. *Biochem. Pharmacol.* (2011), **82**, 789-796.

121 Kiebist, J., Schmidtke, K. U., Schramm, M., Konig, R., Quint, S., Kohlmann, J., Zuhse, R., Ullrich, R., Hofrichter, M. & Scheibner, K. Biocatalytic Syntheses of Antiplatelet Metabolites of the Thienopyridines Clopidogrel and Prasugrel Using Fungal Peroxygenases. *J. Fungi* (2021), **7**, 17.

122 Karich, A., Ullrich, R., Scheibner, K. & Hofrichter, M. Fungal Unspecific Peroxygenases Oxidize the Majority of Organic EPA Priority Pollutants. *Front. Microbiol.* (2017), **8**, 15.

123 Karich, A., Kluge, M., Ullrich, R. & Hofrichter, M. Benzene oxygenation and oxidation by the peroxygenase of Agrocybe aegerita. *AMB Express* (2013), **3**, 8.

124 Peter, S., Kinne, M., Wang, X. S., Ullrich, R., Kayser, G., Groves, J. T. & Hofrichter, M. Selective hydroxylation of alkanes by an extracellular fungal peroxygenase. *FEBS J.* (2011), **278**, 3667-3675.

125 Kiebist, J., Holla, W., Heidrich, J., Poraj-Kobielska, M., Sandvoss, M., Simonis, R., Grobe, G., Atzrodt, J., Hofrichter, M. & Scheibner, K. One-pot synthesis of human metabolites of SAR548304 by fungal peroxygenases. *Bioorg. Med. Chem.* (2015), **23**, 4324-4332.

126 Gutierrez, A., Babot, E. D., Ullrich, R., Hofrichter, M., Martinez, A. T. & del Rio, J. C. Regioselective oxygenation of fatty acids, fatty alcohols and other aliphatic

compounds by a basidiomycete heme-thiolate peroxidase. *Arch. Biochem. Biophys.* (2011), **514**, 33-43.

127 Kluge, M., Ullrich, R., Scheibner, K. & Hofrichter, M. Stereoselective benzylic hydroxylation of alkylbenzenes and epoxidation of styrene derivatives catalyzed by the peroxygenase of Agrocybe aegerita. *Green Chem.* (2012), **14**, 440-446.

128 Olmedo, A., del Rio, J. C., Kiebist, J., Ullrich, R., Hofrichter, M., Scheibner, K., Martinez, A. T. & Gutierrez, A. Fatty Acid Chain Shortening by a Fungal Peroxygenase. *Chem.-Eur. J.* (2017), **23**, 16985-16989.

129 Babot, E. D., del Rio, J. C., Canellas, M., Sancho, F., Lucas, F., Guallar, V., Kalum, L., Lund, H., Grobe, G., Scheibner, K., Ullrich, R., Hofrichter, M., Martinez, A. T. & Gutiereza, A. Steroid Hydroxylation by Basidiomycete Peroxygenases: a Combined Experimental and Computational Study. *Appl. Environ. Microbiol.* (2015), **81**, 4130-4142.

130 Lucas, F., Babot, E. D., Canellas, M., del Rio, J. C., Kalum, L., Ullrich, R., Hofrichter, M., Guallar, V., Martinez, A. T. & Gutierrez, A. Molecular determinants for selective C-25-hydroxylation of vitamins D-2 and D-3 by fungal peroxygenases. *Catal. Sci. Technol.* (2016), **6**, 288-295.

131 Babot, E. D., del Rio, J. C., Kalum, L., Martinez, A. T. & Gutierrez, A. Regioselective Hydroxylation in the Production of 25-Hydroxyvitamin D by Coprinopsis cinerea Peroxygenase. *ChemCatChem* (2015), **7**, 283-290.

132 Peter, S., Kinne, M., Ullrich, R., Kayser, G. & Hofrichter, M. Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase. *Enzyme Microb. Technol.* (2013), **52**, 370-376.

133 Zhang, W. Y., Li, H. H., Younes, S. H. H., de Santos, P. G., Tieves, F., Grogan, G., Pabst, M., Alcalde, M., Whitwood, A. C. & Hollmann, F. Biocatalytic Aromaticity-Breaking Epoxidation of Naphthalene and Nucleophilic Ring-Opening Reactions. *ACS Catal.* (2021), **11**, 2644-2649.

134 Aranda, C., Olmedo, A., Kiebist, J., Scheibner, K., del Rio, J. C., Martinez, A. T. & Gutierrez, A. Selective Epoxidation of Fatty Acids and Fatty Acid Methyl Esters by Fungal Peroxygenases. *ChemCatChem* (2018), **10**, 3964-3968.

135 Gonzalez-Benjumea, A., Carro, J., Renau-Minguez, C., Linde, D., Fernandez-Fueyo, E., Gutierrez, A. & Martinez, A. T. Fatty acid epoxidation by Collariella virescens peroxygenase and heme-channel variants. (2020), **10**, 717-725.

136 Gonzalez-Benjumea, A., Marques, G., Herold-Majumdar, O. M., Kiebist, J., Scheibner, K., del Rio, J. C., Martinez, A. T. & Gutierrez, A. High Epoxidation Yields of Vegetable Oil Hydrolyzates and Methyl Esters by Selected Fungal Peroxygenases. *Front. Bioeng. Biotechnol.* (2021), **8**, 12.

137 Aranda, E., Kinne, M., Kluge, M., Ullrich, R. & Hofrichter, M. Conversion of dibenzothiophene by the mushrooms Agrocybe aegerita and Coprinellus radians and their extracellular peroxygenases. *Appl. Microbiol. Biotechnol.* (2009), **82**, 1057-1066.

138 Bassanini, I., Ferrandi, E. E., Vanoni, M., Ottolina, G., Riva, S., Crotti, M., Brenna, E. & Monti, D. Peroxygenase-Catalyzed Enantioselective Sulfoxidations. *Eur. J. Org. Chem.* (2017), 7186-7189.

139 Kinne, M., Poraj-Kobielska, M., Ralph, S. A., Ullrich, R., Hofrichter, M. & Hammel, K. E. Oxidative Cleavage of Diverse Ethers by an Extracellular Fungal Peroxygenase. *J. Biol. Chem.* (2009), **284**, 29343-29349.

140 Mireles, R., Ramirez-Ramirez, J., Alcalde, M. & Ayala, M. Ether Oxidation by an Evolved Fungal Heme-Peroxygenase: Insights into Substrate Recognition and Reactivity. *J. Fungi* (2021), **7**, 10.

141 Valderrama, B., Ayala, M. & Vazquez-Duhalt, R. Suicide inactivation of peroxidases and the challenge of engineering more robust enzymes. *Chem. Biol.* (2002), **9**, 555-565.

142 Alexander, K., Katrin, C., Rene, U. & Martin, H. Exploring the catalase activity of unspecific peroxygenases and the mechanism of peroxide-dependent heme destruction. *J. Mol. Catal. B-Enzym.* (2016), **134**, 238-246.

143 Pesic, M., Willot, S. J. P., Fernandez-Fueyo, E., Tieves, F., Alcalde, M. & Hollmann, F. Multienzymatic in situ hydrogen peroxide generation cascade for peroxygenase-catalysed oxyfunctionalisation reactions. (2019), **74**, 100-103.

144 Ni, Y., Fernandez-Fueyo, E., Baraibar, A. G., Ullrich, R., Hofrichter, M., Yanase, H., Alcalde, M., van Berkel, W. J. H. & Hollmann, F. Peroxygenase-Catalyzed Oxyfunctionalization Reactions Promoted by the Complete Oxidation of Methanol. *Angew. Chem.-Int. Edit.* (2016), **55**, 798-801.

145 Zhang, W. Y., Burek, B. O., Fernandez-Fueyo, E., Alcalde, M., Bloh, J. Z. & Hollmann, F. Selective Activation of C-H Bonds in a Cascade Process Combining Photochemistry and Biocatalysis. *Angew. Chem.-Int. Edit.* (2017), **56**, 15451-15455.

146 Zhang, W. Y., Fernandez-Fueyo, E., Ni, Y., van Schie, M., Gacs, J., Renirie, R., Wever, R., Mutti, F. G., Rother, D., Alcalde, M. & Hollmann, F. Selective aerobic oxidation reactions using a combination of photocatalytic water oxidation and enzymatic oxyfunctionalizations. *Nat. Catal.* (2018), **1**, 55-62.

147 Yayci, A., Baraibar, A. G., Krewing, M., Fueyo, E. F., Hollmann, F., Alcalde, M., Kourist, R. & Bandow, J. E. Plasma-Driven in Situ Production of Hydrogen Peroxide for Biocatalysis. *ChemSusChem* (2020), **13**, 2072-2079.

148 de Santos, P. G., Lazaro, S., Vina-Gonzalez, J., Hoang, M. D., Sanchez-Moreno, I., Glieder, A., Hollmann, F. & Alcalde, M. Evolved Peroxygenase-Aryl Alcohol Oxidase Fusions for Self-Sufficient Oxyfunctionalization Reactions. *ACS Catal.* (2020), **10**, 13524-13534.

149 Vina-Gonzalez, J., Gonzalez-Perez, D., Ferreira, P., Martinez, A. T. & Alcalde, M. Focused Directed Evolution of Aryl-Alcohol Oxidase in Saccharomyces cerevisiae by Using Chimeric Signal Peptides. *Appl. Environ. Microbiol.* (2015), **81**, 6451-6462.

150 Stepankova, V., Bidmanova, S., Koudelakova, T., Prokop, Z., Chaloupkova, R. & Damborsky, J. Strategies for Stabilization of Enzymes in Organic Solvents. *ACS Catal.* (2013), **3**, 2823-2836.

151 Klibanov, A. M. Improving enzymes by using them in organic solvents. *Nature* (2001), **409**, 241-246.

152 Fernandez-Fueyo, E., Ni, Y., Baraibar, A. G., Alcalde, M., van Langen, L. M. & Hollmann, F. Towards preparative peroxygenase-catalyzed oxyfunctionalization reactions in organic media. *J. Mol. Catal. B-Enzym.* (2016), **134**, 347-352.

153 Martin-Diaz, J., Molina-Espeja, P., Hofrichter, M., Hollmann, F. & Alcalde, M. Directed evolution of unspecific peroxygenase in organic solvents. *Biotechnol. Bioeng.* (2021), **118**, 3002-3014.

154 Vind, J., Kiemer, L. & Amourgi, E. Polypeptides having Peroxygenase Activity. (2016).

155 Invitrogen. Vol. Cat. no. K1740-01, MAN0000042 (ed Invitrogen) (Life Technologies, 2010).

156 Dubendorff, J. W. & Studier, F. W. Controlling basal expression in an inducible t7 expression system by blocking the target t7 promoter with lac repressor. *J. Mol. Biol.* (1991), **219**, 45-59.

157 Gillam, E. M. J., Guo, Z. Y., Martin, M. V., Jenkins, C. M. & Guengerich, F. P. Expression of cytochrome-P450 2d6 in *escherichia-coli*, purification, and spectral and catalytic characterization. *Arch. Biochem. Biophys.* (1995), **319**, 540-550.

158 Grobe, G., Ullrich, R., Pecyna, M. J., Kapturska, D., Friedrich, S., Hofrichter, M. & Scheibner, K. High-yield production of aromatic peroxygenase by the agaric fungus *Marasmius rotula*. *AMB Express* (2011), **1**, 11.

159 Gouterman, M. Spectra of porphyrins. *J. Mol. Spectrosc.* (1961), **6**, 138.

160 Rittle, J. & Green, M. T. Cytochrome P450 Compound I: Capture, Characterization, and C-H Bond Activation Kinetics. *Science* (2010), **330**, 933-937.

161 Egawa, T., Proshlyakov, D. A., Miki, H., Makino, R., Ogura, T., Kitagawa, T. & Ishimura, Y. Effects of a thiolate axial ligand on the pi ->pi* electronic states of oxoferryl porphyrins: a study of the optical and resonance Raman spectra of compounds I and II of chloroperoxidase. *J. Biol. Inorg. Chem.* (2001), **6**, 46-54.

162 Liu, W. C., Inwood, S., Gong, T., Sharma, A., Yu, L. Y. & Zhu, P. Fed-batch high-cell-density fermentation strategies for Pichia pastoris growth and production. *Crit. Rev. Biotechnol.* (2019), **39**, 258-271.

163 Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* (2010), **43**, 186-190.

164 Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D-Struct. Biol.* (2006), **62**, 72-82.

165 Kabsch, W. XDS. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* (2010), **66**, 125-132.

166 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D-Struct. Biol.* (1997), **53**, 240-255.

167 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D-Struct. Biol.* (2004), **60**, 2126-2132.

168 Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. Sect. D-Struct. Biol.* (2003), **59**, 1131-1137.

169 Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P. R., Jenkins, H. T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S. J., Nicholls, R. A., Noble, M., Pannu, N. S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., Uski, V., von Delft, F., Waterman, D., Wilson, K., Winn, M. & Wojdyr, M. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. Sect. D-Struct. Biol.* (2018), **74**, 68-84.

170 Berka, K., Hanak, O., Sehnal, D., Banas, P., Navratilova, V., Jaiswal, D., Ionescu, C. M., Varekova, R. S., Koca, J. & Otyepka, M. MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. *Nucleic Acids Res.* (2012), **40**, W222-W227.

171 Mirdita, M., Sc hütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M. ColabFold - Making protein folding accessible to all. (2022).

172 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* (2021), **596**, 583.

173 Steinegger, M. & Soding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* (2017), **35**, 1026-1028.

174 Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronnerberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvanakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J. & Hassabis, D. Protein complex prediction with AlphaFold-Multimer. (2022).

175 Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K. & Hofrichter, M. A spectrophotometric assay for the detection of fungal peroxygenases. *Anal. Biochem.* (2012), **421**, 327-329.

176 Molina-Espeja, P., Beltran-Nogal, A., Alfuzzi, M. A., Guallar, V. & Alcalde, M. Mapping Potential Determinants of Peroxidative Activity in an Evolved Fungal Peroxygenase from Agrocybe aegerita. *Front. Bioeng. Biotechnol.* (2021), **9**, 11.

177 Magnusson, A. O., Szekrenyi, A., Joosten, H. J., Finnigan, J., Charnock, S. & Fessner, W. D. nanoDSF as screening tool for enzyme libraries and biotechnology development. *FEBS J.* (2019), **286**, 184-204.

178 Polizzi, K. M., Bommarius, A. S., Broering, J. M. & Chaparro-Riggers, J. F. Stability of biocatalysts. *Curr. Opin. Chem. Biol.* (2007), **11**, 220-225.

179 Alexander, C. G., Wanner, R., Johnson, C. M., Breitsprecher, D., Winter, G., Duhr, S., Baaske, P. & Ferguson, N. Novel microscale approaches for easy, rapid determination of protein stability in academic and commercial settings. *BBA-Proteins Proteomics* (2014), **1844**, 2241-2250.

180 McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr. F-Struct. Biol. Commun.* (2014), **70**, 2-20.

181 Adman, E., Watenpaugh, K. D. & Jensen, L. H. NH...S hydrogen-bonds in *peptococcus-aerogenes* ferredoxin, *clostridium-pasteurianum* rubredoxin, and chromatium high potential iron protein. *Proc. Natl. Acad. Sci. U. S. A.* (1975), **72**, 4854-4858.

182 Piontek, K., Ullrich, R., Liers, C., Diederichs, K., Plattner, D. A. & Hofrichter, M. Crystallization of a 45 kDa peroxygenase/peroxidase from the mushroom *Agrocybe aegerita* and structure determination by SAD utilizing only the haem iron. *Acta Crystallogr. F-Struct. Biol. Commun.* (2010), **66**, 693-698.

183 Lowenthal, M. S., Davis, K. S., Formolo, T., Kilpatrick, L. E. & Phinney, K. W. Identification of Novel N-Glycosylation Sites at Noncanonical Protein Consensus Motifs. *J. Proteome Res.* (2016), **15**, 2087-2101.

184 Tsai, C. J., Lin, S. L., Wolfson, H. J. & Nussinov, R. Studies of protein-protein interfaces: A statistical analysis of the hydrophobic effect. *Protein Sci.* (1997), **6**, 53-64.

185 Zhou, H. Y. & Zhou, Y. Q. Quantifying the effect of burial of amino acid residues on protein stability. *Proteins* (2004), **54**, 315-322.

186 Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D. V., Stroe, O., Wood, G., Laydon, A., Aidek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D. & Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* (2022), **50**, D439-D444.

187 Sigmund, M. C. & Poelarends, G. J. Current state and future perspectives of engineered and artificial peroxygenases for the oxyfunctionalization of organic molecules. *Nat. Catal.* (2020), **3**, 690-702.

188 Rauch, M. C. R., Tieves, F., Paul, C. E., Arends, I., Alcalde, M. & Hollmann, F. Peroxygenase-Catalysed Epoxidation of Styrene Derivatives in Neat Reaction Media. *ChemCatChem* (2019), **11**, 4519-4523.

189 Summers, B. Towards the Industrial Application of the Baeyer-Villiger Monooxygenase MO14 from *Rhodococcus jostii* Ph.D. thesis, The University of York, (2014).

190 Jensen, C. N., Cartwright, J., Ward, J., Hart, S., Turkenburg, J. P., Ali, S. T., Allen, M. J. & Grogan, G. A Flavoprotein Monooxygenase that Catalyses a Baeyer-Villiger Reaction and Thioether Oxidation Using NADH as the Nicotinamide Cofactor. *ChemBioChem* (2012), **13**, 872-878.

191 Xuan, Q. Q., Zhao, C. & Song, Q. L. Umpolung of protons from H_2O: a metal-free chemoselective reduction of carbonyl compounds via B(2)pin(2)/H_2O systems. *Org. Biomol. Chem.* (2017), **15**, 5140-5144.

192 Betori, R. C., May, C. M. & Scheidt, K. A. Combined Photoredox/Enzymatic C-H Benzylic Hydroxylations. *Angew. Chem.-Int. Edit.* (2019), **58**, 16490-16494.

193 Luo, F., Pan, C. D., Li, L. P., Chen, F. & Cheng, J. Copper-mediated methylthiolation of aryl halides with DMSO. *Chem. Commun.* (2011), **47**, 5304-5306.

194 Peter, S., Karich, A., Ullrich, R., Grobe, G., Scheibner, K. & Hofrichter, M. Enzymatic one-pot conversion of cyclohexane into cyclohexanone: Comparison of four fungal peroxygenases. *J. Mol. Catal. B-Enzym.* (2014), **103**, 47-51.

195 Trott, O. & Olson, A. J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* (2010), **31**, 455-461.

196 Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem Inf. Model.* (2021), **61**, 3891-3898.

197 Hansch, C., Leo, A. & Taft, R. W. A survey of hammett substituent constants and resonance and field parameters. *Chem. Rev.* (1991), **91**, 165-195.

198 Churakova, E., Kluge, M., Ullrich, R., Arends, I., Hofrichter, M. & Hollmann, F. Specific Photobiocatalytic Oxyfunctionalization Reactions. *Angew. Chem.-Int. Edit.* (2011), **50**, 10716-10719.

199 Dent, J. Review article: pharmacology of esomeprazole and comparisons with omeprazole. *Aliment. Pharmacol. Ther.* (2003), **17**, 5-9.

200 Darwish, M., Kirby, M., Hellriegel, E. T. & Robertson, P. Armodafinil and Modafinil Have Substantially Different Pharmacokinetic Profiles Despite Having the Same Terminal Half-Lives Analysis of Data from Three Randomized, Single-Dose, Pharmacokinetic Studies. *Clin. Drug Invest.* (2009), **29**, 613-623.

201 Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. X. & Vogel, A. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angew. Chem.-Int. Edit.* (2005), **44**, 4192-4196.

202 Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. (2007), **2**, 891-903.

203 Reetz, M. T., Soni, P. & Fernandez, L. Knowledge-Guided Laboratory Evolution of Protein Thermolability. (2009), **102**, 1712-1717.

204 Valetti, F. & Gilardi, G. Improvement of biocatalysts for industrial and environmental purposes by saturation mutagenesis. (2013), **3**, 778-811.