Investigating Nicotinamide Cofactor Promiscuity in a New Class of Flavoprotein Monooxygenases

Chantel Nixon Jensen-Loughrey

Doctor of Philosophy

The University of York

Department of Chemistry

August 2014

Abstract

This thesis concerns the investigation into the structure and function of the 38.6 kDa FAD-containing flavoprotein *Stenotrophomonas maltophilia* flavin-containing monooxygenase, SMFMO, which was encoded from a gene from the marine bacterium *Stenotrophomonas maltophilia*. The enzyme was found to catalyse the asymmetric oxidation of prochiral sulfides and the regioselective Baeyer-Villiger oxidation of bicyclo[3.2.0]hept-2-en-6-one. SMFMO was unusual amongst FPMOs as it demonstrated an ability to employ either NADH or NADPH as nicotinamide cofactor in order to reduce the flavin for catalysis. In an effort to determine the residues responsible for the cofactor promiscuity of SMFMO the structure of SMFMO was determined and revealed that the cofactor promiscuity of SMFMO may be due to the substitution of an arginine residue, responsible for the recognition of the 2'-phosphate on the NADPH ribose in related NADPH dependent FMOs, with a glutamine residue in SMFMO. In an attempt to explore the cofactor determinants in SMFMO, the two residues Gln193 and His194 in the cofactor binding site of SMFMO were mutated in order to mimic the cofactor binding site of the NADPH-dependent FMO, mFMO, from *Methylophaga aminisulfidivorans* sp. SK1, in which structurally homologous residues Arg234 and Thr235 bind the 2'-phosphate on NADPH. mFMO possesses an asparagine residue which is thought to be involved in the stabilisation of the flavin hydroperoxide intermediate, in SMFMO this residue is replaced by Phe52. Mutation of the Phe52 residue revealed that this residue is a determinant in enantioselectivity. The natural variants of SMFMO, PFMO from *Pseudomonas stutzeri* and CFMO from *Cellvibrio* sp., also had the ability to use both nicotinamide cofactors equally to reduce the flavin. The structure of PFMO revealed that the residues Gln194 and Glu195, structurally homologous to Gln193 and His194 in SMFMO, were orientated away from the 2'-phosphate site and thus the Glu195 would not repel the negatively charged phosphate as originally thought.

Table of Contents

List of Figures

xii

List of Tables

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Gideon Grogan, it has been a pleasure to work for you. All your help with interpreting results and suggestions have helped me a great deal throughout my Ph.D. Thank you for sending me away to exotic places; to conferences in Sicily and Manchester and of course Diamond. Your friendliness, kindness and approachability have made the past four years a wonderful experience. Also, thank you for putting up with me and what only can be described as my 'blonde moments'- I'm sure Bafrica is a fantastic place.

Thank you to PML and my industrial supervisors Dr. Mike Allen and Sohail Ali. All your input throughout my Ph.D has helped me in countless ways and is greatly appreciated.

Thanks to Professor Neil Bruce for all your help and support during my Ph.D.

I would like to give a massive thank you to all the past and present members of the Grogan group (a.k.a the Grogan's angels). Thanks for the tea breaks, pub lunches, drinks, fun group meetings ending in the Willow and of course friendship.

Working in the YSBL family has been a delight and I could not have asked for a better place to be a part of. A huge thank you to Sally, Simon, Johan, Sam and Andrew, I'm extremely grateful for all the help you have given me along the way. Of course, thank you to everyone in YSBL for making the past four years wonderful.

Thank you to Professor Ian Fairlamb for allowing me to invade his lab for my organic chemistry needs.

I would like to give a huge thanks to my family, without them I wouldn't be the person I am today. Thank you to my mum and dad for supporting me throughout life and allowing me to do and be what I wanted, you are the best parents anyone could wish for. I would like to give a special thank you to my Nana, who has been there for me my whole life, and Natalie and Rob for being a great sister and brother-in-law.

Last but by no means least I would like to thank my husband Jonny. All your love and support has helped make this experience perfect – I love you.

I would like to dedicate this thesis to my Granda, John Nixon, though no longer with us I know you were with me every step of the way, I hope I've made you proud.

Declaration

Initial screening and preparation of genomic DNA was achieved by Plymouth Marine Laboratories (PML Applications Ltd). Sequencing and assembly of genomic DNA was carried out at the GenePool Genomics and Bioinformatics Facility at the University of Edinburgh. Cloning, expression and purification were performed by Dr. Jared Cartwright (Technology Facility, University of York). Initial gas chromatography assays and crystallisation experiments were performed by Dr. Gideon Grogan prior to the start of this Ph.D.

Data processing and structure solution for all protein structures was conducted by Dr. Gideon Grogan. LC-MS analysis was carried out by Karl Heaton (Chemistry Department, University of York).

Unless otherwise stated, I declare that this thesis and the work presented in it are my own and have been generated by me as a result of my own original research.

Chapter 1: Introduction

1.1 Enzymes within chemistry

1.1.1 The history of enzymes

Enzymes have been employed by humans for thousands of years in the form of fermentation in order to produce and preserve foodstuffs such as beer, wine and bread.[1] However, it was not until the 19th century when milestones in biocatalysis where reached. In 1858 Louis Pasteur discovered that treating an aqueous solution of racemic tartaric acid ammonium salt with a culture of the mould *Penicillium glaucum*, led to the consumption of (+)-tartaric acid over the consumption of (-)–tartaric acid,[2] a process which is considered a forerunner in enzyme catalysed kinetic resolution which is widely recognised in academia and industry today.[1b] Emil Fischer's carbohydrate work in 1894 discovered that the enzyme 'invertin' acts only upon α-D-glucosides, however, the enzyme 'emulsin' acts only upon β-D-glucosides. He recognised the ability of an enzyme to be selective and hypothesised the 'lock and key' characteristic of stereoselective enzyme catalysis.[3] Another landmark for biocatalysis was reached in 1897 by Eduard Büchner who observed that cell-free extracts of yeast containing no living cells were able to carry out the fermentation of sugar to alcohol and carbon dioxide, which proved that biological transformations do not necessarily require living cells.[1b, 4]

The appreciation that enzymes actually mediate biological catalysis allowed the growth of biochemistry, however, it was not until the end of the 20th century that enzymes as catalysts in organic chemistry gained importance.

1.1.2 Biocatalysts within organic chemistry

Over the last twenty to thirty years the use of enzymes as catalysts in chemical synthesis has increased, one reason being an enzymes ability to obey the twelve principles of green chemistry.[5] Biocatalysts are natural products that possess no toxicity issues, are biodegradable and are generally employed in aqueous media. The majority of biocatalysts work at ambient reaction conditions that produce non-toxic side products

that are easily disposed of thus reducing the risk of hazards. In addition, biocatalysts are catalytic and are highly specific catalysing chemo, regio and stereoselective reactions.[6]

Enzymes have the ability to act as chiral catalysts owing to the chiral environment of an enzyme's active site. The chirality of the active site ensures that the enzyme is naturally able to bind one enantiomer of the prochiral substrate over the other thus catalysing chiral functionalization of prochiral substrates and kinetic resolutions.[1a, 6] The fact that enzymes have the ability to resolve racemic mixtures of chemicals into their optically active compounds has become increasingly important in drug discovery. Due to the unique specificity of an enzyme, biocatalysts are capable of regioselective reactions in addition to stereoselective reactions and are able to functionalise one chemically equivalent site on a compound from many. An example of a regio and stereoselective reaction by a biocatalyst is the oxidative hydroxylation of the steroid progesterone to the 11α-product by the fungi *Rhizopus arrhizus* or *Aspergillus niger* (Figure 1.1), a compound that can be chemically altered to form the popular products cortisone and hydrocortisone.[1b, 7] The hydroxylation only occurs at the 11 position of the steroid and only in the α-orientation. Later it was discovered that hydroxylations of this nature are catalysed by cytochrome P450 enzymes.[1b, 8] This biocatalytic and semi-synthetic route offered a viable alternative to the 40-step synthesis of hydrocortisone proposed by Robert Woodward and co-workers in 1952.[9]
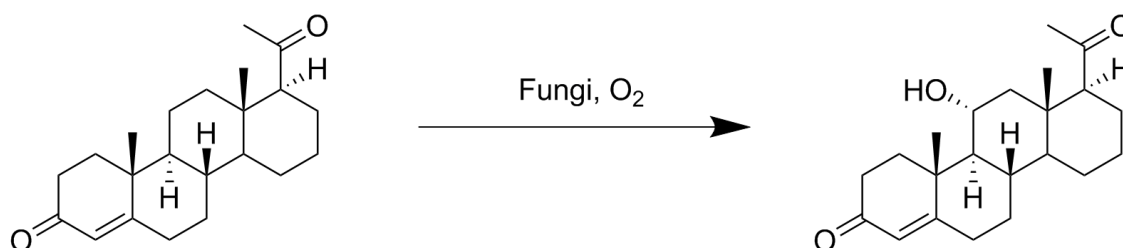
Figure 1.1 Regio- and stereoselective hydroxylation of progesterone to the 11α product by *Rhizopus arrhizus* or *Aspergillus niger* strains.

Natural products have played a huge role in modern medicine and isolation of such compounds has led to the discovery of biologically active compounds from quinine, morphine and penicillin to the potent anti-cancer drug taxol isolated from yew bark (Figure 1.2).
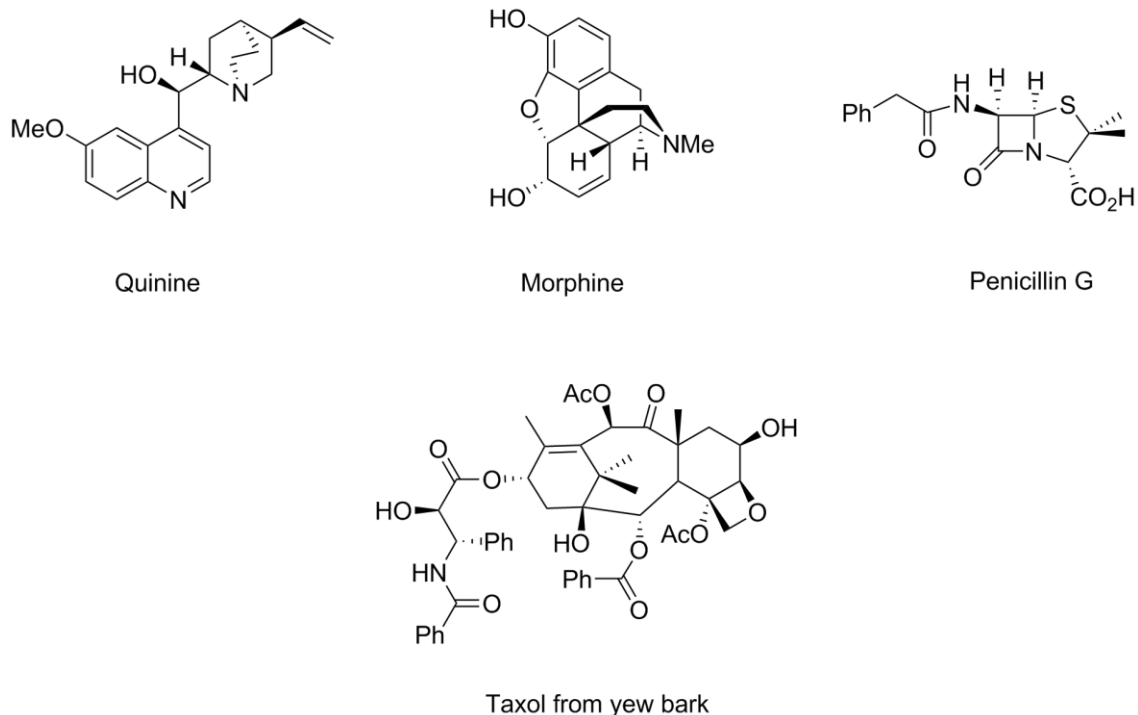
Figure 1.2 Structures of the natural products quinine, morphine, penicillin and taxol.

The structures of such compounds are so complex that synthesis by a chemical route is not feasible due to cost. However, the natural products were originally synthesised in nature using enzyme catalysed biosynthetic pathways and efforts to isolate such enzymes in order to produce natural products *in vitro* are of considerable interest in the biotechnology industry. An example of this is the production of penicillins using a naturally occurring enzyme penicillin acylase (Figure 1.3).[1a]
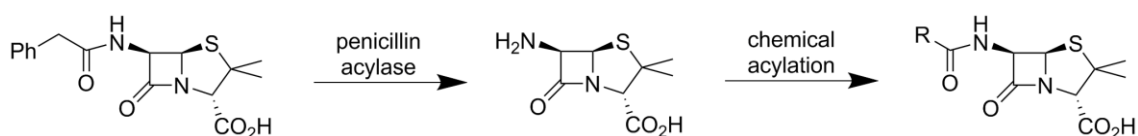
Figure 1.3 Example of the industrial production of semi-synthetic penicillins using penicillin acylase. E.g. for Penicillin V, R= PhOCH$_2$.

Within the chemical and pharmaceutical industries biocatalysis is emerging the method of choice for the synthesis of chiral compounds, as enantiomerically pure amino acids, amino alcohols, amines, alcohols and epoxides are extremely important classes of compounds in the pharmaceutical and agricultural industry.[10]

1.2 Biocatalysis using monooxygenases

The specific and efficient insertion of one oxygen atom into an organic compound is a difficult reaction to accomplish by chemical routes.[11] Chemical catalysts such as per-acids are employed to achieve oxygen insertion but the 'green' nature and unique specificity of enzymes that perform monooxygenations are unparalleled.[11-12] Such enzymes are called monooxygenases and are of increasing interest due to a wide range of reactivities and selectivities. Unfortunately, a small number of monooxygenases have been explored due to difficulties of enzyme expression and isolation. In addition, the majority of these enzymes depend on expensive coenzymes.[11] However, many of these issues have been improved due to better expression systems and coenzyme regeneration, thus exploring monooxygenases for their synthetic value has increased.[11, 13]

A popular example of a monooxygenase is the heme-containing family, P450 monooxygenases (EC 1.14.13, EC 1.14.14 and EC 1.14.15).[11, 14] P450s are relatively abundant and have been shown to catalyse a number of specific oxygenations. An interesting property of P450s is the ability to catalyse the specific hydroxylation of unactivated carbon atoms.[1a] Regioselective hydroxylations that are catalysed by P450 have been employed to alter steroids and sterols. The best known example of the P450 family is P450$_{cam}$, which catalyses the hydroxylation of camphor (Figure 1.4). It was shown that the reaction proceeds with retention of stereochemistry at the site of oxygen insertion.[1a, 8, 15]
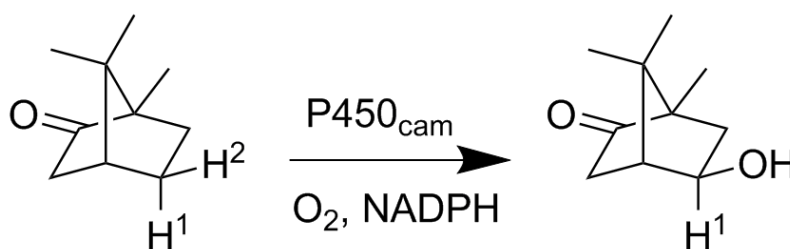
Figure 1.4 Stereospecific hydroxylation of camphor catalysed by P450$_{cam}$.

In addition to P450 monooxygenases, there are other monooxygenase classes such as non heme-dependent monooxygenases (EC 1.14.16), copper-dependent monooxygenases (EC 1.14.17 and EC1.14.18) and flavin-dependent monooxygenases (EC 1.13.12 and EC 1.14.13) which all rely on the presence of cofactor in some way.[1a,

[11, 14] In the past 10 years few monooxygenases that are independent of cofactor have been discovered.[11] For example, the monooxygenase ActVa-Orf6 in *Streptomyces coelicolor*[16] and the quinol monooxygenase YgiN from *E. coli*[17] both oxidise multiringed aromatic compounds without the aid of cofactor.

Many monooxygenase belong to the flavin-dependent family and have been attracting attention as selective oxidation catalysts in order to be used for the synthesis of expensive chemical building-blocks or pharmaceuticals, owing to their wide range of oxygenation reactions and enantio and regioselective properties.[11, 18]

1.3 Flavoprotein monooxygenases

1.3.1 Introduction

Flavoprotein monooxygenases or FPMOs are a class of oxidative enzymes that have important roles in eukaryotic metabolism pathways[19] and in the production of secondary metabolites in lower organisms.[20] In addition, they have the potential to acts as green catalysts for applications in asymmetric organic synthesis. Flavoprotein monooxygenases catalyse a wide range of reactions (Figure 1.5) including hydroxylations, epoxidations[21], Baeyer-Villiger oxidations[22] and sulfoxidations.[23]
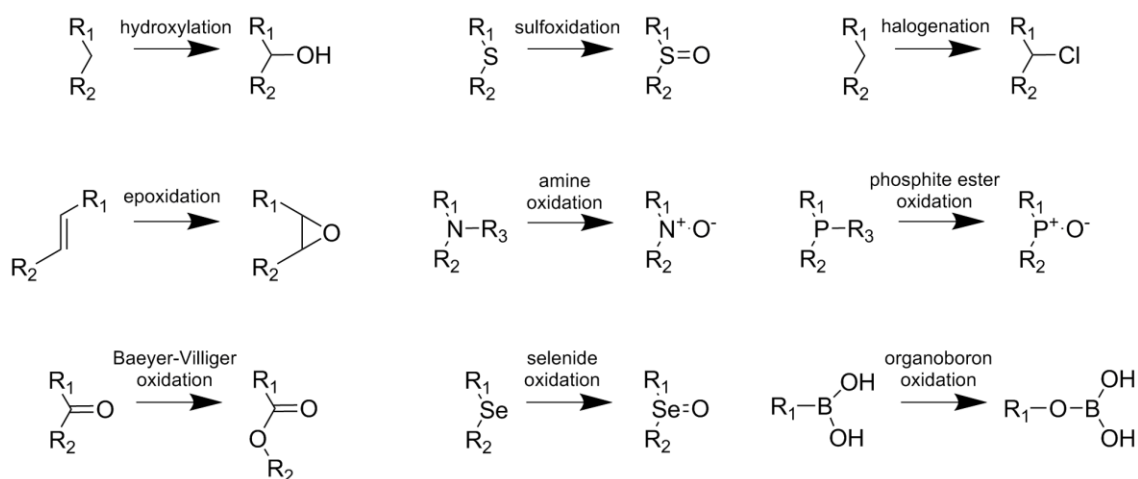
Figure 1.5 Reactions that are catalysed by Flavoprotein Monooxygenases.

As with many enzymes, the specific oxidation reaction and selectivity depend on the physical properties and chemical nature of the monooxygenase's active site. Although,

owing to sequence and structural homology several FPMOs subclasses can be classified, revealing that each subclass is only able to catalyse a limited range of oxidation reactions. This suggests that the structural fold of the FPMO dictates what type of oxygenation reaction is catalysed.[11]

1.3.2 The mechanism of a FPMO

In chemical synthesis molecular oxygen cannot be employed as an oxidant as the concerted reaction between O_2 and carbon in organic compounds is spin-forbidden, due to a pair of degenerate electrons in the anti-bonding orbital (π^*).[1a] In many cases, enzymes have discovered a way to utilise molecular oxygen as a substrate and oxygenate a compound. However, the oxygen must be activated by the enzyme in some way in order to allow the reaction.

The Pauli exclusion principle and Hund's rule state that electrons must occupy orbitals of increasing energies thus electrons occupy orbitals with the lowest energy first. In addition, a maximum of two electrons can occupy one molecular orbital and all orbitals at one energy level must be filled before pairing of electrons. Following this principle reveals that the ground state of molecular oxygen exists as a triplet ($^3\Sigma g^-$) with two degenerate unpaired electrons of parallel spin in the π_{2p}^* anti-bonding orbital, which prevents O_2 from reacting with other molecules which are generally in singlet state. O_2 can be activated to form two excited singlet states $O_2^*(^3\Sigma g^+)$ and $O_2^*(^1\Delta g)$ which are higher in energy and less stable than the ground state O_2 and differ in spin and occupancy of the π^* anti-bonding orbitals. In the excited states a spin flip occurs so the π^* anti-bonding electrons are of opposite spin (Figure 1.6). The excited $O_2^*(^3\Sigma g^+)$ has one electron in each of the π^* orbitals with opposite spins but is short lived and relaxes to the lowest-lying $O_2^*(^1\Delta g)$ excited state which is recognised as singlet oxygen and possess a pair of electrons in the same anti-bonding orbital with opposite spins and high-energy making it extremely reactive.[24]
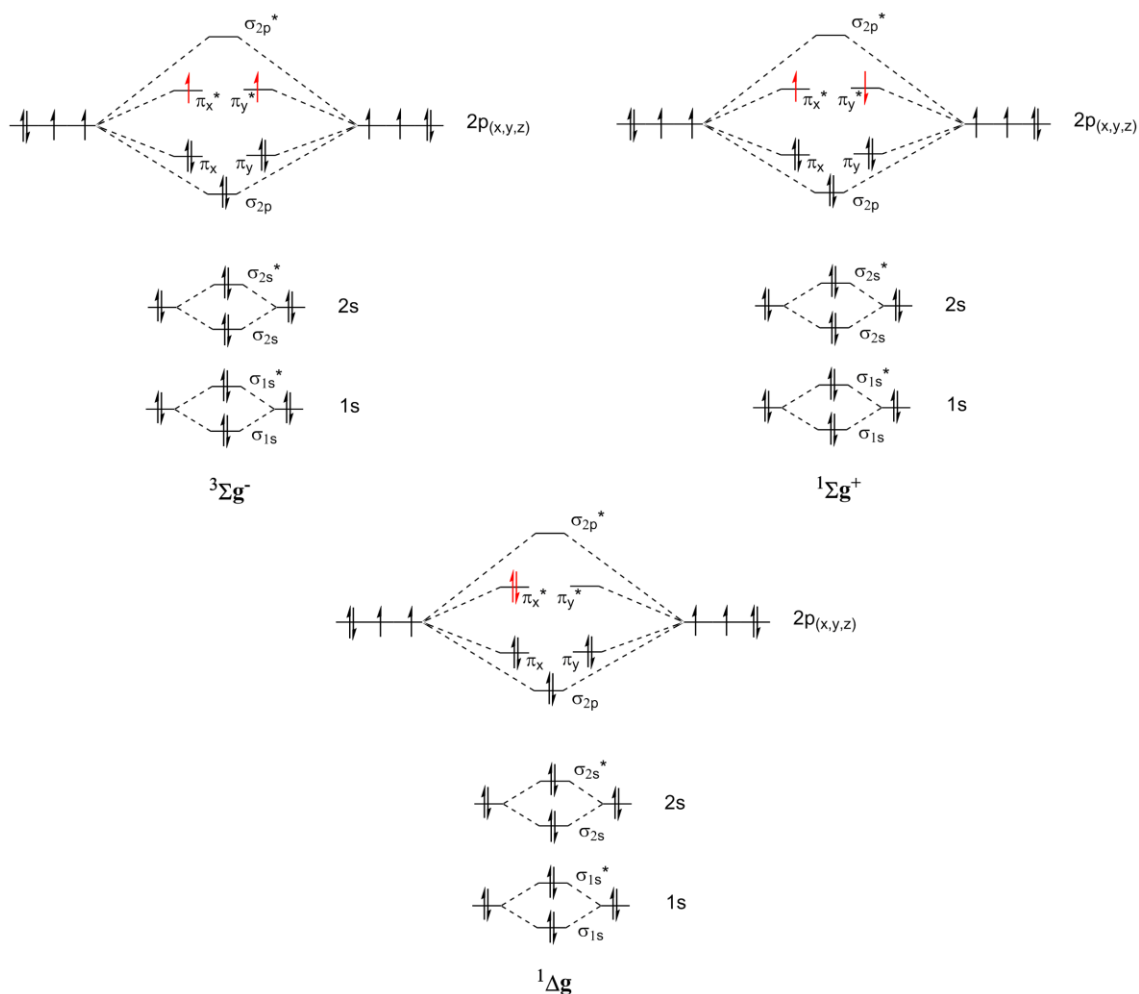
Figure 1.6 Molecular orbital diagrams illustrating the triplet ground state ($^3\Sigma g^-$) and singlet excited states ($^3\Sigma g^+$ and $^1\Delta g$) of molecular oxygen.
The triplet ground state can be excited to the singlet states for reaction. The $O_2{}^*(^1\Delta g)$ is the common excited state in which the electrons in the anti-bonding orbitals (in red) undergo a spin flip and orbital transfer.

To activate the molecular oxygen enzymes often use a transition metal such as iron which may or may not be bound to an organic cofactor, for example, heme in P450 monooxygenases. In the case of FPMOs however, a purely organic cofactor is used for oxygenation reactions.[11]

FPMOs depend on a flavin cofactor (Section 1.3.3) which must be reduced for reactivity with molecular oxygen. The electron rich flavin intermediate is able to use O_2 as a substrate and transfer one electron to the oxygen,[25] forming superoxide and the flavin radical. A spin inversion then occurs resulting in the formation of reduced oxygen.[26] For the majority of FPMOs, a covalent adduct between the $C_{4\alpha}$ of the flavin and molecular oxygen occurs forming the reactive intermediate $C_{4\alpha}$–hydroperoxyflavin. Peroxyflavins

are unstable and typically decay to form hydrogen peroxide and oxidised flavin (Figure 1.7).
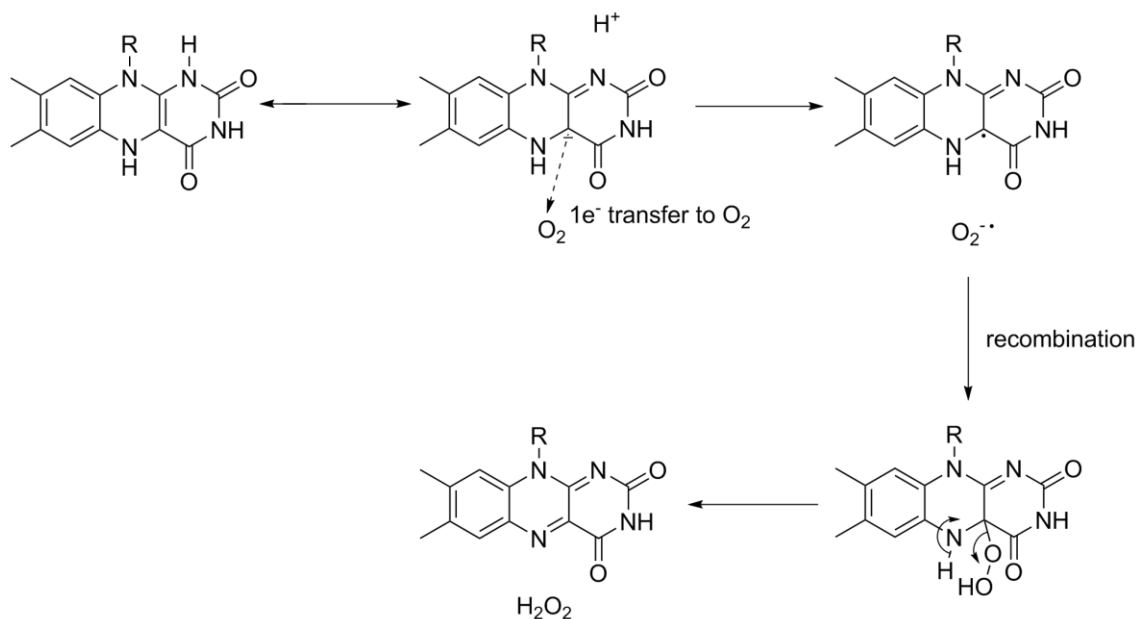
Figure 1.7 Reaction of flavin with molecular oxygen in oxidases. Peroxyflavin is unstable and will break down to form hydrogen peroxide and oxidised flavin.

Flavoprotein monooxygenases however are able to stabilise such a species and it can be used to insert a single oxygen atom into an organic compound.[27] Depending on the protonation state of the peroxyflavin intermediate, either a nucleophilic or electrophilic attack on the substrate is carried out (Figure 1.8).[11]
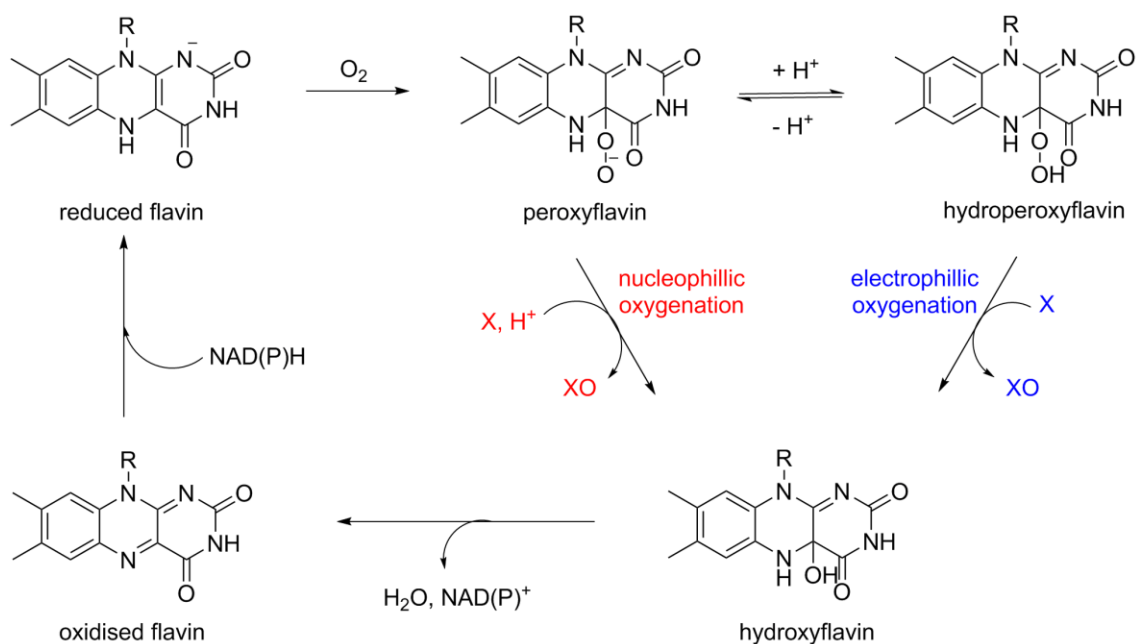
Figure 1.8 General mechanism of oxygenation reactions catalysed by flavoprotein monooxygenases.
The bound flavin is reduced by NADPH and reacts with O_2 to give (hydro)peroxyflavin which undergoes a nucleophilic or electrophilic attack on the substrate.

1.3.3 Cofactors and their roles

Many enzymes, including FPMOs depend on small molecules, known as cofactors, to aid catalysis. FPMOs depend on a riboflavin-derived flavin cofactor. The structure of riboflavin consists of an isoalloxidine heterocyclic ring, which contains the $C_{4\alpha}$ carbon atom responsible for oxygenation. FPMOs can depend on one of two flavins: flavin mononucleotide (FMN) in which the riboflavin is attached to a phosphorylated ribitol sidechain or flavin adenine dinucleotide (FAD) were the riboflavin has a diphosphate linkage attached to an adenosine molecule (Figure 1.9). An obvious property of riboflavin is that it has a bright yellow colour, a characteristic that is observable in FPMOs in solution.[1a]
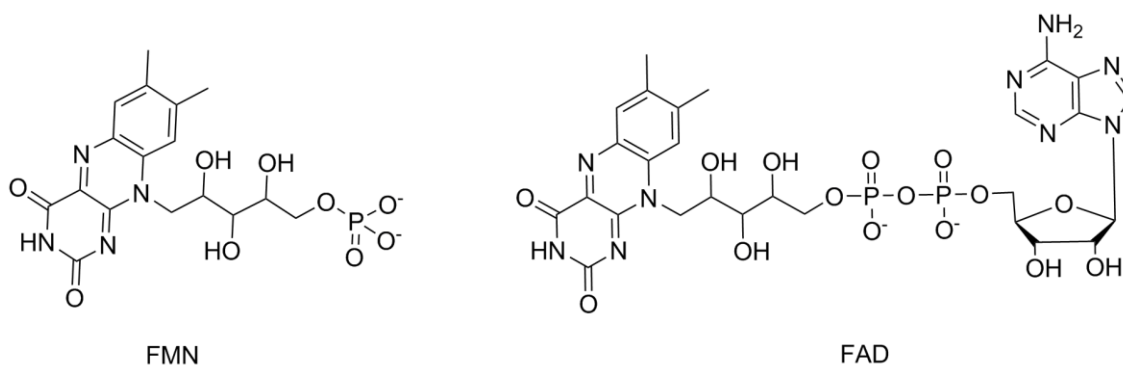
Figure 1.9 Structures of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), cofactors present in FPMOs.

FPMOs can be broadly classified into two major subgroups, external flavoprotein monooxygenases which require external cofactors such as NAD(P)H to reduce the flavin (EC 1.14.13) or internal flavoprotein monooxygenases in which the flavin is reduced by the substrate (EC 1.13.12). A known example of an internal FPMO is lactate monooxygenase, in which lactate is oxidised to pyruvate and the flavin is reduced.[28] However, examples of internal FPMOs are extremely rare. In addition, there is another group of flavoenzymes that contain flavin but it is not directly involved in the oxygenation reaction and the oxygen atom that is inserted is from water, for example vanillyl-alcohol oxidase.[11]

The majority of FPMOs contain a non-covalently bound FMN or FAD, although in some cases the flavin is covalently bound. For example in vanillyl-alcohol oxidase the FAD is covalently linked to a histidine through the 8α- methyl group of FAD to the N3 atom of the histidine.[29] However, for the majority of internal and external FPMOs the flavin moiety is bound non-covalently.[1a, 11]

External flavoprotein monooxygenases depend on one of two nicotinamide cofactors to reduce the flavin, nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADPH), which bears a 2'-phosphate on the adenosine moiety (Figure 1.10).[1a, 6]
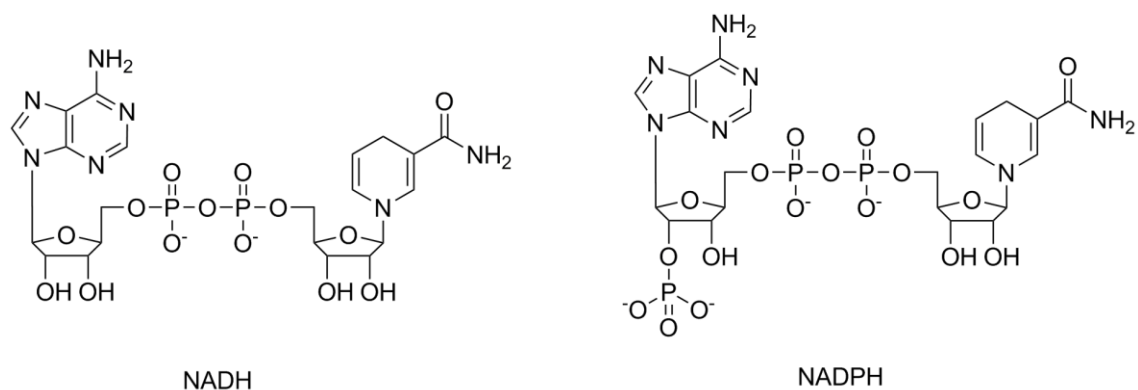
Figure 1.10 Structures of nicotinamide cofactors NADH and NADPH employed by FPMOs to reduce flavin.

The redox-active part of the cofactors is the nicotinamide heterocyclic ring. NAD(P)$^+$ is the oxidised form which is a pyridinium salt and can be reduced to a 1,4-dihydro-pyridine NAD(P)H (Figure 1.11).[1a]
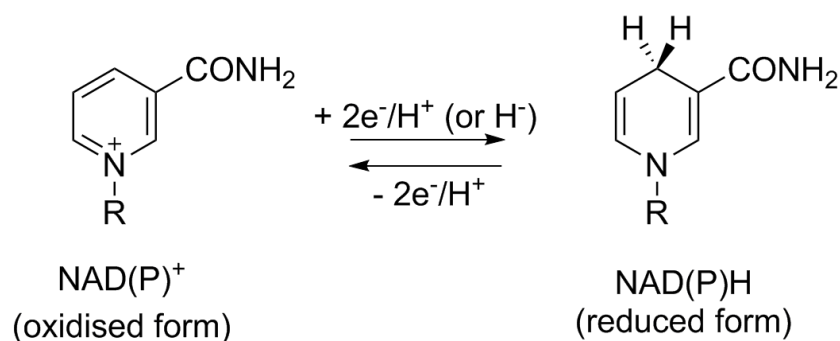
Figure 1.11 Diagram illustrating the oxidised and reduced forms of the nicotinamide cofactors.

In whole cell biotransformations, the organism generally produces and recycles such small molecule cofactors. In biotransformations involving isolated enzymes however, the cofactors may need to be added separately. As NAD(P)H is not bound to the enzyme it needs to be added in stoichiometric amounts, leading to high expense, or more commonly, recycled using and appropriate auxiliary enzyme and auxiliary substrate.

NADH retails at £148.50 per 250 mg which is approximately a third of the price when compared to NADPH (£364 per 250 mg).[30] Therefore, NADH is more economically

11

interesting for use as cofactor in FPMOs; however, FPMOs that employ NADPH as cofactor rarely use NADH as efficiently, if at all.[6]

1.3.4 Classification of FPMOs

Classification of flavin-dependent enzymes has be carried out using specific criteria, such as the kind of chemical reaction catalysed, the characteristics of redox substrates, homology in sequence or topology of 3D structure.[31] In 2006, a review by van Berkel and Fraaije grouped external flavoprotein monooxygenases into six classes, A-F. The distinction between the different subclasses was based on sequence similarity and specific structural features which are reflected in sequence similarity and conserved sequence motifs.[11]

Class A flavoprotein monooxygenases are encoded by a single gene and have a tightly bound FAD cofactor. This type of FPMO can employ either NADH or NADPH as nicotinamide cofactor to reduce the flavin. This type of enzyme are typically involved in the microbial degradation of aromatic compounds by *ortho-* or *para-* hydroxylation of the aromatic ring.[32] They have limited substrate scope with typical substrates being aromatic compounds that contain an activating hydroxyl or amino group. The first enzyme of this subclass to be characterised was salicylate 1-monooxygenase (EC 1.14.13.1) from *Pseudomonas putida* in 1972.[33] However, the most extensively studied Class A FPMO is 4-hydroxybenzoate 3-monooxygenase (EC 1.14.13.2) from *Pseudomonas*, involved in the degradation of aromatic compounds.[27, 34] In addition to hydroxylations, enzymes belonging to this subclass are also involved in the catalysis of epoxidations reactions and Baeyer-Villiger oxidations. The latter involving the bacterial enzyme MtmOIV found to be involved in the biosynthesis of mithramycin, an anticancer drug and calcium reducing compound.[35] Class A FPMOs have been found to be structurally composed of one dinucleotide binding domain binding FAD.[11]

Class B are encoded by a single gene and possess a tightly bound FAD. Such enzymes depend on NADPH to reduce the bound flavin and keep the NADPH/NADP$^+$ cofactor bound during catalysis. In this subclass the flavin-containing monooxygenases are referred to as multi-functional as they are able to oxidise (hetero)-atoms in addition to carbon atoms. Three sequence-related FPMOs where found to make up this subclass: flavin-containing monooxygenase (FMOs), microbial N-hydroxylating

monooxygenases (NMOs) and Baeyer-Villiger monooxygenase (BVMOs).[36] Class B FPMOs will be discussed further in Section 1.4.

Class C FPMOs are encoded by multiple genes encoding one or two monooxygenase components and a reductase component. The monooxygenases employ FMN as cofactor which is generated by reductase. The reductase can utilise either NADPH or NADH as nicotinamide cofactor. Bacterial luciferases (EC 1.14.14.3) are the most comprehensively studied monooxygenase of this subclass and are useful for the oxidation of long chain aliphatic aldehydes.[37] Several BVMOs also fall under this subclass such as, 2,5-diketocamphane 1,2-monooxygenase and 3,6-diketocamphane 1,6-monooxygenase (EC 1.14.15.2) from *Pseudomonas putida* ATCC 17453, involved in the degradation of the compound camphor.[38] The structural core of class C monooxygenases subunits have been shown to display TIM-barrel fold.[37]

Class D FPMOs are encoded by two genes which encode a monooxygenase and a reductase. This type of enzyme employs reduced FAD as flavin cofactor which is generated by the reductase which can use either NADPH or NADH as nicotinamide cofactor. Monooxygenases in this subclass are shown to be restricted to one type of oxygenation, hydroxylation. As with class A FPMOs, the substrate scope of class D FPMOs is limited to aromatic compounds such as, 4-hydroxyphenylacetate, phenol[39] and 4-nitrophenol.[40] The best known monooxygenases in this subclass is 4-hydroxyphenylacetate 3-monooxygenases (EC 1.14.13.3) from *E. coli* W[41] and *Acinetobacter baumanii*.[42] There is no structure available for this subclass but sequence homology suggests a structural resemblance to acyl-CoA dehydrogenase fold.[11]

Similar to class D monooxygenases, class E FPMOs are also encoded by two genes encoding a monooxygenase and a reductase which generates reduced FAD. This subclass can utilise either nicotinamide cofactor, NADH or NADPH. The two-component monooxygenase subclass is represented by styrene monooxygenase from *Pseudomonas sp*. VLB120 which catalyses the conversion of styrene onto (*S*)-styrene oxide at an enantiomeric excess higher than 99 %.[43] A recombinant *E. coli* system was developed that expressed StyA (monooxygenase) and StyB (reductase)[44] and the whole-cell biocatalyst allowed for a wide range substrate scope that also gave access to chiral aryloxides.[45] The system was successfully scaled up and it allowed approximately 400 g of (*S*)-styrene oxide to be produced at pilot scale.[46] Class E FPMOs are relatively rare as a very limited number of monooxygenases falling under this subclass are known.

Sequence analysis indicates the presence of one dinucleotide domain indicating a possible evolutionary link with the class A FPMOs.[11] However, the structure of a class E styrene monooxygenase (3IHM) was found to exist as a homodimer in which each monomer forms two domains, an FAD domain and what is likely to be the styrene binding domain.[47]

Class F FPMOs are also two-component monooxygenases which are encoded by two genes encoding a monooxygenase and a reductase. This subclass uses reduced FAD which is generated by the reductase that can employ either NADH or NADPH as nicotinamide cofactor. Enzymes in this subclass catalyse halogenation reactions. The most extensively studied flavin-dependent halogenase is tryptophan 7-halogenase.[48] The reaction catalysed by class F FPMOs does not result in an oxygenated product, however, the proposed catalytic mechanism is similar to that of a flavoprotein monooxygenase. The structure of such monooxygenases was found to have two domains, a helical domain and a FAD binding domain, which is compatible with a 'monooxygenase' type of catalytic mechanism.[11, 49]

1.4 Class B flavoprotein monooxygenases

As discussed above class B FPMOs is comprised of three sequence-related monooxygenases, FMOs, NMOs and BVMOs. Each member of these families are single-component FAD-dependent enzymes and rely on NADPH as nicotinamide cofactor. The protein sequences of such monooxygenases possess two Rossmann-fold motifs (GXGXXG) which are responsible for two binding domains for FAD and NADPH.

1.4.1 Baeyer-Villiger monooxygenases

1.4.1.1 The Baeyer-Villiger reaction

Over a century ago in 1899, it was discovered that a mixture of sodium persulfate and concentrated sulfuric acid (Caro's acid) could oxidize menthone to its corresponding lactone (Figure 1.12).[50]
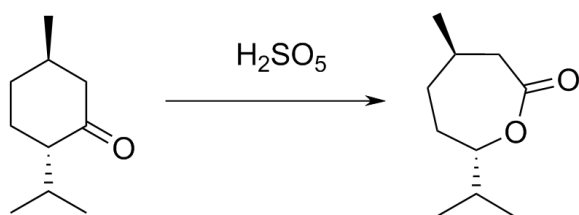
Figure 1.12 Menthone oxidation using Caro's acid.

This reaction became known as the Baeyer-Villiger reaction and upon replacement of the persulfuric acid with an organic peracid it is now one of the most successful and frequently used reactions in organic synthesis. The oxidation reaction's success is largely due to the fact that a range of carbonylic compounds can be converted into their subsequent esters or lactones via an oxygen-insertion process. The reaction tolerates a great proportion of functional groups and can be regio and stereo selective. In addition, its success can also be attributed to the variety of oxidants that can be successfully employed (Figure 1.13).[50b, 51]
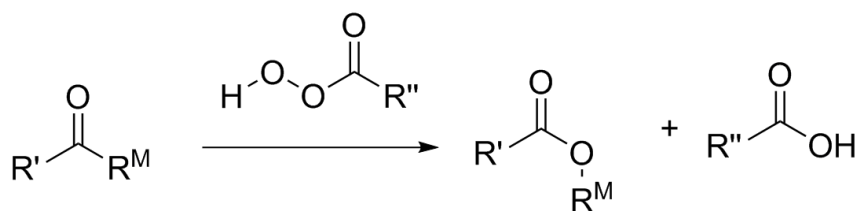
Figure 1.13 Baeyer-Villiger oxidation reaction involving a peracid.

Many efforts were made in order to understand the mechanism of the reaction discovered by Adolf Baeyer and Victor Villiger and the accepted mechanism was proposed by Criegee in 1948.[52] This two-step mechanism precedes *via* a nucleophilic attack by a per-oxo species on a ketone's carbonyl group which results in the formation of the tetrahedral "Criegee" intermediate. The intermediate then undergoes a rearrangement to produce the subsequent ester or lactone. For the migrating step to occur, stereoelectronic requirements must be fulfilled, which are: the migrating group R^M must be antiperiplanar to the peroxy bond (O-O) and antiperiplanar to the lone pair of the hydroxyl group (Figure 1.14).[12, 22c, 50b]

"Criegee" intermediate
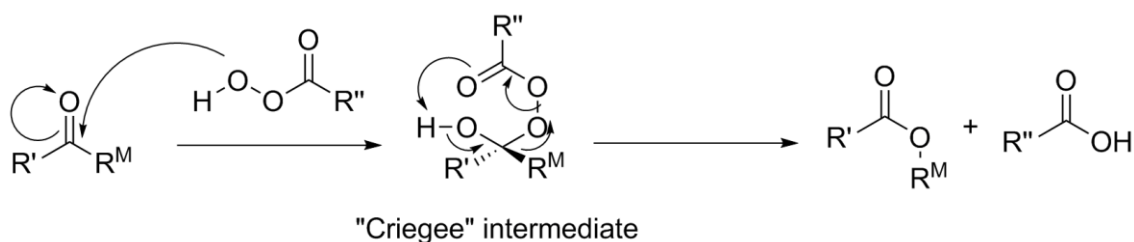
Figure 1.14 The reaction mechanism for a chemical Baeyer-Villiger oxidation. The peracid acts as a nucleophile and attacks the ketone to form an intermediate which rearranges to give the product and carboxylic acid.

A large number of peracids can be used as oxidants in the BV reaction such as trifluoroperoxyacetic acid, *m*-chloroperoxybenzoic acid, peroxyacetic acid and hydrogen peroxide. However, the use of such peracids leads to disadvantages within the BV reaction, including the formation of the corresponding carboxylic acid which then needs to be disposed of. In addition, these acids are expensive and hazardous thus diminishing their commercial availability. Hydrogen peroxide is a little different to the other peracids as it is perceived to be a much "greener" approach to the BV reaction. It is a cheap, safe and clean oxidant producing water as a side product. Unfortunately, this water can cause a problem in which the resulting ester is hydrolysed leading to a lower substrate scope. Furthermore, due to its lower oxidative activity a catalyst must be used which can lead to further disadvantages. Thus, attention turned to a much greener approach in the form of Baeyer-Villiger monooxygenases.[22, 50b]

1.4.1.2 The biocatalytic Baeyer-Villiger reaction

Class B BVMOs catalyse the Baeyer-Villiger oxidation of ketones (or aldehydes) to its corresponding ester. In addition, owing to the advanced regio- and stereoselectivity nature of enzymes, BVMOs can also catalyse the asymmetric oxygenation of racemic chiral ketones to useful asymmetric lactones (Figure 1.15).[11, 22]
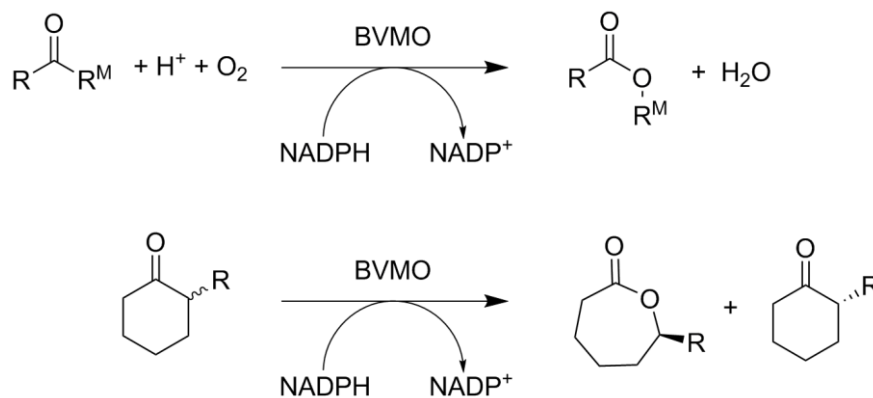
Figure 1.15 BVMO catalysed Baeyer-Villiger reactions.

As with the many FPMOs, the enzyme catalysed Baeyer-Villiger reaction is initiated by the reduction of FAD by NADPH followed by an immediate oxidation by molecular O_2 to give a flavin peroxidate anion, which is thought to be the active catalyst in the oxygen-insertion process.[53] In the absence of substrate the peroxide lives in its most stable reduced form. The peroxide acts as a nucleophile and attacks the carbonyl group of the ketone substrate forming the tetrahedral 'Criegee intermediate'. Rearrangement of the intermediate forms the desired lactone product and hydroxyflavin. The loss of H_2O regenerates FAD and the product and cofactor are released (Figure 1.16).[22a]
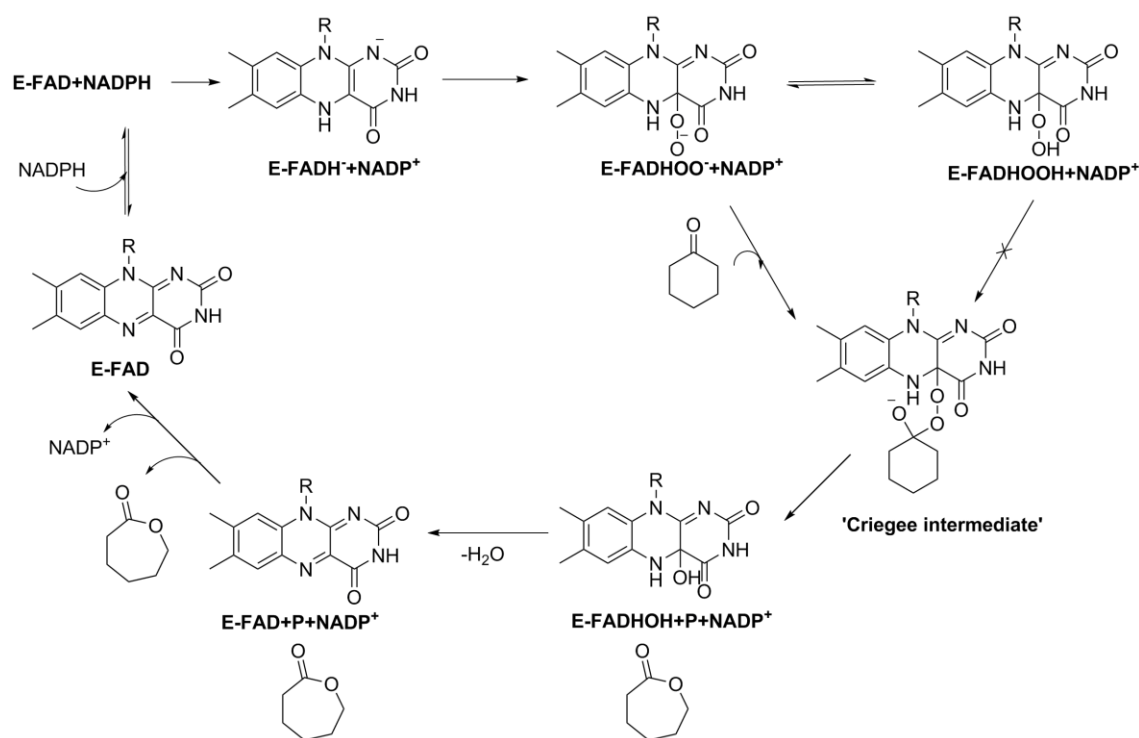
Figure 1.16 The catalytic cycle of a class B Baeyer-Villiger monooxygenase. NADPH reduces the flavin which reacts with molecular oxygen to produce peroxyflavin. The peroxyflavin attacks the ketone to form an intermediate that undergoes a rearrangement to yield the corresponding lactone.

The accepted mechanism of BVMOs today stems from the results generated for the BVMO cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* NCIMB 9871 (CHMO$_{9871}$) and proposed steps and rate constants by Walsh *et al.* in 1982 and more recently by Massey *et al.* in 2001.[53-54]

1.4.1.3 Known Baeyer-Villiger monooxygenases

Numerous class B BVMOs have been cloned and characterised.[36, 55] Cyclohexanone monooxygenase (CHMO) was the first BVMO to show catalytic potential and since 1976 when Trudgill *et al.* originally isolated and characterized CHMO from *Acinetobacter* sp. NCIMB 9871 (E.C. 1.14.13.22) it has become the most studied BVMO to date.[11, 56] Currently over one hundred substrates have been identified for this enzyme. CHMO has been shown to accept heteroatom compounds in addition to a wide range of carbonylic compounds (Figure 1.17).[22b]
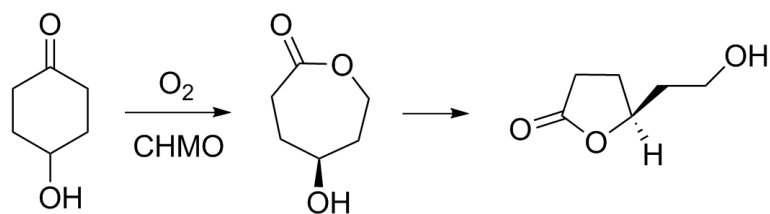
Figure 1.17 An example of a Baeyer-Villiger reaction catalysed by CHMO.
The BV reaction of the ketone produces the lactone which spontaneously rearranges to give the final product.

Its ability to accept such a vast array of compounds along with its high regioselective and enantioselective properties has allowed CHMO and thus BVMOs to become very desirable green oxidative biocatalysts. A number of CHMO homologs have been discovered and overexpressed allowing comparative biocatalytic studies which revealed that although sequence similarity can be high between the homologs, the regioselectivity can vastly differ.[57] In addition, directed evolution studies by Reetz *et al.* identified that the enantioselectivity of CHMO can be dramatically improved by mutating a single amino acid.[58] The structure of CHMO from a mesophilic species of *Rhodococcus* complexed with $NADP^+$ and FAD was solved (Figure 1.18).
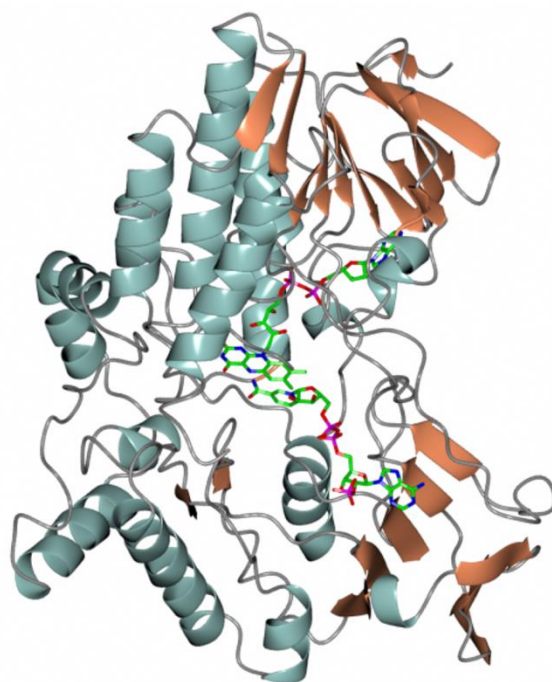
Figure 1.18 Structure of the class B BVMO, CHMO from *Rhodococcus* complexed with FAD and NADP$^+$ (3GWF).
The protein back bone is shown in ribbon format. The NADP$^+$ and FAD are shown in cylinder format with carbon atoms in green.

The structure is shown to consist of two domains, including both FAD and NADPH binding domains, with the active site present in a cleft at the domain interface. In addition, it was found that the conserved BVMO motif is on the surface of the enzyme away from the implied active site and it was suggested that this fingerprint sequence may be involved in conformational changes of the protein during the catalytic cycle.[59]

Since the discovery of CHMO, many BVMOs have now been discovered and characterised such as, cyclopentanone monooxygenase (CPMO) from *Comamonas* sp. NCIMB 9872 (EC 1.14.13.16),[55] 4-hydroxyacetophenone monooxygenase (HAPMO) *Pseudomonas fluorescens* ACB (EC 1.14.13.84)[60] and phenylacetone monooxygenase (PAMO) from the thermophile *Thermobifida fusca* (EC 1.14.13.92).[61]

PAMO was discovered by genome mining and its gene was discovered in the genome of the thermophile *Thermobifida fusca* which allowed PAMO to be relatively thermostable.[61b] In addition, PAMO has been shown to tolerate solvents which are positive characteristics as other BVMOs can be limited due to instability.[11, 62] The main substrate scope for this enzyme are aromatic ketones although, aromatic sulfides,

sulfoxides, aliphatic ketones and organoboron compounds are also oxidised by this enzyme.[61b, 63] In 2004, PAMO was the first monooxygenase in this subclass for which the crystal structure was determined (Figure 1.19).[61a]
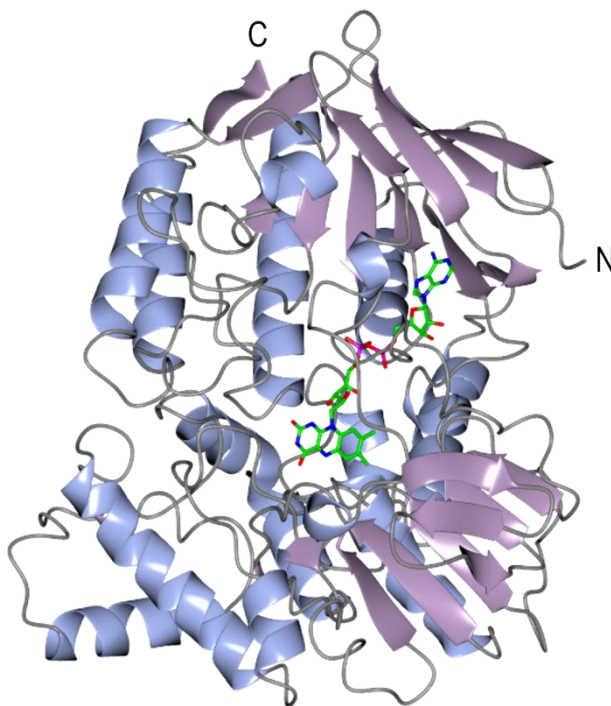
Figure 1.19 Structure of the class B BVMO, PAMO from *Thermobifida fusca* complexed with FAD (1W4X).
The protein back bone is shown in ribbon format. The FAD is shown in cylinder format with carbon atoms in green.

The structure was found to possess two Rossmann-fold domains binding the FAD and NADPH cofactor. Within the structure the dinucleotide binding domains are flanked by two helical domains. Both CHMO and PAMO structures are very similar and possess unique characteristics for this type of FPMO.[11]

1.4.1.4 Conserved motifs

In 2002, Fraaije and co-workers published findings identifying an amino acid sequence named the 'BVMO motif' (FXGXXXHXXXW(P/D)), which is highly conserved throughout Bayer-Villiger monooxygenases. Two other conserved regions known as Rossmann motifs (GXGXXG) were also found to be present throughout the BVMO

family and were reported to aid binding of the ADP moiety present in each cofactor. The length of BVMOs generally falls between 500 and 550 amino acids with a molecular weight around 60 kDa. The first Rossmann sequence lies near the N-terminus whilst the second is approximately 200 amino acids away. The BVMO motif sequence is said to be located closely upstream to the second Rossmann motif and it is this BVMO motif that allows for the identification of new BVMOs from sequenced genomes.[36]

1.4.2 Flavin monooxygenases

Flavin monooxygenases (FMOs) (EC 1.14.13.8) formerly named 'mixed-function oxidases' were originally acknowledged in liver enzymes. FMOs are present in all mammals and other eukaryotic organisms. In humans several isoforms exist, with each allowing a specific substrate and tissue localisation. Within the human genome, six FMO genes and five pseudo-genes have been identified.[11] Of these, FMO3 is recognised as the most important isoform in the liver as it plays an important role in the detoxification of drugs, biodegradation of aromatic compounds[64] and other xenobiotics in the human body.[65] It has been reported that mutations in FMO3 cause the inheritable disease trimethylaminuria (TMAU) or 'fish odour syndrome' due to the build-up of trimethylamine in body fluids.[65] FMOs are said to complement the activity of the cytochrome P450 system and catalyse the oxygenation of carbon bound reactive heteroatoms such as, nitrogen, sulfur, phosphorus, selenium or iodine.[11, 66] All mammalian FMOs have a strong membrane association with poor water solubility[64] and so far no structure of a mammalian FMO is available.[67]

1.4.2.1 Known FMOs

In addition to mammalian FMOs, non-mammalian eukaryotic FMOs have also been identified. Such FMOs were isolated from yeast from *Saccharomyces cerevisiae* which was shown to be involved in the oxidation of biological thiols such as glutathione.[68] It was revealed that the yeast FMO is required for proper folding of the disulphide containing proteins by generating an oxidising environment in the endoplasmic reticulum.[69] Another yeast FMO from *Schizosaccharomyces pombe* has also been identified which was used in the oxidation of methimazole. Crystal structures of the

yeast FMO from *S. pombe* were determined and revealed a two domain structure, a FAD binding domain and NADPH binding domain. In addition, structures of the yeast enzyme bound to substrate revealed that the substrate binds to the same site as the nicotinamide cofactor.[70] A plant FMO was also identified from *Arabidopsis* which was found to catalyse the hydroxylation of tryptamine involved in the biosynthesis of auxin.[71]

The first bacterial FMO was cloned by Choi *et al*. from the methylotropic bacterium *Methylophaga* sp. strain SK1 and named mFMO. Overexpression and characterisation of mFMO revealed the enzyme was dependent on NADPH as nicotinamide cofactor and had the ability to oxidise a variety of amines. It was also found that mFMO could readily oxidise indole as large amounts of indigo were formed by the recombinant *E. coli* cells expressing the enzyme.[72] In addition, mFMO fused with phosphite dehydrogenase (PTDH) has recently been employed for the asymmetric oxidation of a range of prochiral sulfides to give predominantly the (*S*)-sulfoxide product with moderate to high enantiomeric excess (Figure 1.20).[73]
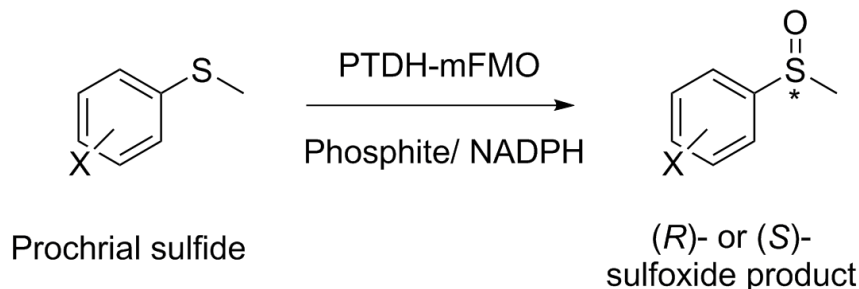
Figure 1.20 Sulfoxidation reactions of a range of prochiral sulfides catalysed by PTDH-mFMO.

Structural studies allowed for the structure of mFMO to be discovered and although smaller than class B BVMOs at 53 kDa the overall structure was relatively similar. The mFMO structure was found to possess two distinct domains, a larger FAD binding domain and a smaller NADP$^+$ binding domain (Figure 1.21).[67] In addition, the molecular determinants responsible for the binding of the 2'ribose phosphate of NADPH in mFMO, and those which discriminate between NADPH and NADH have been identified.[74]
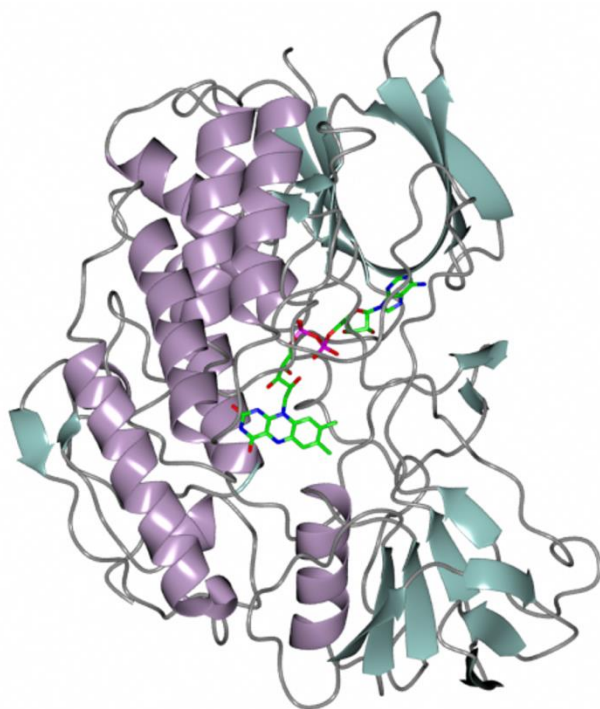
Figure 1.21 Structure of the class B FMO, mFMO from *Methylophaga* sp. strain SK1complexed with FAD (2XLT).
The protein back bone is shown in ribbon format. The FAD is shown in cylinder format with carbon atoms in green.

1.4.2.2 The catalytic mechanism of FMOs

FMOs can act on a range of non-polar substrates that contain soft nucleophile heteroatoms.[64] As with class B BVMOs, class B FMOs are single component polypeptides that use NADPH as nicotinamide cofactor to reduce the bound FAD,[75] forming $FADH_2$, which reacts with molecular oxygen to yield the $C_{4\alpha}$-hydroperoxyflavin intermediate. In the case of sulfides, this intermediate undergoes a nucleophilic attack by the sulphur atom. The intermediate is responsible for the insertion of the oxygen atom into the substrate, which allows for the production of the hydroxyflavin intermediate. The loss of water allows for the restoration of cofactor.[64, 67] An extraordinary feature of the FMOs catalytic cycle is the ability to stabilise the hydroperoxyflavin intermediate until a substrate is present.[67, 76] It is thought that NADPH is essential for intermediate stabilisation and thus remains bound throughout the catalytic cycle and thus prevents the enzyme from working as a NADPH oxidase, which would produce hydrogen peroxide and oxidised flavin (Figure 1.22).[67, 77] However, recent reports suggest that structural changes in the active site weaken the

binding of $NADP^+$ and interactions between the peroxyflavin and residues present in the active site accelerate the cofactors release and also stabilise the hydroperoxyflavin until substrate is introduced.[74] In the absence of substrate the FMO returns to its oxidised form releasing hydrogen peroxide.[78]
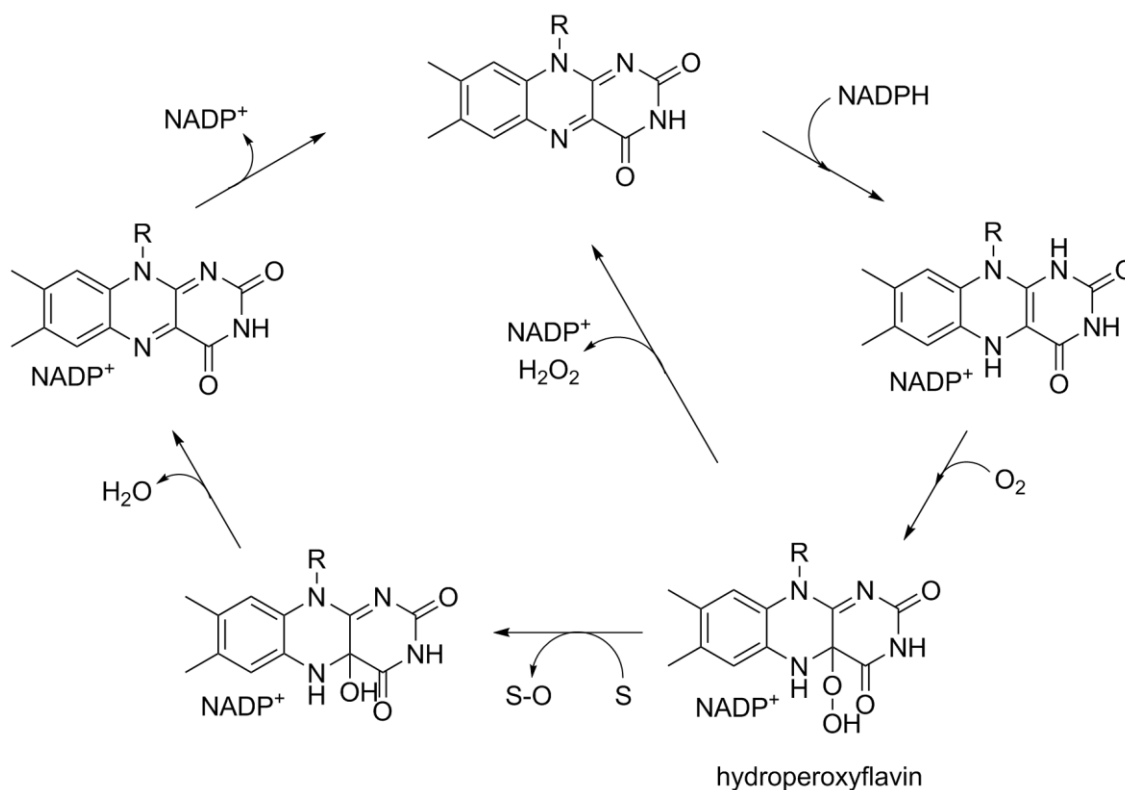
Figure 1.22 The catalytic cycle of a class B Flavin Monooxygenase.
FAD is reduced by NADPH and reacts with O_2 to form the reactive hydroperoxyflavin intermediate which undergoes a nucleophilic attack from the substrate (S) and oxygen insertion occurs (S-O).

1.4.2.3 Conserved motifs

Similar to class B BVMOs, FMOs possess two Rossmann motifs (GXGXXG) which are responsible for the recognition of the ADP moiety in FAD and the nicotinamide cofactor NADPH. FMOs also have an amino acid sequence that is individual to FMOs (FXGXXXHXXXY/F), known as the 'FMO motif' and thus can be used to distinguish between class B BVMOs and FMOs. The 'FMO motif' can be used to identify possible new FMOs however, in contrast to eukaryotic FMOs bacterial FMOs are quite rare.[11, 36]

1.4.3 N-hydroxylating monooxygenases

Microbial N-hydroxylating monooxygenases (NMOs) catalyse the N-hydroxylation of long chain primary amines and include the ornithine hydroxylase from *Pseudomonas aeruginosa* (PvdA),[79] the ornithine hydroxylase from Aspergillus *fumigatus* (SidA),[80] and the lysineN(6)-hydroxylases.[81] NMOs play a role in the biosynthesis of certain bacterial and fungal siderophores.[11] Cloning and sequencing of several NMO genes revealed that the monooxygenases share sequence homology with FMOs[82] and rely on NADPH for catalysis.

The NMO PvdA catalyses the FAD-dependent hydroxylation of the ornithine side chain amine using NADPH as the electron donor and molecular oxygen (Figure 1.23), however the FAD is not stably bound.[79c]
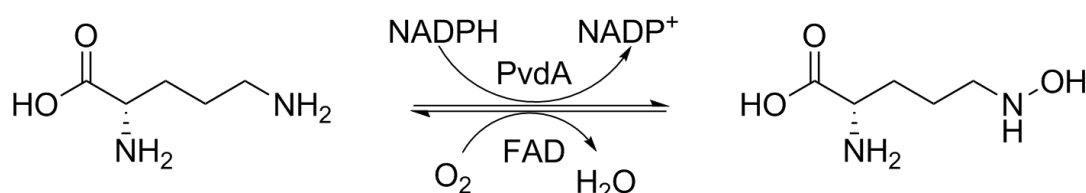
Figure 1.23 Reaction scheme for the hydroxylation of ornithine by the NMO PVdA from *Pseudomonas aeruginosa*.

Recently two structures of PvdA have been determined with the FAD in both oxidised and reduced form and are the first structures of a class B NMO. PvdA is comprised of three domains, two Rossmann-like domains for the recognition of FAD and NADPH and a substrate binding domain (Figure 1.24). The structures possess bound NADPH and substrate (hydroxy-ornithine) in a solvent-exposed active site. Structural and biochemical evidence suggest that $NADP^+$ remains bound throughout the oxidative half of the reaction, which is said to stabilise the flavin intermediates.[83]

Figure 1.24 The three domain structure of PvdA from *Pseudomonas aeruginosa* complexed with FAD and NADPH (3S5W, 3S61).
The protein back bone is shown in ribbon format. The FAD and NADPH are shown in cylinder format with carbon atoms in green.

A recent report by Robinson *et al*. involving a class B NMO, SidA, suggests a novel role for NADP$^+$ in the reaction of flavin-dependent monooxygenases, in which the nicotinamide cofactor donates a proton from the 2'-hydroxyl of the nicotinamide ribose to the peroxyflavin to form the hydroperoxyflavin.[84]

Members of the NMO family have been studied less compared to BVMOs and FMOs partly owing to their low affinity for FAD which limits mechanistic studies.[85] In NMO sequences, only the histidine residue is conserved in the sequence region corresponding to the BVMO and FMO sequence motifs.[36, 86] In addition, the Rossmann motifs in NMOs are not the typical GXGXXG as in the other class B FPMOs, but exist as GXG. For example in the case of PvdA, the motif sequence is GXGXXN.[83]

1.4.4 The drawback of class B FPMOs

From an industrial application point of view, a major drawback of class B FPMOs is the dependence of the enzymes on NADPH as nicotinamide cofactor. NADPH is more

expensive than its non-phosphorylated neighbour NADH and using structural analysis of the interactions between the active site and the phosphate of NADPH attempts have been made to investigate and engineer the cofactor preference of class B BVMOs.[87] However, progress in this area has been unsuccessful, indicating that engineering an NADPH dependent enzyme to employ NADH is not a simple task. Recently, NADH-dependent FMO activity has been reported in a group of enzymes from *Rhodococcus jostii* RHA1 in which the enzymes were able to employ either NADH or NADPH as nicotinamide cofactor yielding similar conversion rates and enantioselectivities.[18]

Other than enzyme engineering, the cofactor associated drawback can be eliminated by cofactor recycling systems for both NADPH and NADH such as, glucose-6-phosphate dehydrogenase or formate dehydrogenase which regenerate the cofactor without interfering with the oxygenation reaction.[13] Thus, the cofactor does not have to be employed in stoichiometric amounts.

Another example of reducing cofactor costs is that involving the class E monooxygenase StyA were the reductase component (StyB), NAD and the formate dehydrogenase NADH recycling system was replaced by pentamethyl-cyclopentadienyl rhodium bipyridine $[Cp*Rh(bpy)(H_2O)]^{2+}$. The redox-catalyst receives electrons from the oxidation of formate to carbon dioxide and is able to regenerate reduced FAD directly.[88] In addition the flavin can also be reduced by direct electrochemical reduction using an electrode.[11]

In addition, the use of whole cells rather than isolated enzyme has the advantage of in cell cofactor regeneration.[11]

1.5 Applications of class B FPMOs

Class B FPMOs are capable of preforming a range of oxygenation reactions under mild conditions and producing only water as side product which is of great demand in many industrial processes. The products produced by these enzymes act as important intermediates in enantioselective organic synthesis and drug discovery.

Cyclohexanone monooxygenase is an excellent example of class B FPMOs being used at an industrial scale. In 2005, Hilker and co-workers established that *E. coli* cells expressing CHMO could be used to produce industrial important lactones at kilogram

scale. The study demonstrated the scale up of the Baeyer-Villiger oxidation of racemic bicyclo[3.2.0]hept-2-en-6-one (Figure 1.25) in which high yields were obtained by using a resin-based in situ substrate feeding and product removal methodology, a glycerol feed control and an improved oxygenation device. The method allowed both regioisomeric lactones ((-)-(1*S*, 5*R*) and (-)-(1*R*, 5*S*)) to be obtained in almost enantiopure form (*ee* >98 %) and good yield. The study represented the first Baeyer-Villiger bio-oxidation reaction at kilogram scale and opened the way for further (industrial) upscaling of important monooxygenase reactions.[89]
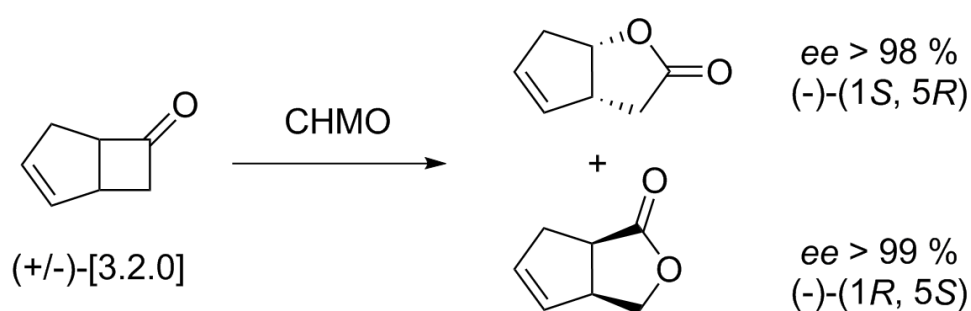
Figure1.25 Whole cell Baeyer-Villiger oxidation of racemic bicyclo[3.2.0]hept-2-en-6-one catalysed by CHMO.

The class B BVMO, CHMO, can also be used in the synthesis of modern drugs, for example esomeprazole. Esomeprazole is a proton pump inhibitor which is prescribed for the treatment of dyspepsia, peptic ulcer disease (PUD), gastroesophageal reflux disease (GORD/GERD) and Zollinger-Ellison syndrome. Esomeprazole is the (*S*)-enantiomer of the racemic omeprazole (Figure 1.26).
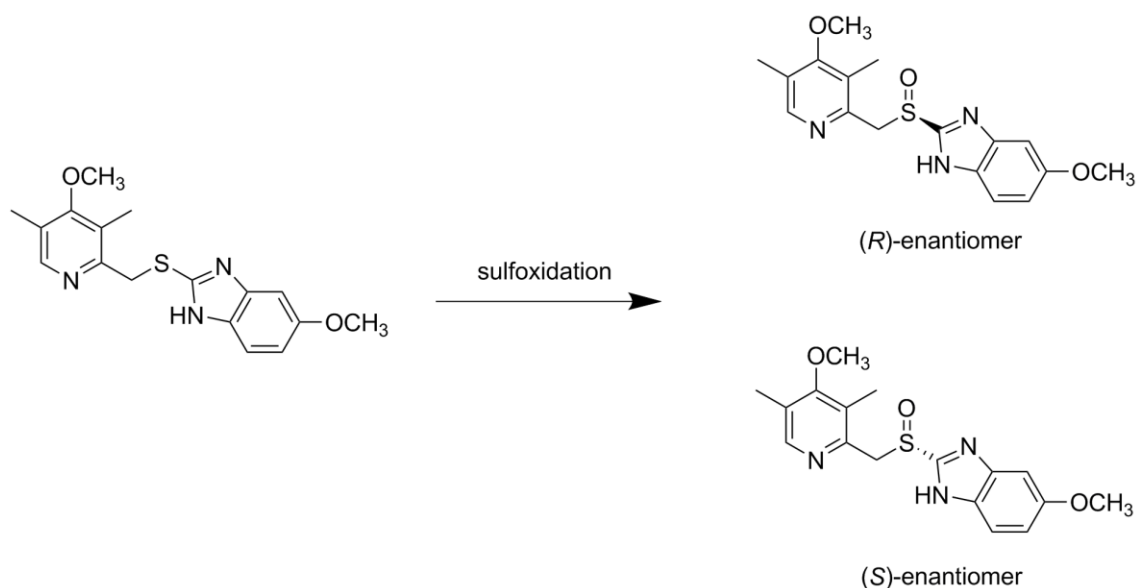
Figure 1.26 Reaction scheme illustrating the sulfoxidation of racemic omeprazole to enantiomeric products.

Typically esomeprazole is synthesised by chemical asymmetric oxidation of sulfides to sulfoxides using a Kagan-Sharpless type oxidation which gives esomeprazole in approximately 94 % enantiomeric excess.[90] In an effort to increase efficiency, enantiopurity and decrease the number of processing steps, a recent patent was released employing mutants of CHMO to produce esomeprazole and other prazole compounds.[91]

Class B FPMOs have great potential for use as biocatalysts in industry. Together with newly developed engineering strategies and structural analysis, selected FPMOs could be engineered for increased efficiency in catalysis by ensuring effective cofactor consumption, oxygen transfer and product recovery and therefore enable more applications for flavoprotein monooxygenases as biocatalysts.

In addition to being highly suitable for biocatalysts, BVMOs can also be of medical relevance. For example, it has been reported that the *etaA* gene of *Mycobacterium tuberculosis* is responsible for activating thiocarbamide prodrugs[92] and mutations in this gene allows *M. tuberculosis* to become drug-resistant. Heterologous expression of EtaA revealed that it represents a BVMO oxidising various ketones.[93] In addition, studies have shown that the efficiency of antitubercular prodrugs can be increased by designing drugs for the upregulation of the *etaA* gene.[94] As BVMOs are widespread within bacteria and no BVMOs are present in human, plant or animal genomes, BVMOs could be a very appealing (pro)drug target.[22c]

Chapter 2: Preliminary Investigation of SMFMO

This thesis focuses on the characterisation of an unusual FMO from the marine bacterium *Stenotrophomonas maltophilia*. Prior to the commencement of this work, preliminary experiments were performed.

Initial screening and preparation of genomic DNA was achieved by Plymouth Marine Laboratories (PML Applications Ltd). Sequencing and assembly of genomic DNA was carried out at the GenePool Genomics and Bioinformatics Facility at the University of Edinburgh. Cloning, expression and purification were performed by Dr Jared Cartwright (Technology Facility, University of York). Initial gas chromatography assays and crystallisation experiments were performed by Dr Gideon Grogan prior to the start of this Ph.D.

2.1 Identification and isolation of SMFMO gene

Stenotrophomonas maltophilia flavoprotein monooxygenase (SMFMO) was encoded within the marine bacterial strain *Stenotrophomonas maltophilia* (PML168), isolated from a rock pool on Wembury Beach. Initial screening of PML168 indicated that the strain had general alkaline phosphomonoesterase, alkaline phosphodiesterase, carboxyesterase, epoxide hydrolase, halocarboxylic acid dehalogenase/dehydrogenase, peroxidase, laccase and lactone hydrolase enzymatic activities.[95] In addition, PML168 was notable for performing Baeyer-Villiger oxidation of the substrate 3-acetyl indole, described in Figure 2.1. This industrially relevant enzymatic activity prompted the analysis of the *Stenotrophomonas maltophilia* genome sequence, which revealed an interesting 357 amino-acid (38.6 kDa) enzyme (SMFMO, UNIPROT B2FRL2). The enzyme sequence contains two Rossman fold motifs (GXGXXG) responsible for the binding of the ADP moiety of the FAD and NADPH[96] and an amino acid motif distinctive of an FMO (**FAGIQLHSAHY**),[11, 36] shown in Figure 2.2.
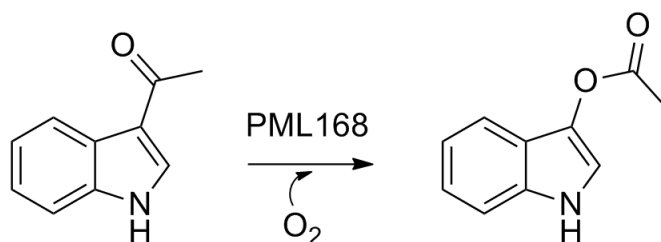
Figure 2.1 Baeyer Villiger oxidation of 3-acetyl indole to yield the corresponding ester. PML168 was notable for performing the oxidation of 3-acetyl indole in the presence of molecular oxygen.

```
SMFMO 357 aa

MDSVDVVVIGGGQSGLSAGYFLRRSGLSYVILDAEASPGGAWQHAWHSLHLFSPAGWSSI
PGWPMPASQGPYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERLRVVARDGRQWL
ARAVISATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAIIGGGNSGAQILA
EVSTVAETTWITQHEPAFLADDVDGRVLFERATERWKAQQEGREPDLPPGGFGDIVMVPP
VLDARARGVLAAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSHLKGLDLVTPQG
QVEVDGSGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTAYCADHQDR
```

Figure 2.2 *Stenotrophomonas maltophilia* (SMFMO) amino acid sequence.
The two Rossman motifs are indicated in green and the 'FMO motif' is indicated in red.

2.2 Cloning, expression and purification

The gene encoding SMFMO, codon optimised for use within *E. coli*, was synthesised by Geneart (Regensburg, Germany). The gene was amplified by PCR and cloned into the pET-YSBLIC-3C vector using the LIC procedure described in Section 3.1.1 − 3.1.5. The resulting recombinant plasmid was employed to transform *E. coli* XL1-Blue cells (Novagen) and the resulting plasmid recovered using miniprep, as described in Section 3.1.6, was used for gene expression (Section 3.2). The expressed 38.6 kDa protein was purified using nickel-affinity chromatography and gel filtration (Section 3.3, Figure 2.3) yielding SMFMO that is bright yellow in colour.
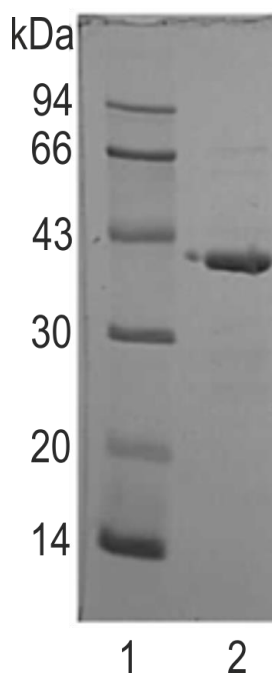
Figure 2.3 SDS-PAGE gel showing pure recombinant SMFMO at 38.6 kDa.
The enzyme was purified using nickel affinity chromatography followed by gel
filtration. Lane 1: Low molecular weight markers from BioRad. Lane 2: SMFMO after
purification.

2.3 Enzyme assays

Preliminary gas chromatography analysis, described in Section 3.4, of SMFMO with the
commercially available substrate **15** demonstrated that the enzyme possessed class B
FPMO-type activity, when supplied with a nicotinamide cofactor NADH or NADPH
(Figure 2.4). Interestingly, the enzyme was able to oxidise the thioether substrate using
either cofactor. A second substrate, **1**, was also employed and revealed that SMFMO
had the ability to catalyse a Baeyer-Villiger oxidation, illustrated in Figure 2.4.
Similarly, activity was observed when using either NADH or NADPH as nicotinamide
cofactor to reduce the flavin.
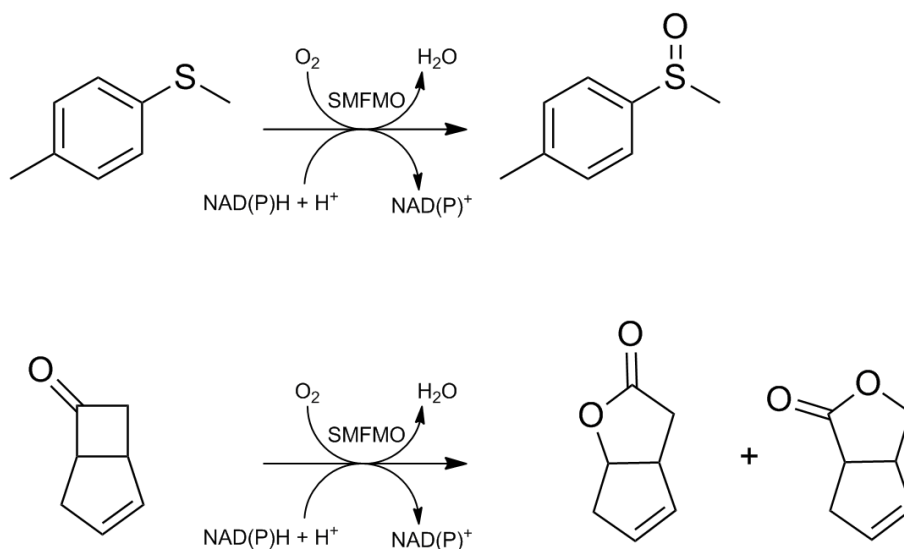
Figure 2.4 Sulfoxidation and Baeyer Villiger oxidation reactions catalysed by SMFMO. Initial assays established that SMFMO had the ability to catalyse the oxidation of **15** and **1** to their corresponding products, in the presence of either NADH or NADPH.

2.4 Crystallisation and structure determination

Crystallisation conditions for SMFMO were established (Section 3.7) and were successful in producing yellow needle like protein crystals (Figure 2.5). The crystals allowed for a 2.7Å dataset to be obtained and the structure to be solved but poor quality of the density maps meant that building of the structure proved difficult.
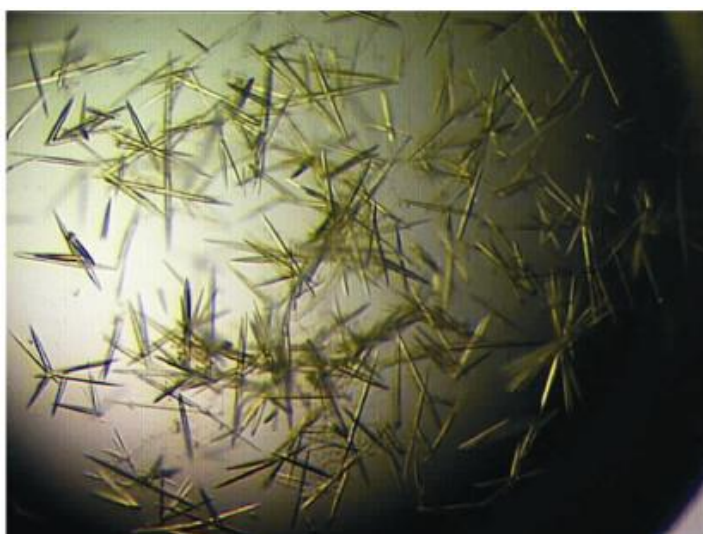
Figure 2.5 Initial yellow SMFMO protein crystals obtained by Gideon Grogan.
Crystals were obtained using conditions containing lithium sulphate (1.8 M) and protein (4 mg mL^{-1}).

2.5 PhD project Aims

The ability for SMFMO to use either NADH or NADPH as nicotinamide cofactor in oxidation reactions gave this enzyme unique characteristics within class B FPMOs, along with its smaller molecular weight. The gene was easily expressed and protein purification was simple, thus, allowing for a complete investigation into the nature of SMFMO and its structure.

The aims of the PhD project were:

1. Express the SMFMO gene and purify the protein to yield pure isolated protein at a high concentration.

2. Characterise the kinetic ability of SMFMO to reduce the flavin using either NADH or NADPH as nicotinamide cofactor. In addition, determine if SMFMO conforms to Michaelis-Menten enzymatic behaviour using UV-spectrometry assays.

3. Characterise the activity of SMFMO and thus the enzymes ability to oxidise substrates in the presence of either nicotinamide cofactor using biochemical techniques such as, gas chromatography (GC) or chiral GC.

4. Determine the substrate specificity of SMFMO by utilising a library of commercially-available ketones to be screened against the target enzyme for activity with both NADH and NADPH.

5. Screen the activity of SMFMO against a range of prochiral sulfides employing either nicotinamide cofactor to reduce the flavin.

6. Determine the structure of SMFMO and characterise the enzyme using structural analysis.

7. Engineer SMFMO to investigate the cofactor promiscuity of the enzyme using mutagenesis techniques such as, site-directed mutagenesis. In addition, obtain a structure of the SMFMO variants for structural analysis.

8. Evolve SMFMO to investigate the binding properties of the active site and improve the enantiomeric excess of the successful reactions using mutagenesis techniques such as, site-saturation mutagenesis.

9. Study closely related enzymes to SMFMO for possible putative NADH-specific FPMOs.

Chapter 3: Methods

All chemicals were purchased from Sigma-Aldrich. The SMFMO gene was a gift from Jared Cartwright (Technology Facility, University of York) and all other genes were synthesised by GeneArt. The pET-YSBLIC-3C plasmid was obtained from a YSBL stock. All primers were synthesised by MWG Biotech. Competent cell strains were purchased from Novagen. All plasmids obtained from cloning and mutagenesis were sequenced by the Genomics Laboratory (Technology Facility) and GATC Biotech.

3.1 Gene Cloning

3.1.1 Polymerase chain reaction (PCR)

A polymerase chain reaction (PCR) is generally utilised to increase the amount of DNA for a specific gene of interest in a short amount of time. The ability DNA has to reversibly unwind from double to single stranded DNA at increased temperatures is exploited within the PCR. Forward and reverse oligonucleotide primers are employed, which are designed to anneal to the coding and non-coding strands of the separated template DNA.

A PCR is carried out using a thermal cycler which mediates the rapid rise and fall in temperature needed throughout the reaction. The reaction mixture is initially heated to $95^{\circ}C$ allowing the template DNA to essentially melt from double to single strands. The reaction temperature is then lowered to ~$50^{\circ}C$ enabling the primers to anneal to the single strands of the template. The 'annealing temperature' is usually lower than the melting temperature (Tm) of the primers. The reaction mixture is then heated to $72^{\circ}C$, the 'extension temperature', at which DNA polymerase is at its optimum activity. The DNA polymerase uses the deoxynucleotide triphosphate monomeric units (dNTPs) present in the reaction mixture to synthesise the new strands of DNA in the 5´ to 3´ direction. Newly synthesised DNA can then be used as template DNA within the PCR. This melt-anneal-extend cycle is repeated approximately 30 times in order to amplify the gene of interest. One cycle is shown in Figure 3.1.
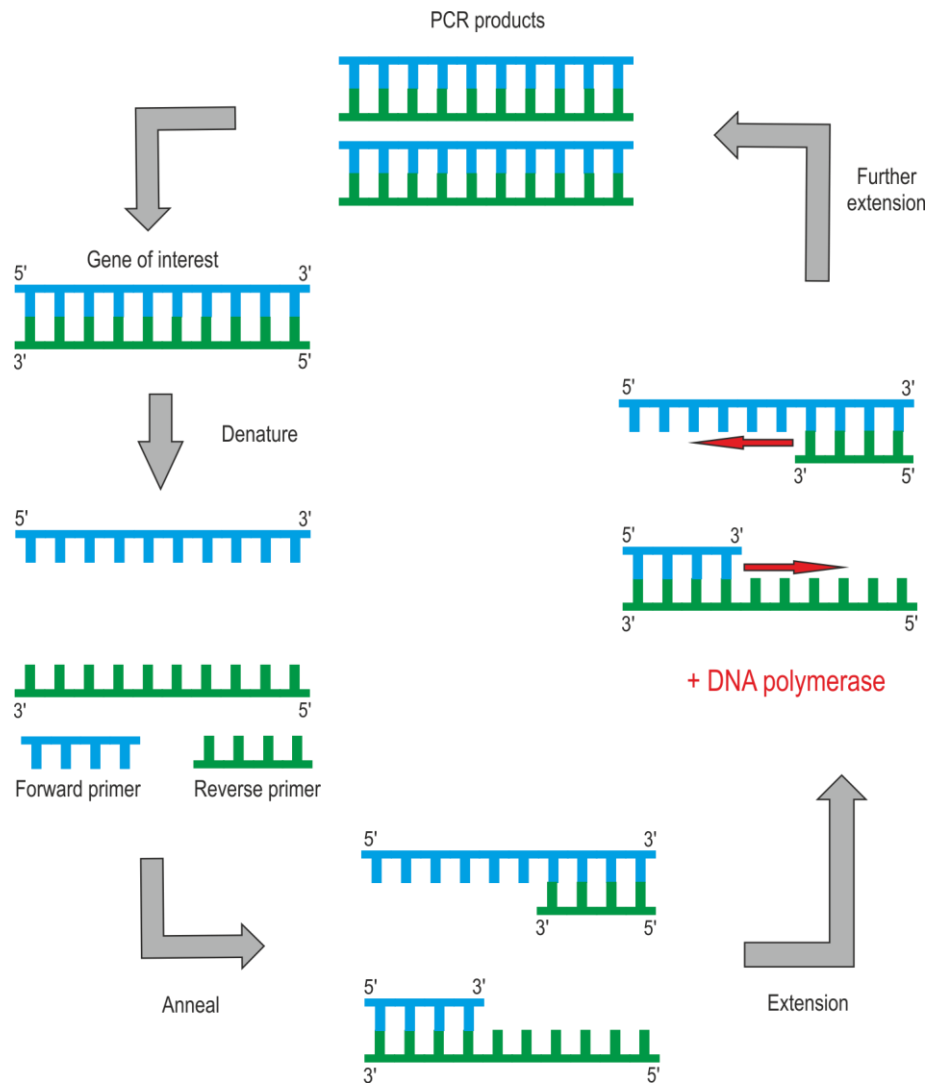
Figure 3.1: Diagram of the melt-anneal-extend cycle within the polymerase chain reaction (PCR).

PCR Method

Unless stated all PCRs were carried out using the following method. To a thin walled Eppendorf tube (500 µL) the components of the PCR (Table 3.1) were added. Milli-Q water was added to give a total volume of 50 µL. The DNA polymerase was added last and the reaction mixture was thoroughly mixed with a pipette.

Table 3.1 Components of PCR reaction used to amplify target genes.

Component	Component concentration	Volume /μL
DNA template	25 ng μL^{-1}	1
Forward primer	20 μM	1
Reverse primer	20 μM	1
dNTPs	2 mM	5
KOD (Hot Start) polymerase buffer		5
DMSO		1
MgSO$_4$	25 mM	3
KOD (Hot Start) polymerase		1

The PCRs were placed in the thermal cycler and ran using the programme described in Table 3.2

Table 3.2 PCR programme used to amplify target genes.

Step	Temperature /^{o}C	Time /s	Cycles
Initial denature	94	120	1
Denature	94	30	
Anneal	50	30	35
Extend	72	30	
Final extend	72	180	1
Hold	4	-	-

3.1.2 Ligation independent cloning (LIC)

Ligation-independent cloning is a method that is independent of ligase activity and allows complementary strands of DNA to anneal spontaneously.[6] Within LIC special nucleotide ends known as 'LIC ends' are added to the primers used in the gene of interest (G.O.I) amplification step. The primers are particularly easy to design as no restriction sites need to be incorporated and only the LIC extensions need to be added. After the gene of interest is amplified it is cleaned and incubated with T4 DNA polymerase and dATP. The polymerase acts as a 3´-5´ exonuclease, in which it digests nucleotides from the 3´ends of the DNA sequence until it reaches an 'A' base.[97] At this point the polymerase ceases digestion exposing a single strand on each end of the gene.

Similarly, the LIC vector is digested with a restriction endonuclease (BseRI) which yields a linear plasmid with two base pair overhangs. The product is then treated with T4 DNA polymerase and dTTP. Once again, the polymerase digests nucleotides from the 3´ends until a 'T' base is reached. This allows the gene and vector to have 5´-single stranded overhangs that are complementary to one another. Incubation of the gene and vector allows for the spontaneous annealing of the complementary 'sticky ends', thus, resulting in the desired recombinant plasmid (Figure 3.2).
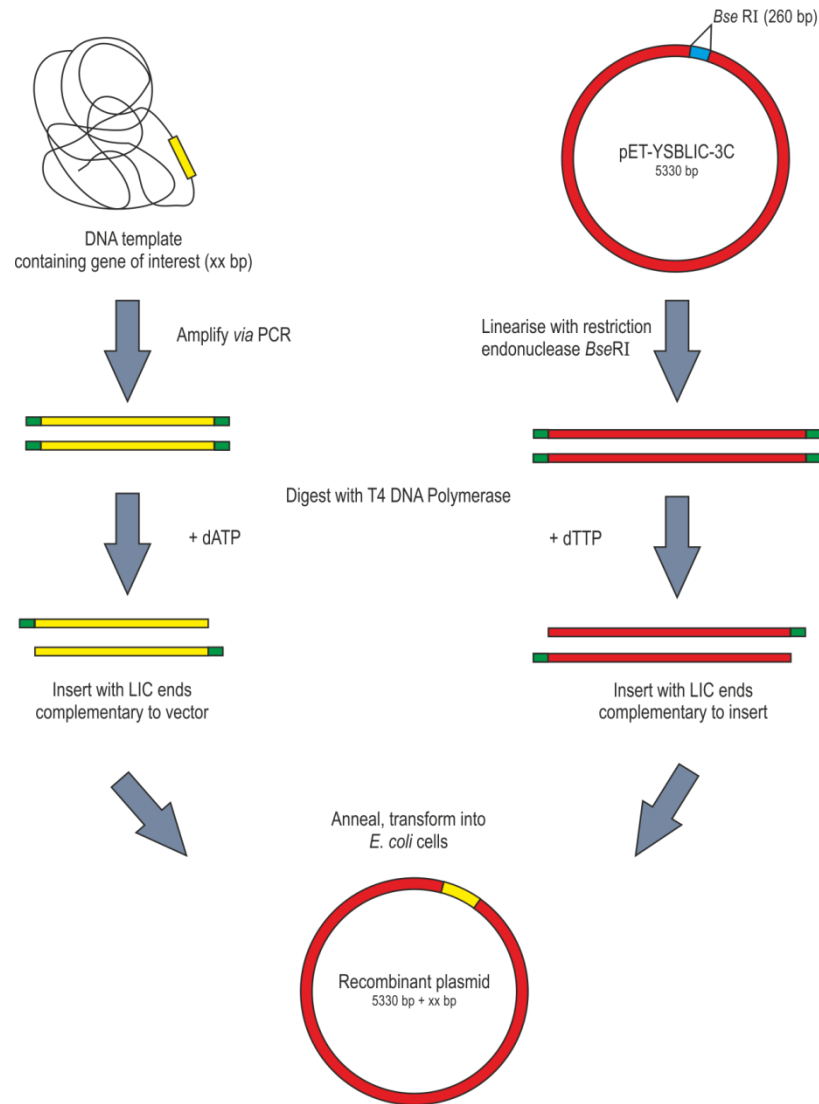
Figure 3.2 Diagram outlining the ligation-independent cloning (LIC) process.
The gene of interest is amplified with LIC primers by PCR. Both gene and linearised vector are incubated with T4 DNA polymerase and corresponding dATP/dTTP, yielding complementary overhangs which anneal spontaneously at room temperature to give the desired recombinant plasmid.

The pET-YSBLIC 3C vector

The pET-YSBLIC vector is derived from the pET-28a vector (Novagen), which was designed to allow the fast production of any protein. The pET-28a plasmid contains a kanamycin antibiotic resistance marker, cloning site, *lacI* gene which encodes the *lac* repressor protein, T7*lac* promoter and a *lac* operator used to block transcription. Within the cloning site an N-terminal His$_6$ tag is encoded and a human rhinovirus 3C protease cleaving site. For the pET-YSBLIC vector the cloning site has been altered to include a

LIC site with a BseRI cleavage site. The T7*lac* promoter and *lac* operator control the expression of the gene within the plasmid.

The desired gene must be expressed in a strain of *E. coli* (DE3) that has been altered to encode the gene for T7 RNA polymerase, which is under the control of the *lacUV5* promoter and *lac* operator system switching on induction when lactose is present. Isopropyl *β*-1-thiogalactopyranoside (IPTG) is a non-metabolisable allolactose analogue and is used in place of lactose to give overexpression of recombinant enzymes (Figure 3.3).
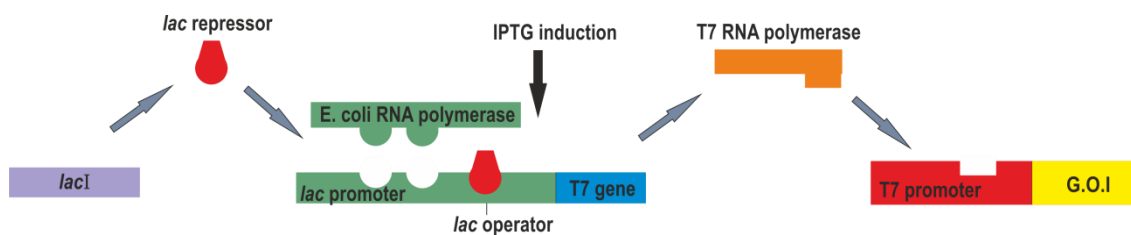
Figure 3.3 Diagram illustrating IPTG induced gene expression in the pET system. IPTG binds to the *lac* repressor and is released from the *lac* operator. *E. coli* RNA polymerase can then bind to the *lac* promoter initiating the production of T7 polymerase, resulting in the subsequent expression of the target protein.

If IPTG is absent from the growth medium, the primary carbon source is glucose and the *lac* repressor is bound to the *lac* operator, therefore transcription of the T7 polymerase gene and the desired gene is blocked. In the presence of IPTG, the *lac* repressor is released from the *lac* operator in the *E. coli* genome. *E. coli* RNA polymerase can now bind to the *lac* promoter and catalyse the production of T7 RNA polymerase. The T7 polymerase can then bind to the T7 promoter within the pET plasmid, thus initiating the transcription and translation of the desired protein.

3.1.3 Agarose gel electrophoresis

Preparation of LIC insert

The desired gene was amplified using PCR as described in section 3.1.1.

The PCR products were purified by agarose gel electrophoresis. The bands corresponding to the amplified gene were cut out and a GenElute™ gel extraction kit (Sigma-Aldrich™) was used to recover the amplified gene of interest.

Preparation of 1 % agarose gel

To agarose (0.6 g) $1 \times$ TAE buffer (60 mL) was added and heated in a microwave for 70 s until all the solid had dissolved. Once cool, *Sybr® Safe* DNA stain (2 µL, 10,000 x concentrated in DMSO, Invitrogen) was added and swirled to mix. The mixture was then poured into an agarose cast and the desired comb size was inserted. The gel was then allowed to set for 1 h.

Table 3.3 Agarose gel electrophoresis buffers

Buffer	Component	Concentration
TAE buffer pH 8.0	Tris	40 mM
	Acetic acid	20 mM
	EDTA	1 mM

Sample preparation

DNA samples were mixed with $6\times$ agarose gel loading buffer (New England Biolabs®) in a 1:5 ratio and loaded on to the gel. A DNA ladder was prepared by mixing 1 kb DNA ladder (1.5 µL) with 6 x loading dye (1 µL) and water (4 µL).

Agarose gel running conditions

The gel was run at 100 V, 400 mA for 1 h. After the run the gel was removed from the tank and placed under a transilluminator or Safe Imager for analysis.

3.1.4 LIC T4 polymerase reactions

Insert LIC T4 pol reaction

To a thin walled Eppendorf tube (0.5 mL) the following components were added:

Table 3.4 Components for LIC T4 polymerase reactions

Component	Concentration	Volume /μL
DNA insert	0.2 *p*mol	5
10 × T4 buffer		2
dATP	25mM	2
DTT	100 mM	0.5
T4 DNA polymerase		0.5

Water was added to make the final volume of the reaction 20 μL. The reagents were mixed and incubated first at 22°C for 30min followed by incubation at 75°C for 20 min using a PCR machine.

3.1.5 LIC annealing reaction

Pure T4 treated linearised pET-YSBLIC-3C vector was sourced from Laila Roper within the Grogan group.

To an Eppendorf tube the previously T4 treated insert (3 μL) and pure T4 treated linearised pET-YSBLIC-3C vector (1.5μL, 4.0 *p*mol) were added and the mixture was incubated at room temperature for 10 min. EDTA (1.5 μL, 25 mM) was added and the mixture was incubated for a further 10 min.

3.1.6 Transformation and miniprep

E. coli XL10 Gold Ultra-competent cells were gently thawed on ice. Cells (75 μL) were aliquoted into a pre-cooled Eppendorf tube (0.5 mL) along with β-mercaptoethanol (4 μL) and the mixture was swirled every 2 min for 10 min. The LIC annealing reaction (6

µL) was added and the sample was left on ice for 30 min. The sample was subjected to 'heat shock' at 42°C for 45 s and returned back to ice. LB (500 µL) was added and the sample was incubated at 37°C for 2 h. After incubation, the mixture was centrifuged (3 min, 5000 rpm) and the pellet was resuspended in supernatant (300 µL). The mixture was then spread between two kanamycin (30 µg mL^{-1}) LB agar plates and incubated at 37°C overnight.

A colony from the previous plates was picked and placed in a sterile vial containing LB medium (5 mL) and kanamycin (30µg mL^{-1}). The mixture was incubated overnight at 37°C with shaking. A GenElute™ plasmid Miniprep kit (Sigma-Aldrich®) was used to retrieve the DNA.

3.1.7 Restriction Digest

To an Eppendorf tube (0.5 mL) mini-prep DNA (7 µL), NcoI (NEB) (1 µL), NdeI (NEB) (1 µL) and 10 × NEB buffer 2 (1 µL) were added and incubated at 37°C for 2 h. Each reaction mixture (10 µL) was loaded onto a 1 % agarose gel and analysed (section 3.1.3).

3.2 Protein Production

3.2.1 Transformation with SMFMO plasmid

The pET-YSBLIC-3C vector containing the B2FRL2 gene (1 µL) was added to *E. coli* competent cell strain BL21 (DE3) (25 µL) and placed on ice for 20 min. The mixture was placed in a water bath (42°C) and heat shocked for 45 s before resting on ice for a further 5 min. LB media (1 mL) and incubated at 37°C, 90 min. After incubation, the mixture was centrifuged (3 min, 5000 rpm) and the pellet was resuspended in supernatant (300 µL). The mixture was then spread between two kanamycin (30 µg mL^{-1}) LB agar plates and incubated at 37°C overnight.

3.2.2 Preparation of starter culture

A colony from the previous transformations was picked and placed a sterile vial containing LB medium (5 mL) and kanamycin (30 µg mL^{-1}) before incubating at 37°C, 180 rpm overnight.

3.2.3 Small scale expression tests

Starter culture (50 µL) was used to inoculate LB medium (3 × 5 mL) containing kanamycin (30 µg mL^{-1}). The mixture was incubated at 37°C until OD$_{600nm}$ = 0.6. The cultures were induced with IPTG and incubated separately at 16°C, 30°C and 37°C overnight with shaking. Controls were also set up in the same way but without the induction of IPTG.

Analysis of soluble and insoluble fractions

After incubation the cell cultures were centrifuged (10 min, 5000 rpm) and the supernatant was discarded. The resulting cell pellet was resuspended in Milli-Q water (1 mL). The cells were lysed by sonication (3 × 30 s) and centrifuged (10 min, 5000 rpm). The supernatant (soluble fraction) was decanted into a clean Eppendorf tube and the pellet (insoluble fraction) was resuspended in Milli-Q water (500 µL). Both soluble and insoluble fractions were analysed by SDS-PAGE.

3.2.4 Sodium dodecyl sulphate-polyacrylamide electrophoresis (SDS-PAGE)

12% SDS-PAGE gel preparation

The 12% resolving gel was prepared using sterile de-ionised water (3.2 mL), resolving buffer (2.5 mL) and acrylamide (4.5 mL). APS (50 µL, 10 % w/v) and TEMED (8.5 µL) were added before mixing and pouring into a gel frame. A small amount of butan-1-ol was added in order to even out the top of the gel and remove any bubbles. The gel was left to set for 15 min. Once set, the butan-1-ol was removed. The stacking gel was prepared by mixing deionised water (3.2 mL), stacking buffer (1.3 mL), Acrylamide (0.5 mL) and bromophenol blue (10 µL). APS (25 µL, 10 % w/v) and TEMED (8.5 µL) were then added and mixed thoroughly before pouring into the gel frame. A plastic gel

comb was added immediately to the stacking solution and allowed to set for 15 min. Solutions used are shown in Table 3.5.

Sample preparation

Samples were prepared by mixing protein sample (10 µL) with SDS-PAGE loading buffer (10 µL). The samples were heated to 95°C for 5 min. A Bio-Rad low molecular weight marker was also prepared.

Running of SDS-PAGE gel

The gel was removed from the frames and placed in a gel tank. The tank was filled with running buffer and the prepared samples and markers were loaded on to the gel using a needled syringe. The gel was run for 55 min at 200 V.

Staining of Gels

Once run, the gel was removed from the gel tank and placed in a plastic tub. Coomassie blue stain was added to the tub until the gel was covered and heated in a microwave on high for 40 s. The stain was decanted and the gel was rinsed lightly with water.

Destaining of Gels

Destain was added to the tub until the gel was covered and heated in a microwave on high power for 40 s. A tissue was then folded and placed on top of the gel and the tub was covered with a lid. The tub was placed on a rocker overnight and after destaining the gel was photographed and stored in water.

Table 3.5 Solutions for SDS-PAGE

Solution	Component	Concentration
Resolving buffer pH 8.8	Tris	1.5 M
	SDS	0.1 % w/v
Stacking Buffer pH 6.8	Tris	0.5 M
	SDS	0.4 % w/v
SDS-PAGE loading buffer	Tris-HCl (pH 8.0)	60 mM
	Glycerol	10 % w/v
	SDS	2 % w/v
	Bromophenol blue	0.02 % w/v
	β-mercaptoethanol	10 % w/v
Running buffer	Tris	25 mM
	Glycine	200 mM
	SDS	0.1 % w/v

3.2.5 Large scale gene expression

Transformation and starter cultures were prepared as previously described (Section 3.2.1, Section 3.2.2).

Lysogeny broth medium preparation (LB)

LB was prepared by dissolving the following components (Table 3.6) in distilled water (1 L). The media was transferred to 2 L baffled Erlenmeyer flasks and autoclaved.

Table 3.6 Components for lysogeny broth (LB) medium

Component	Concentration /g L^{-1}
Sodium Chloride	10
Yeast extract	5
Tryptone	10

Gene expression

Starter culture (5 mL) was added to LB media (500 mL) containing kanamycin (500 μL) and incubated at 37°C, with shaking at 180 rpm. Once the OD_{600nm} reached approx. 0.6 the cultures were induced with IPTG (1 mM) and incubated overnight at 16°C, with shaking at 180 rpm. The following day the cells were collected by centrifugation (15 min, 5000 rpm) and the supernatant was discarded. The cell pellet was resuspended in Buffer A (100 mL).

3.3 Protein purification

General buffers used throughout protein purification are listed in Table 3.7. All buffers were filtered and degassed before use.

Table 3.7 Buffer used throughout protein purification

Buffer	Component	Concentration /mM L^{-1}
A pH 7.0	Tris/HCl	50
	NaCl	300
B pH 7.0	Tris/HCl	50
	NaCl	300
	Imidazole	30
C pH 7.0	Tris/HCl	50
	NaCl	300
	Imidazole	500

3.3.1 Cell lysis by sonication

Cell pellets derived from 2 L of culture were resuspended in buffer A (100 mL) and transferred to a 200 mL glass beaker. The cells were then sonicated on ice (3 × 30 s with 30 s intervals). The sonicated extract was centrifuged (35 min, 18,000 rpm) and the pellet was discarded. The supernatant was -filtered using a Millex syringe filter (0.45 µm pore size followed by 0.22 µm) into a sterile pot and the filtered protein was kept on ice.

3.3.2 Nickel (Ni^{2+}) affinity chromatography

Expression using the pET-YSBLIC-3C vector gives rise to a N-terminal His$_6$ tag, which can be used to purify a protein using Ni^{2+} affinity chromatography. Polymer beads within the column are charged with nickel ions and the crude protein extract is loaded.

The His$_6$ tagged protein will chelate the nickel ions, resulting in the binding of the protein to the charged beads. Other proteins, which are not hexahistidine tagged, are eluted without binding. After washing with buffer, the bound protein is eluted with an increasing concentration of imidazole.

Preparation of Ni^{2+} column

The Ni^{2+} column (5 mL, HisTrap column, GE Healthcare) was prepared by following the series of steps in Table 3.8.

Table 3.8 Steps for preparation of Ni^{2+} column

Step	Solvent	Volume /mL
1	dH$_2$O	20
2	EDTA (0.1 M) with NaCl (0.5 M)	20
3	NaCl (0.5 M)	20
4	dH$_2$O	20
5	NiSO$_4$	20
6	dH$_2$O	20
7	Buffer A	20

The crude protein extract was loaded onto the prepared column followed by buffer A (20 mL).

AKTA-FPLC

The AKTA-FPLC pumps were washed with H$_2$O and the Ni^{2+} column was attached. The bound protein was eluted with an increasing gradient of imidazole, from 30 mM to 500 mM (0 - 100 % Buffer C). The FPLC was programmed to collect 4 mL fractions at

a flow rate of 1 mL min^{-1} until 150 mL of gradient buffer had been collected. Fractions determined to contain protein according to the FPLC chromatogram were analysed by SDS-PAGE (Section 3.2.4). The protein fractions were combined and centrifuged (5000 rpm) using a Centricon® concentrator tube (10 kDa MWCO) until the volume reached ~ 2 mL.

3.3.3 Size exclusion chromatography (SEC)

After nickel affinity chromatography further purification was needed in order to obtain pure protein free of nonspecific aggregates. SEC exploits columns packed with inert polymer beads that contain small pores. Once a protein is loaded onto the column it is eluted using an uncharged buffer. Large proteins are unable to enter the small pores, and must travel around the beads and therefore elute first. Smaller proteins enter the pores which hinder their route through the column and thus they are eluted last.

A Superdex® 75 gel filtration column (16 mm × 60 cm) from GE Healthcare was employed for all purifications.

Preparation of gel column

The column was equilibrated overnight with filtered H_2O (120 mL) followed by buffer A (120 mL).

Sample loading and gel filtration

A protein injection loop (2 mL) was attached to the AKTA-FPLC and washed with H_2O (10 mL) and buffer A (10 mL) using a syringe. The Ni^{2+} purified protein (~ 2mL) was injected on to the loop and loaded onto the column. The FPLC was instructed to collect 4 mL fractions at a flow rate of 0.5 mL min^{-1} until 120 mL of buffer A had been collected. The fractions displaying high UV absorbance at 280 nm were analysed by SDS-PAGE (Section 3.2.4).

Protein concentration

The fractions containing pure protein as identified by the SDS-PAGE analysis were combined and concentrated using a Centricon® concentrator tube (10 kDa MWCO) and spun in a centrifuge (5000 rpm) until the concentration of protein reached 3 mg mL^{-1}. The concentration of protein was determined using a UV spectrophotometer measuring absorbance. Using the Beer- Lambert equation A= ε.c.l where ε is the absorption coefficient (estimated using the EXPasy ProtoParam tool) and l is 1 cm the concentration of protein can be calculated. The purified protein was snap frozen using liquid nitrogen (500 μL aliquots) and stored at -80°C for later use.

3.3.4 His$_6$-tag cleavage

The pET-YSBLIC-3C vector contains a sequence that encodes a protease cleavage site (Leu-Glu-Val-Leu-Phe-Gln/Gly-Pro-Ala) which is recognised by the human rhinovirus 3C protease (HRV 3C). The HRV 3C protease can be used to remove a protein histadine tag.

Protease digestion

Protein (0.5 mL, 3 mg mL^{-1}) was diluted with buffer A to a concentration of 1 mg mL^{-1}.

HRV 3C protease (3.5 μL, 1 mg per 100 mg of protein) was added to the protein and mixed. The mixture was stored at 4°C overnight.

The following day samples of the cleaved and uncleaved protein were analysed by SDS-PAGE (Section 3.2.4).

The cleaved protein was purified using nickel affinity chromatography (Section 3.3.2).

3.4 Enzyme assays

Isolated protein was used in all activity assays, unless otherwise stated. For spectrophotometric assays a GCB-DBUV spectrophotometer (Cintra10) was used. Kinetic constants were determined using the method employed for MtmOIV by Röhr and co-workers.[35a]

3.4.1 Spectrophotometric assays

UltraViolet-Visible (UV-Vis) spectrophotometry is exploited within analytical chemistry to determine characteristics of different solutions ranging from transition metal ions to biological macromolecules.

During a UV-Vis experiment a UV-Vis spectrophotometer is utilized and the intensity of light passing through the sample is measured. This is then compared to the initial intensity of light before passing through a sample and the observed value is known as the transmittance (%T), where $T = \frac{I}{I_0}$. Absorbance, A, is the usual measurement obtained from a UV-Vis experiment and is derived from the transmittance of a particular sample:

$$A = -\log\left(\frac{\%T}{100\%}\right)$$

The Beer-Lambert law states that the absorbance of a solution is directly proportional to the concentration of the absorbing species and the path length.

$$A = \log_{10}\left(\frac{I_0}{I}\right) = \varepsilon.\, C.\, l$$

A = measured absorbance

I_0 = Intensity of incident light at a given wavelength (λ)

I = Transmitted intensity

l = Path length

C = Concentration of sample

ε = Molar absorption or extinction coefficient for a given solvent, temperature and pressure.

For a given path length, the Beer-Lambert law can be used to determine the concentration of the absorbing species in a particular solution.

Using UV-Vis spectroscopy, time courses can be carried out in which the absorbance is measured every second throughout a reaction. This is extremely useful for determining kinetic data for enzymes as the concentration between two different time points can be determined.

Enzyme kinetics

The introduction of substrate into an enzyme's active site allows for the formation of an enzyme-substrate complex. The enzyme then catalyses the chemical reaction through an array of transition states and an enzyme-product complex is formed. The complex dissociates to give free enzyme and product. Other substrate molecules can then bind and the catalytic process can continue.

$$E + S \xrightarrow{k_1} [[ES] \rightarrow [ES]^* \rightarrow [EP]] \xrightarrow{k_2} E + P$$

At some point within an enzymatic reaction, all the enzymes active sites will be occupied with substrate. In other words, the enzymes active sites are saturated and the enzyme is working at its maximum velocity or V_{max}. Thus, V_{max} has the units of moles or substrate converted (product evolved) per unit time.

For an enzyme with a certain substrate, the maximal rate (V_{max}) allows for the determination of the *Michaelis-Menten constant* or K_M, where, K_M is the concentration of substrate at which the enzyme catalyses its reaction at half V_{max}. The K_M value can be taken as an approximate indication for how tightly the enzyme binds its substrate. A low number implies that the enzyme has a greater affinity for its substrate, whilst, a high number represents a lower affinity.

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

$$V_{max} = k_{cat}[E_0]$$

Another kinetic term that is encountered within Michaelis-Menten kinetics is k_{cat}. k_{cat} is the *turnover number* of an enzyme, corresponding to the number of molecules turned over by one molecule of enzyme per second, which gives it the units of s^{-1}. In addition, k_{cat}/K_M, known as the *specificity constant*, is often used to give the most accurate measure of catalytic efficiency. The *specificity constant* can be used to compare the catalytic properties of two enzymes, such as mutants or the effect of two different substrates on the enzyme.

3.4.2 NADH/NADPH oxidation assays

Buffer A was used as a blank and the instrument was set to a time scan (10 min). To a quartz cuvette (1 mL), buffer A (Table 3.7, 50 µmol) and varying concentrations of NAD(P)H (1 mM stock) were added and placed within the instrument and the run started. After 50 s, enzyme was added (3.9 nmol) and the decrease in absorbance at 340 nm was monitored. The concentrations of cofactor are described in Table 3.9.

Table 3.9 Components of NADH/NADPH reduction assays

Run	NADH / µM	NADPH /µM
1	2.5	-
2	5	-
3	10	10
4	20	20
5	40	40
6	60	60
7	80	80
8	100	100
9	120	120

3.4.3 Bicyclo[3.2.0]hept-2-en-6-one (substrate 1) oxidation assays

Buffer A was used as a blank and the instrument was set to a time scan (10 min) at a wavelength of 340 nm. To the cuvette (1 mL), buffer A (50 µmol), NAD(P)H (250 µM) and varying concentrations of bicyclo[3.2.0.]hept-2-en-6-one were added and placed within the instrument and the run started. After 50 s, enzyme was added (3.9 nmol) and the decrease in absorbance at 340 nm was monitored.[87b] A repeat was taken for every

run. The concentrations for substrate for NADH assays and NADPH assays are described in Table 3.10 and 3.11.

Table 3.10 Bicyclo[3.2.0]hept-2-en-6-one concentrations (NADH)

Run	Substrate /mM
1	0.010
2	0.025
3	0.050
4	0.075
5	0.100
6	0.150
7	0.200
8	0.250

Table 3.11 Bicyclo[3.2.0]hept-2-en-6-one concentrations (NADPH)

Run	Substrate /mM
1	0.5
2	2
3	6
4	10
5	18
6	20

3.4.4 Biotransformations

For each reaction the appropriate enzymatic system was used to recycle cofactor. For NADH, formate dehydrogenase in the presence of sodium formate was utilised. For NADPH, glucose-6-phosphate dehydrogenase in the presence of glucose-6-phosphate was employed as the recycling system.[13] Buffer A used throughout is described in Table 3.7. The substrates employed for reactions are shown in figure 3.4.
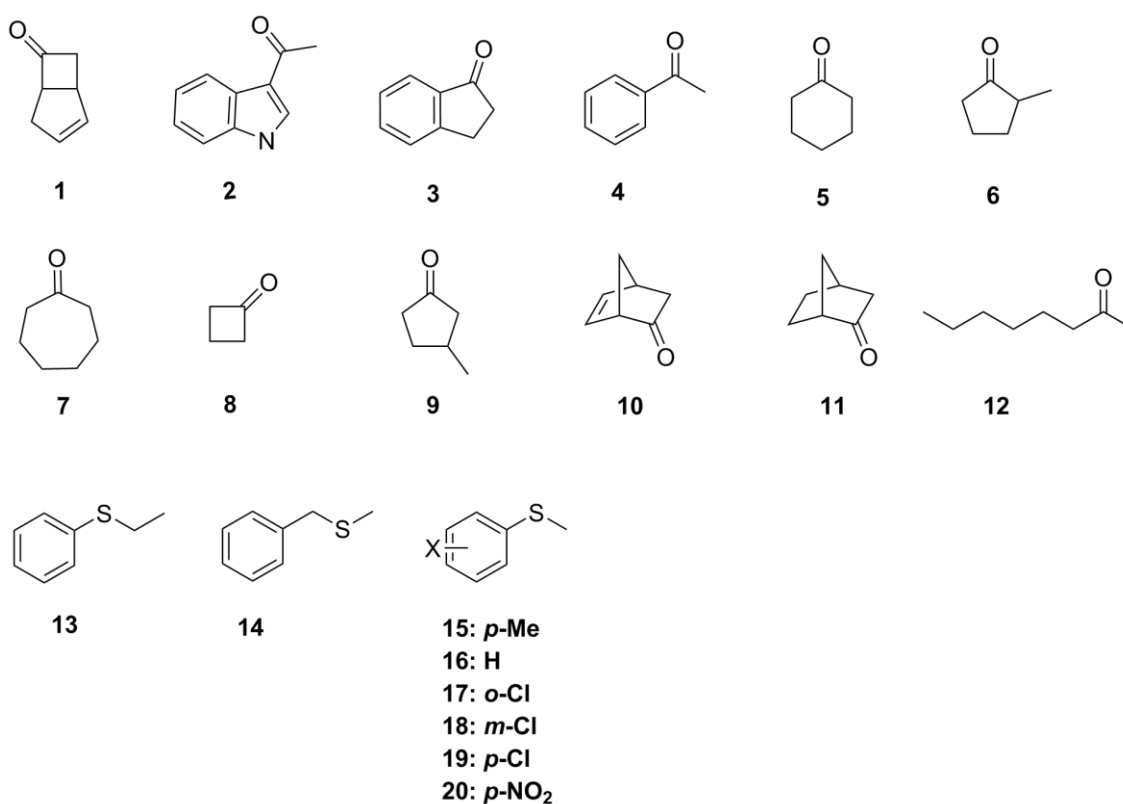
Figure 3.4 Test substrates used in SMFMO characterisation.

For substrates **1**, **2**, **3**, **6**, **8**, **9**, **15** the following reactions were carried out.

NADH

To a round bottomed flask (10 mL) containing buffer A (5 mL) substrate (5 mg in 100 µL of EtOH), NADH (5 mg), formate dehydrogenase (5 mg), sodium formate (6.8 mg) and enzyme (1 mL at 5 mg mL^{-1}) were added and stirred for 24 h at room temperature (r.t).

NADPH

To a round bottomed flask (10 mL) containing buffer A (5 mL) substrate (5 mg in 100 μL of EtOH), NADPH (5 mg), glucose-6-phosphate-dehydrogenase (0.14 mg), glucose-6-phosphate (0.52 mg) and enzyme (1 mL at 5 mg mL^{-1}) were added and stirred for 24 h at r.t.

For each cofactor experiment, aliquots (500 μL) were taken at time points t=0, t=0.5, t=1, t=2, t=4, t=6, t=8 and t=24 h and the organic layer was extracted using ethyl acetate (EtOAc) (500 μL). The organic layer was transferred to a glass screw top vial (2 mL, Agilent Technologies) and stored at 4°C before analysis by GC.

For substrates **4**, **5**, **7**, **10**, **11**, **13**, **14**, **16 - 20** the following reactions were employed.

NADH

To a screw cap glass vial (5 mL) containing buffer A (1 mL) substrate (1 mg in 50 μL of EtOH), NADH (1 mg), formate dehydrogenase (1 mg), sodium formate (1.2 mg) and enzyme (200 μL at 5 mg mL^{-1}) were added and placed in an incubator (r.t.) with shaking for 24 h.

NADPH

To a screw cap glass vial (5 mL) containing buffer A (1 mL) substrate (1 mg in 50 μL of EtOH), NADPH (1 mg), glucose-6-phosphate-dehydrogenase (0.14 mg), glucose-6-phosphate (0.52 mg) and enzyme (200 μL at 5 mg mL^{-1}) were added and placed in an incubator (r.t.) with shaking for 24 h.

For each cofactor, a sample after 24 h (500 μL) was taken and the organic layer was extracted using EtOAc (500 μL). The organic layer was transferred to a glass screw top vial (2 mL, Agilent Technologies) and analysed by GC.

3.4.5 Gas chromatography analysis

For all GC analysis, an Agilent HP-6890 gas chromatograph (Agilent Technologies) was used. The injector temperature was set at 250 $^{\circ}$C and the detector at 320 $^{\circ}$C. Helium was employed as the carrier gas at a pressure of 83 kPa. All sample vials used were screw top vials (2 mL) fitted with a PTFE/red silicone septa screw top (Agilent Technologies).

Achiral GC analysis

The gas chromatograph was fitted with a HP-5 column (30 m $\times$ 0.32 mm $\times$ 0.25 μm). The gradient programmes used for analysing the biotransformation reactions are shown in Table 3.12.

Table 3.12 Gradient programmes preformed for achiral GC analysis

Substrate	Gradient (°C)	Run Time (min)	Retention Time of Substrate (min)	Retention Time of Product (min)
1	130	4	1.80	3.10
2	240	10	2.87	-
3	180	10	2.27	-
4	100 to 250 at 10 °C min^{-1}	15	2.80	-
5	100 to 250 at 10 °C min^{-1}	15	1.90	-
6	80 to 150 at 10 °C min^{-1}	7	1.99	-
7	100 to 250 at 10 °C min^{-1}	15	2.50	-
8	80 to 150 at 10 °C min^{-1}	7	1.99	-
9	80 to 150 at 10 °C min^{-1}	7	2.00	-
10	100 to 250 at 10 °C min^{-1}	15	2.10	3.00
11	100 to 250 at 10 °C min^{-1}	15	2.30	-
13	180	10	1.90	3.00
14	150	10	2.30	5.10
15	150	10	2.40	4.60
16	150	10	2.00	3.30
17	180	10	2.20	3.00
18	180	10	2.20	3.00
19	180	10	2.20	3.00
20	180	10	3.60	-

Chiral gas chromatography

Ketone compound

For substrate **1** columns BGB 173 and BGB 175 (both 30 m × 0.25 mm × 0.25 μm, BGB Analytik) were employed. The enantiomers of the substrate were resolved on the BGB 175 column using a linear temperature gradient of 100 °C to 127 °C at 2 °C min^{-1}. Enantiomers of the lactone products were resolved on the BGB 173 column with a temperature gradient of 90 °C to 34 °C at 1 °C min^{-1}.

Sulfide compounds

For substrates **13**-**19** the BGB 175 column was employed. The gradient programmes used for resolving the enantiomers of the sulfoxide products are described in Table 3.13.

Table 3.13 Gradient programmes for resolving enantiomers of sulfoxide products

Substrate	Gradient /°C	Run Time /min
13	180	30
14	140	155
15	180	20
16	180	30
17	180	30
18	180	30
19	180	30

3.4.6 Gas chromatography-mass spectrometry (GC-MS)

For all GC-MS analysis, an Agilent 7890A GC system fitted with an Agilent 7695 auto-sampler was employed, along with an Agilent 5975C inert XL MSD with triple axis detector. A HP-5 column (30 m × 0.32 mm × 0.25 μm) was also fitted. Helium was

employed as the carrier gas at a pressure of 83 kPa. The gradient programmes for analysing the *m/z* for the product peaks are given in Table 3.14.

Table 3.14 Gradient programmes for analysing the *m/z* for oxidation products

Substrate	Gradient /^{o}C	Run Time /min
1:	100	10
13	150	10
14	120	30
15	120	20
16	120	30
17	150	10
18	150	10
19	150	10

3.4.7 Cyclohexanone monooxygenase (CHMO) –Assignment of configuration.

Absolute configurations of sulfide products were assigned through comparison with products of biotransformations using cyclohexanone monooxygenase (CHMO), along with results reported for the enzyme.[23a, 98]

Transformation with CHMO plasmid

Transformation was carried out according to Section 3.2.1. *E. coli* B834 (DE3) cells were transformed with pMMO4 plasmid which was a gift from M. Mihovilovic. Ampicillin (100 µg ml^{-1}) was employed as antibiotic throughout CHMO expression.

Gene expression

Starter culture and protein production was carried out in accordance to Section 3.2.5.

Biotransformation of sulfide library

The filtered supernatant was concentrated using a Centricon® concentrator tube (10K) and spun in a centrifuge (4500 rpm, 15 min) until the volume reached a third of the original volume. To a screw cap glass vial (5 mL), NADPH (1 mg), glucose-6-phosphate-dehydrogenase (0.14 mg), glucose-6-phosphate (0.52 mg), sulfide (1 mg in 50 µL of EtOH) and cell lysate (1 mL) were added and placed in an incubator (30 $^{\circ}$C) with shaking for 24 h. A t=24 sample (500 µL) was taken and the organic layer was extracted using ethyl acetate (500 µL). The organic layer was transferred to a GC vial (2 mL) and ran on the BGB 175 column using the gradient programmes for of the sulfides previously described (Table 3.13).

3.5 Mutagenesis

3.5.1 Site-directed mutagenesis (SDM)

Site-directed mutagenesis (SDM) is a tool which enables the active sites and mechanistic pathways of enzymes to be investigated rationally. For example, if the hydrophobic properties of a phenylalanine were to be investigated it could be mutated into tyrosine. In a SDM reaction, template DNA (recombinant plasmid containing the target gene) is utilised. In addition, oligonucleotide primers containing the desired mutation, dNTPs and polymerase are also required. During the SDM PCR, the template DNA is denatured or melted and the designed mutant primers anneal to the gene. The polymerase uses the dNTPs to amplify the whole plasmid and the mutant plasmid is constructed containing nicks. However, the original template still remains and needs to be removed. The template DNA is methylated due to a bacterial source, for example a mini-prep from *E. coli*, and the mutant plasmid is not, thus incubation with the restriction endonuclease DpnI can digest the template DNA. After digestion, transformation of the SDM product into an *E. coli* cloning strain will repair the nicked plasmid.
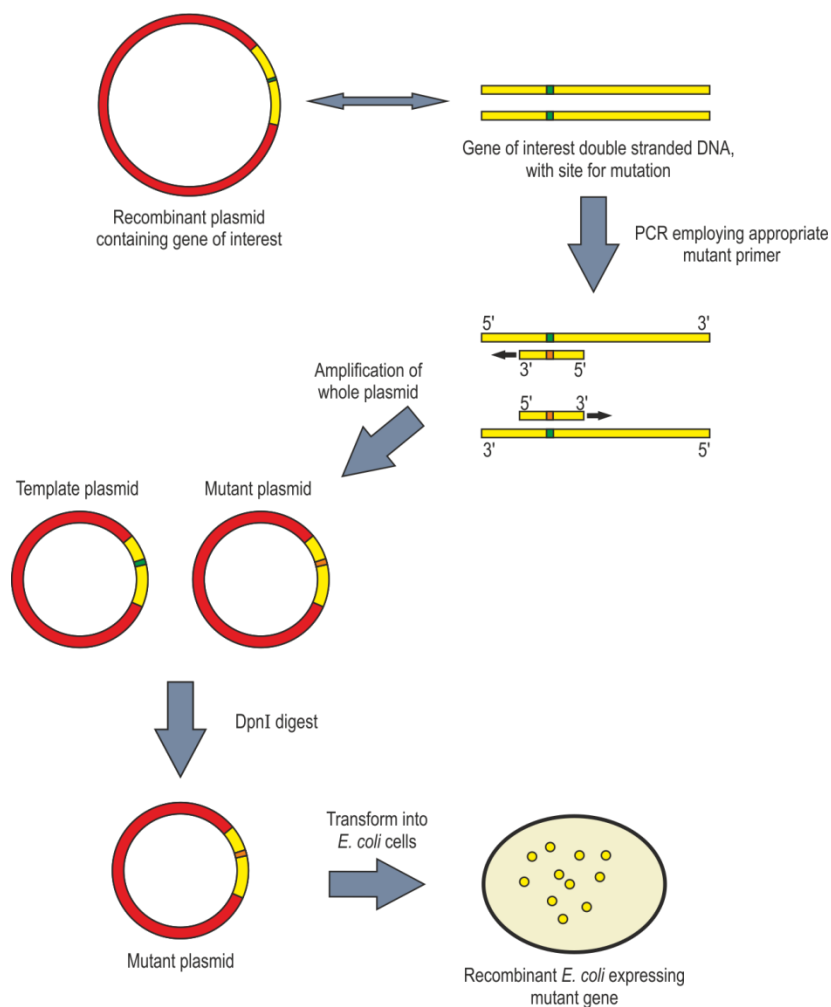
Figure 3.5 Diagram illustrating a method for site-directed mutagenesis (SDM). Template DNA is melted, the mutant primers anneal and the whole plasmid is amplified resulting in a nicked mutant plasmid. DpnI digestion removes template DNA and the *E. coli* cell is transformed by the mutant plasmid, repairing the nicks.

SDM primer design

A forward and reverse primer was designed for each mutation. Primers were designed using HiTel primer design software. All SDM experiments were carried out using a Qiuckchange^{TM} kit (Stratagene).

Table 3.15 Primers designed for SDM

Mutant	Forward	Reverse
F52N	GCATTCTCTGCATCTGAATAGCCCAGCGGGCTGG	CCAGCCCGCTGGGCTATTCAGATGCAGAGAATGC
F52A	CATTCTCTGCATCTGGCGAGCCCAGCGGGCTG	CAGCCCGCTGGGCTCGCCAGATGCAGAGAATG
N173F	CAATTATCGGTGGCGGTTTTCTGGCGCACAGATCC	GGATCTGTGCGCCAGAAAAACCGCCACCGATAATTG
N173Y	CAATTATCGGTGGCGGTTATTCTGGCGCACAGATCC	GGATCTGTGCGCCAGAATAACCGCCACCGATAATTG
N173H	GCAATTATCGGTGGCGGTCATTCTGGCGCACAGATCC	GGATCTGTGCGCCAGAATGACCGCCACCGATAATTGC
Q193R	GAAACGACTTGGATCACACCGTCACGAGCCGGCCTTTCTGGC	GCCAGAAAGGCCGGCTCGTGACGTGTGATCCAAGTCGTTTC
Q193E	GAAACGACTTGGATCACAGAACACGAGCCGGCCTTTCTG	CAGAAAGGCCGGCTCGTGTTCTGTGATCCAAGTCGTTTC
H194T	GACTTGGATCACACAGACCGAGCCGGCCTTTCTGG	CCAGAAAGGCCGGCTCGGTCTGTGTGATCCAAGTC
QH193RT	GACTTGGATCACACGTACCGAGCCGGCCTTTCTG	CAGAAAGGCCGGCTCGGTACGTGTGATCCAAGTC

PCR method

To a thin walled Eppendorf tube (0.5 mL) the components of Table 3.16 were added and mixed thoroughly.

Table 3.16 SDM reaction components

Component	Volume /µL	Concentration
DNA template	1	48 ng µL^{-1}
Forward primer	0.5	20 µM
Reverse primer	0.5	20 µM
dNTPs	5	2 mM
10 × turbo CX buffer	5	
DMSO	1	
Pfu turbo^{TM} (HS) DNA polymerase	1	

Water was added to make the final volume of the reaction 50 µL. The PCR tubes were placed in a thermal cycler and ran using the programme described in Table 3.17.

Table 3.17 SDM PCR programme

Step	Temperature /^{o}C	Time /min	Cycles
Initial denature	95	5	1
Denature	95	1	
Anneal	50	1	30
Extend	72	10	
Final denature	95	72 s	1
Final extend	72	5	1
Hold	4	-	-

DpnI digestion

Once the PCR was completed, $10 \times$ NEB buffer 2 (5 µL) and DPN I (1 µL, 20 k U mL^{-1}) were added to the reaction mixture and incubated at 37°C for 5 h.

Mutagenesis transformation and miniprep

The digested PCR mixture (1.5 µL) was transformed into *E. coli* XL10 Gold Ultra-competent cells (45 µL) and minipreped as described in Section 3.1.6.

3.5.2 Site saturation mutagenesis (SSM)

In addition to single rational mutagenesis, it may be useful to mutate a single residue to a number of different amino acids ranging from hydrophobic to hydrophilic, acidic to basic, small to bulky. This method is known as saturation mutagenesis and employs primers with 'NNK' and 'NDT' codons which replace the targeted codon. For NNK (N: adenine/cytosine/ guanine/thymine; K: guanine/thymine) 32 codons are utilised and allows for all twenty amino acids. For NDT, however, (D: guanine/thymine; T: thymine) only 12 codons are involved and therefore the number of amino acids available is reduced to 12 (Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly) but still cover a diverse chemical space.[99]

SSM primer design

Degenerate primers were designed to have a NDT codon at the desired site of mutation (Table 3.18). HiTel primer design software was used to design the primers. For all SSM reactions a Quickchange® Lightning Multi Site-Saturated Mutagenesis kit (Stratagene) was employed.

Table 3.18 Primers designed for SSM

Mutation	Primer
F52X	GGCATTCTCTGCATCTGNDTAGCCCAGCGGGCTGG
N173X	GCAATTATCGGTGGCGGTNDTTCTGGCGCACAG
S174X	CGGTGGCGGTAATNDTGGCGCACAGATCCTGGC
F283X	GGTGTACGGGCNDTCGTCCGGCTCTGTC

PCR method

To a thin wall Eppendorf tube (0.5 mL) the components of Table 3.19 were added and mixed thoroughly.

Table 3.19 SSM reaction components

Component	Volume /µL	Concentration
DNA template	1	48 ng µL^{-1}
primer	0.5	20 µM
Quickchange Lightening Multi 10× buffer	2.5	
Quickchange solution	0.75	
dNTPs	1	2 mM
DMSO	1	
Milli-Q water	17.25	
Quickchange Lightening Multi enzyme blend	1	

The PCR tubes were placed in a thermal cycler and ran using the programme described in Table 3.20.

Table 3.20 SSM PCR programme

Step	Temperature /^{o}C	Time /s	Cycles
Initial denature	95	120	1
Denature	95	20	
Anneal	55	30	30
Extend	65	420	
Final denature	95	72	1
Final extend	65	300	1
Hold	4	-	-

DpnI digestion

Once the PCR was completed, $10 \times$ NEB buffer 2 (5 µL) and DpnI (1 µL, 20 k U mL^{-1}) were added to the reaction mixture and incubated at 37°C for 5 h.

Mutagenesis transformation and miniprep

The digested PCR mixture (1.5 µL) was transformed into *E. coli* XL10 Gold Ultra-competent cells (45 µL) and minipreped as described in Section 3.1.6.

3.6 Chemical synthesis of sulfoxides

Unless otherwise stated, all reactions utilised oven dried glassware and magnetic stirrers. Solvents and liquid reagents were transferred using a syringe. All sulfides were commercially available. Evaporation of organic solvents was achieved by rotary evaporation with a water bath temperature of 41°C. Filtration was carried out using a Buchner funnel with a high vacuum. Thin layer chromatography was carried out in 100 % ethyl acetate using ultra-violet, all product spots were UV active.

Procedure
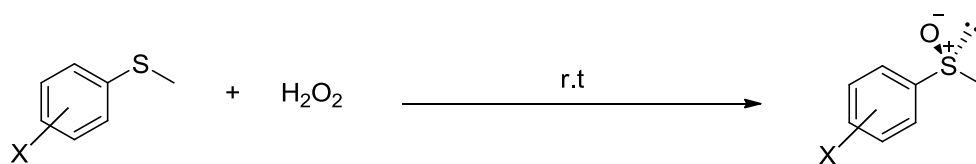
Figure 3.6 Reaction of a sulfide in glacial acetic acid with H_2O_2 to yield corresponding sulfoxide.

To a solution of corresponding sulfide (2.0 mmol) in glacial acetic acid (2 mL) hydrogen peroxide (8 mmol) was added drop wise whilst stirred at room temperature for 50 min. TLC (EtOAc) indicated reaction completion. The reaction mixture was neutralised with sodium hydroxide (NaOH, 2 M), washed with water (25 mL) and extracted twice with dichloromethane (3 × 10 mL). The combined organic layers were then washed with saturated sodium bicarbonate solution (25 mL) and brine (25 mL) and dried over sodium sulphate.[100] Removal of the dried organic solvent yielded the corresponding sulfoxide. The pure sulfoxide products were identified using thin layer chromatography, ^{1}H and ^{13}C NMR and mass spec.

3.7 Protein crystallisation

Protein structures can be determined by subjecting crystals of a protein to X-ray crystallography. A crystal is a 3D ordered array of molecules and the X-rays are diffracted by the electrons in the structure to give a 3D map highlighting the distribution of electrons. If the order of the crystal is poor the X-rays will not be diffracted to a high resolution and thus the data will not give a detailed structure. However, a well ordered crystal will diffract the X-rays with high resolution leading to a detailed structure. The diffraction pattern is visualised as a series of spots or reflections due to the constructive and deconstructive effects. Each reflection provides information on all the atoms in the structure and each atom is responsible for the intensity of each reflection.

To achieve protein crystals, crystallisation screens containing many different conditions are set up. If a screen is successful the 'hit' can be scaled up using the hanging drop vapour diffusion method. At the beginning of the scale up the concentration of the precipitant in the drop is half that of the well. Equilibration then occurs through the

vapour phase and the concentration of the well remains the same. The drop loses water vapour to the well until the concentration of the precipitant is equal to that of the well. If such conditions are favourable during this process the protein will become supersaturated and will be driven out of solution in the form of crystals.

Crystallisation of SMFMO

Pure SMFMO was subjected to crystallisation trials using a range of commercially available screens in 96-well plates employing 300 nL drops at a range of protein concentrations (3, 4, 6 and 10 mg mL^{-1}). The best crystals were obtained using the Clear Strategy Screen[101] in conditions containing lithium sulphate (1.8 M) and protein (4 mg mL^{-1}). Larger crystals for diffraction analysis using optimised conditions were prepared using the hanging-drop vapour diffusion method in 24-well plate Linbro dishes and using crystallisation drops of 2 mL with protein (4 mg mL^{-1}). The best crystals were routinely obtained in crystal drops containing lithium sulphate (0.9 M) in bis-tris propane buffer (100 mM) at pH 5.6.

Crystallisation of Q193R/H194T mutant

The crystallisation conditions for the double mutant were similar to that for the wild type. The crystals used for diffraction experiments were obtained from hanging crystal drops set up in 24-well plates. The reservoir contained 0.9 M lithium sulphate in 100 mM bis-tris propane buffer at pH 5.6. The protein concentration used was 3 mg mL^{-1}.

Crystals were tested for diffraction using a Rigaku Micromax-007HF generator fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. The crystals that diffracted to a resolution of greater than 3 Å were flash-cooled in liquid nitrogen in a cryogenic solution containing the mother liquor containing also 10% w/v glycerol and retained for data collection at the synchrotron.

Chapter 4: Characterisation of SMFMO

4.1 Introduction

Flavoprotein monooxygenases (FPMOs) are of potential interest for their ability to catalyse asymmetric organic reactions, such as Baeyer-Villiger oxidations and heteroatom oxidations. The biocatalytic synthesis of enantiopure products has become increasingly attractive owing to the enzymes selectivity and green reaction conditions. FPMOs commonly use NADPH as cofactor which, from an industrial application point of view, is a drawback to these enzymes as NADPH is expensive. NADH however, is cheaper and efforts have been underway to find FPMOs that can utilise the cheaper, non-phosphorylated cofactor.

SMFMO sparked initial interest due to the enzymes ability to use either NADH or NADPH as cofactor to catalyse oxidation reactions. Additionally, the 38.6 kDa protein was found to be much smaller than typical class B bacterial FMOs and a BLAST search of the Swissprot database revealed closest sequence homology to shorter putative flavin monooxygenases from strains such as *Pseudomonas*. In addition, SMFMO had sequence homology with the pyridine nucleotide disulphide oxidoreductase enzyme family, such as thioredoxin reductases (TrxRs). The TrxRs homology prompted a sequence-based search of the RCSB Protein Structure database which also highlighted that the closest sequence homology to proteins with solved structures also included TrxRs, from *Thermus thermophilius* (2ZBW) and *Helicobacter pylori* (2Q0K[102]), along with BVMOs PAMO (1W4X[61a]) and CHMO from *Rhodococcus* (3GWD[59]), and FMOs from *S. pombe* (2V8) and *Methylophaga spp* (2XVI[74], 2XLT[103]), shown in Figure 4.1.

```
PAMO    MAGQTTVDSRRQPPEEVDVLVVGAGFSGLYALYRLRELG------RSVHVIETAGDVGGV 54
SMFMO   -------------MDSVDVVVIGGGQSGLSAGYFLRRSG------LSYVILDAEASPGGA 41
2ZBW    ----------MAADHTDVLIVGAGPTGLFAGFYVGMRG------LSFRFVDPLPEPGGQ 43
2XVI    -------------MATRIAILGAGPSGMAQLRAFQSAQEKGAEIPELVCFEKQADWGGQ 46
                         . : ::*.* :*:                        **

PAMO    WYWNRYPGARCD-------------IESIEYCYSFSEEVLQEWN--WTERYASQPEILRY 99
SMFMO   WQHAWHSLHLFS-------------PAGWSSIPGWPMPASQGP-------YPARAEVLAY 81
2ZBW    LTALYPEKYIYD-------------VAGFPKVYAK---------------DLVKGLVEQ 74
2XVI    WNYTWRTGLDENGEPVHSSMYRYLWSNGPKECLEFADYTFDEHFGKPIASYPPREVLWDY 106
                     .                                  .

PAMO    INFVADKFDLRSGITFHTTVTAAAFDEATNTWTVDTNHGDRIRAR-----YLIMASGQ-L 153
SMFMO   LAQYEQKYALP----VLRPIRVQRVSHFGERLRVVARDGRQWLAR-----AVISATGT-W 131
2ZBW    VAPFNPVYSLG--------ERAETLEREGDLFKVTTSQGNAYTAK-----AVIIAAGVGA 121
2XVI    IKGRVEKAGVRKYIRFNTAVRHVEFNEDSQTFTVTVQDHTTDTIYSEEFDYVVCCTGH-F 165
            :           :        ... :    * . .          :: ..:*

PAMO    SVPQLPNFPGLKDFAGNLYHTGNWPHEPVDFSGQRVGVIGTGSSGIQVSPQIAKQAAELF 213
SMFMO   GEAYTPEYQGLESFAGIQLHSAHYS-TPAPFAGMRVAIIGGGNSGAQILAEVSTVAETTW 190
2ZBW    FEPRRIGAPGEREFEGRGVYYAVKS--KAEFQGKRVLIVGGGDSAVDWALNLLDTARRIT 179
2XVI    STPYVPEFEGFEKFGGRILHAHDFR-DALEFKDKTVLLVGSSSSAEDIGSQCYKYGAKKL 224
               * ..* *  :       *  . * ::* ..*.:  :    .

PAMO    VFQRTPHFAVPARNAPLDPEFLADLKKRYAEFREESRNTPGGTHRYQGPKSALEVSDEEL 273
SMFMO   ITQHEPAFLADDVDGRVLFERATERWKAQQEGREPD-LPPGG------------------ 231
2ZBW    LIHRRPQFRAHEAS-------VKELMKAHEE---------------------------- 203
2XVI    ISCYRTAPMGYKWPENWDERPNLVRVDTENAYFADGSSEK-------------------- 264
        :            .                   .

PAMO    VETLERYWQEGGPDILAAYRDILRDRDANERVAEFIRNKIRNTVRDPEVAERLVPKGYPF 333
SMFMO   ------------------FGDIV-------MVPPVLDARARGVLAAVPPPARFSPTGMQW 266
2ZBW    -------------------------GRLEVLTPYELRRVEGDERVRWAVVFHNQTQ 234
2XVI    -------------------VDAIILCTGYIHHFPFLNDDLR-LVTNNRLWPLNLYKGVVW 304
                                   .       *              .

PAMO    GTKRLILEIDYYEMFNRDNVHLVDTLSAPIETITPRGVRTSEREYELDSLVLATGFDALT 393
SMFMO   ADG--------------------------------------TERAFDAVIWCTGFRPAL 287
2ZBW    EEL--------------------------------------ALEVDAVLILAGYITKL 254
2XVI    EDNPKFFYIG-----------------------------MQDQWYSFNMFDAQAWYAR 333
                                                     . :   :

PAMO    GALFKIDIRG-VGNVALKEKWAAGPRTYLGLSTAGFPNLFFIAGPGSPSALSNMLVSIEQ 452
SMFMO   SHLKGLDLVTPQGQVEVDG---------SGLRALAVPSVWLLGYGDWNGMASATLIGVTR 338
2ZBW    GPLANWGLALEKNKIKVDT----------TMATSIPGVYACGDIVTYPGKLPLIVLGFG 303
2XVI    DVIMGRLPLPSKEEMKADS-----------MAWREKELTLVTAEEMYTYQGDYIQNLID 381
            . :        ::  .                   .              .

PAMO    HVEWVTDHIAYMFKNGLTRSEAVLEKEDEWVEHVNEIADETLYPMTASWYTGANVPGKPR 512
SMFMO   YAREAVRQVTAYCADHQDR----------------------------------------- 357
2ZBW    EAAIAANHAAAYANPALKVNPGHSSEKAAPGT---------------------------- 335
2XVI    MTDYPSFDIPATNKTFLEWKHHKKENIMTFRDHS-----------YRSLMTGTMAPKHHT 430
           .       . .

PAMO    VFMLYVGGFHRYRQICDEVAAKGYEGFVLT 542
SMFMO   -----------------------------
2ZBW    -----------------------------
2XVI    PWIDALDDSLEAYLSDKSEIPVAKEAL--- 457
```

Figure 4.1 Sequence alignments of four related flavin-containing enzymes.
PAMO = phenylacetone monooxygenase from *Thermobifida fusca*; SMFMO = *Stenotrophomonas maltophilia* Flavin-containing Monooxygenase; 2ZBW = TrxR-like protein from *Thermus thermophilus*; 2XVI = 2XLT = mFMO from *Methylophaga aminisulfidivirans*. Sequences in red and blue show the Rossman motifs and sequences in green indicate the BVMO/FMO motif region.

A phylogenetic tree (Figure 4.2) possessing representative sequences of FMOs, BVMOs and TrxRs was used to analyse SMFMO and indicated that the 38.6 kDa enzyme is as closely related to TrxRs as it is to BVMOs and FMOs, as the SMFMO branch is placed within the middle of all three enzyme families.
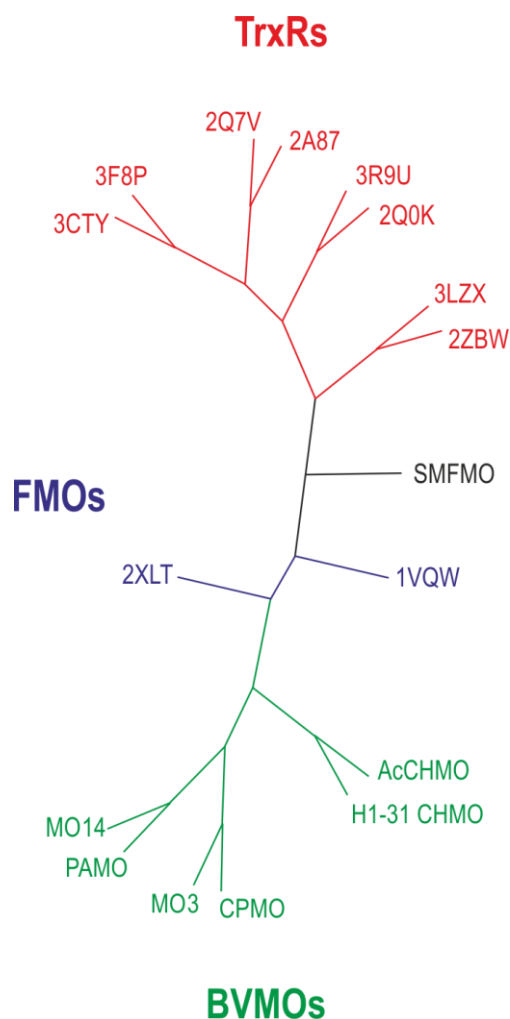
Figure 4.2 A phylogenetic tree indicating the position of SMFMO in relation to some BVMOs, FMOs and flavin-binding TrxRs for which structure and/or function has been identified.

TrxRs: 3F8P= TrxR from *Sulfolobus solfataricus* (Uniprot Q8X236); 3CTY= TrxR from *Thermoplasma acidophilum* (Uniprot Q9HJI4); 2A87= TrxR from *Mycobacterium tuberculosis* (Uniprot P52214); 2Q7V= TrxR from *Deinococcus radiodurans* (Uniprot Q9RSY7); 2Q0K= TrxR from *Helicobacter pylori* (Uniprot P56431); 3R9U= TrxR from *Campylobacter jejuni* (Uniprot Q0PBZ1); 3LZX= Ferredoxin-NADP$^+$ oxidoreductase from *Bacillus subtilis* (Uniprot O05268); 2ZBW= TrxR-like protein from *Thermus thermophilus* (Uniprot Q5L28); **FMOs:** 1VQW= Flavin-containing monooxygenase from *Schizosaccharomyces pombe* (Uniprot Q9HFE4); 2XLT= FMO from *Methylophaga aminisulfidivirans* (Uniprot Q83K4); **BVMOs:** HI-31 CHMO= Cyclohexanone monooxygenase from *Rhodococcus* sp. HI-31 (3GWD, Uniprot C0STX7); AcCHMO= Cyclohexanone monooxygenase from *Acinetobacter calcoaceticus* (Uniprot Q9R2F5); MO14= ro03437 from *Rhodococcus jostii* RHA1 (Uniprot Q0SB46); MO3= ro03247 from *Rhodococcus jostii* RHA1 (Uniprot Q0SBN6); PAMO= phenylacetone monooxygenase from *Thermobifida fusca* (Uniprot Q47PU3); CPMO= Cyclopentanone monooxygenase from *Comamonas* sp. (Uniprot Q8GAW0).

SMFMO's ability to use either NADH or NADPH along with its unusual size prompted further investigation into this enzyme.

4.2 Aims

To express the gene encoding SMFMO, and purify the enzyme, in order to characterise its activity and substrate specificity using biochemical techniques, such as UV spectrophotometry and gas chromatography.

4.3 Materials and methods

4.3.1 Gene expression and protein purification

SMFMO was expressed in *E. coli* strain BL21 (DE3) as described in Section 3.2.5. SMFMO was purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography described in Section 3.3.

4.3.2 Enzyme assays

Kinetic assays

Kinetic studies of SMFMO where carried out in accordance with Section 3.4.2 and Section 3.4.3. Kinetic parameters were calculated using the GraFit data analysis software.

Biotransformations with SMFMO

Biotransformations were carried out as described in Section 3.4.4 using a range of class B FPMO ketone substrates and prochiral sulfides (Figure 4.8). Biotransformations were analysed by GC and GC-MS described in Sections 3.4.5 and 3.4.6.

4.4 Results

4.4.1 Gene expression and protein purification

The SMFMO gene in the pET-YSBLIC-3C vector was expressed in BL21 (DE3) cells as described in Section 3.2.5 with a total volume of 8 L LB media. After expression the protein was purified.

SMFMO purification *via* Ni^{2+} affinity chromatography

The SMFMO protein was purified successfully by Ni^{2+} affinity chromatography described in Section 3.3.2, shown in the chromatogram below (Figure 4.3).
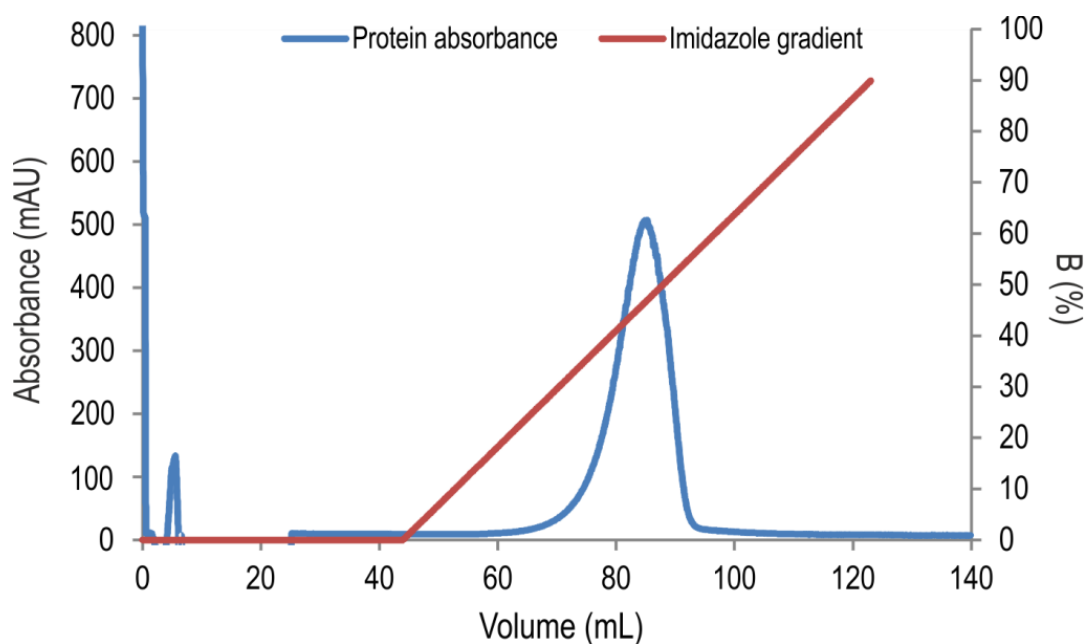
Figure 4.3 Chromatogram for SMFMO purification by Ni^{2+} affinity chromatography. Protein absorbance at 280 nm is indicated by the blue trace.

The protein eluted as a single peak (from 65-95 mL). The fractions eluted were yellow in colour, signifying that FAD is present in the protein. The fractions were analysed by SDS-PAGE (Section 3.2.4) to confirm the presence of pure protein. The gel can be seen below (Figure 4.4).
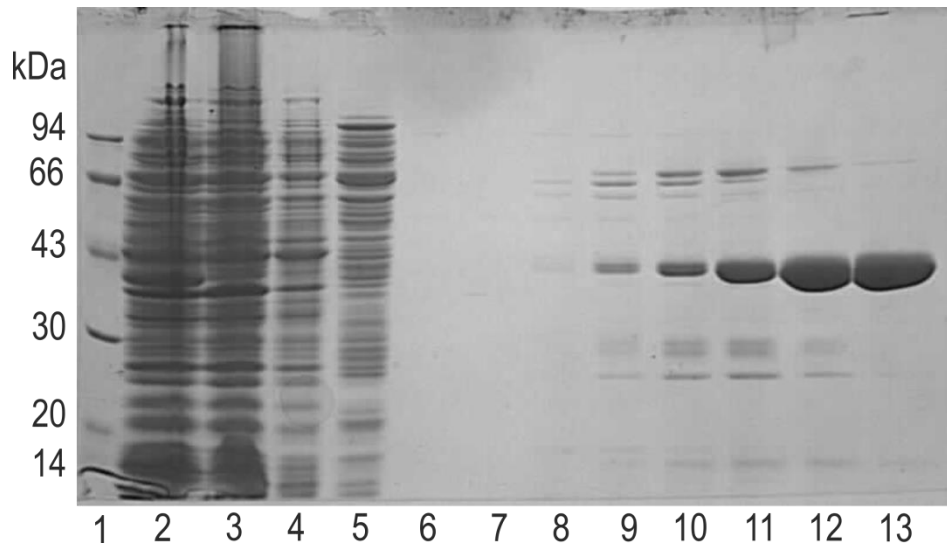
Figure 4.4 SDS-PAGE gel of SMFMO purification after Ni^{2+} affinity chromatography. Lane 1: Bio-Rad low molecular weight marker, lane 2: soluble protein fraction, lane 3: flow through after column loading, lanes 4-5: flow through after buffer A wash, lanes 6-7: collected fractions after FPLC run. The protein can be seen as a thick dark band at 38 kDa in lanes 10-13.

The SMFMO protein can be seen as a dark band at 38 kDa in lanes 10 to 13. The fractions containing protein were pooled and concentrated to a volume of 2 mL (Section 3.3.2).

SMFMO purification by size exclusion chromatography

The SMFMO protein underwent further purification by size exclusion chromatography (Section 3.3.3), shown in the chromatogram below (Figure 4.5).
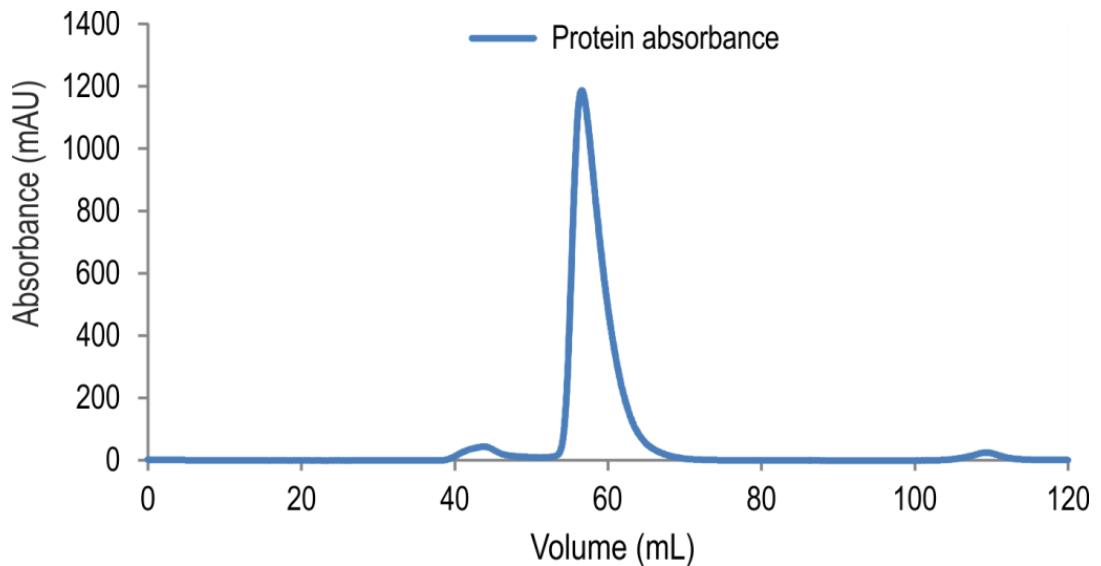
Figure 4.5 Chromatogram of SMFMO purification by size exclusion chromatography. Protein absorbance at 280 nm is indicated by the blue trace.

The protein eluted as a single peak (from 55-70 mL). Using a calibration curve provided by GE Healthcare, the position of the elution peak gave the molecular weight at ~80 kDa, indicating that SMFMO is a dimer in solution. The fractions eluted were bright yellow in colour, indicating FAD was still present within the purified protein. The corresponding fractions were analysed by SDS-PAGE (Section 3.2.4). The gel can be seen in Figure 4.6.
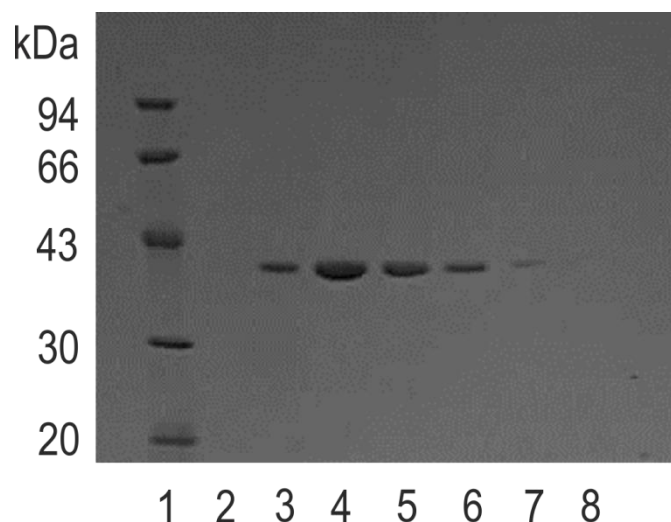
Figure 4.6 SDS-PAGE gel of SMFMO purification by size exclusion chromatography. Lane 1: Bio-Rad low weight molecular marker, lane 2-8: collected fractions after FPLC run. The protein can be seen at 38 kDa.

The SMFMO protein can be seen as a dark band on the gel at 38 kDa. The gel shows the protein is pure as no other bands are present in the lanes. The fractions were combined to give ~21 mg of pure protein (Section 3.3.3).

4.4.2 Kinetic assays

Cofactor utilisation by SMFMO

In preliminary experiments, it was observed that SMFMO was able to utilise either NADH or NADPH as cofactor in the oxidation of standard FPMO substrates. Kinetic studies of cofactor utilisation by SMFMO were performed as described in Section 3.4.2, using a constant concentration of **1** in the presence of increasing concentrations of either cofactor. The results are shown in Figure 4.7.
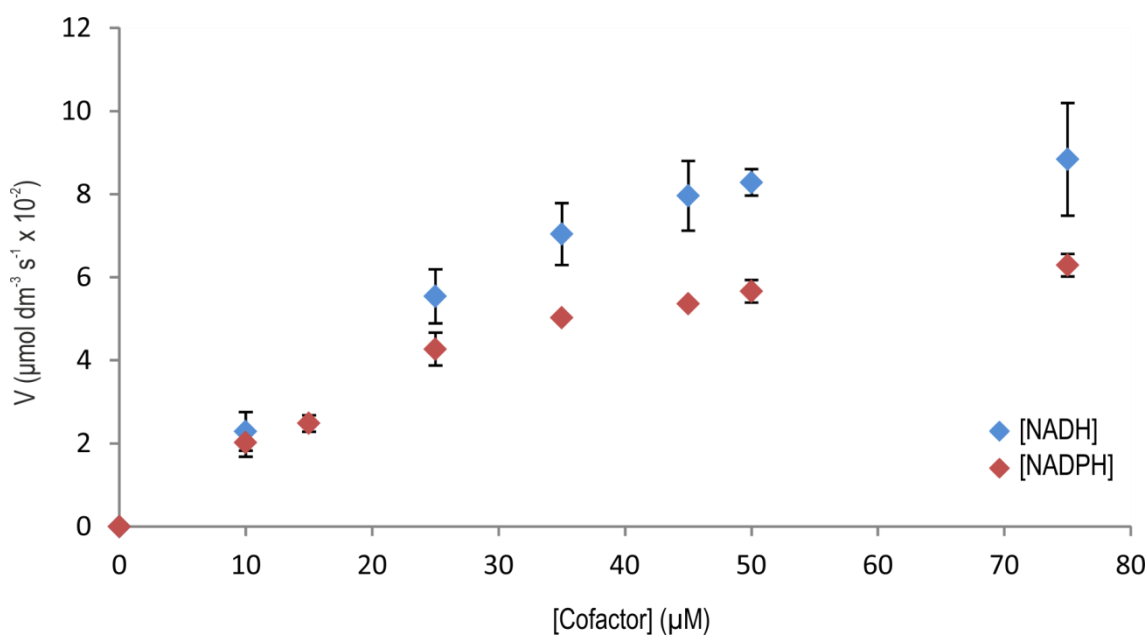
Figure 4.7 Kinetics of cofactor (NAD(P)H) utilisation by SMFMO.
Kinetic parameters calculated: NADH: V_{max} 11.2 × 10^{-2} μM s^{-1}, K_M 23 μM; NADPH: V_{max} 8.4 × 10^{-2} μM s^{-1}, K_M 27 μM.

For the NADH assay, the K_M was 23 μM and V_{max} 11.2 × 10^{-2} μM s^{-1}. Similarly for the NADPH assay, the K_M was calculated to be 27 μM and V_{max} 8.4 × 10^{-2} μM s^{-1} (Table 4.1). The K_M and V_{max} values obtained for NADH and NADPH in the assay were comparable, indicating that the enzyme is able to use either cofactor to reduce the bound

flavin with equal capability. In addition, the values are also comparable to those of the NADPH oxidase activity of the FMO mFMO from *Methylophaga* (13 µM and 0.06 s⁻¹).[67]

Transformation of substrate 1 by SMFMO

It was established that SMFMO was able to utilise both NADH and NADPH as cofactor to reduce the bound FAD in the absence and presence of constant concentrations of **1**. Kinetic studies of SMFMO in the presence of varying concentrations of **1** were performed as described in Section 3.4.3, using a constant concentration of either cofactor.

The studies will indicate if the ketone oxidation would conform to Michaelis-Menten kinetics, indicating that the preliminary studies observed by GC are indeed genuinely 'enzymatic'. The results are shown in the graphs below (Figure 4.8 and Figure 4.9).
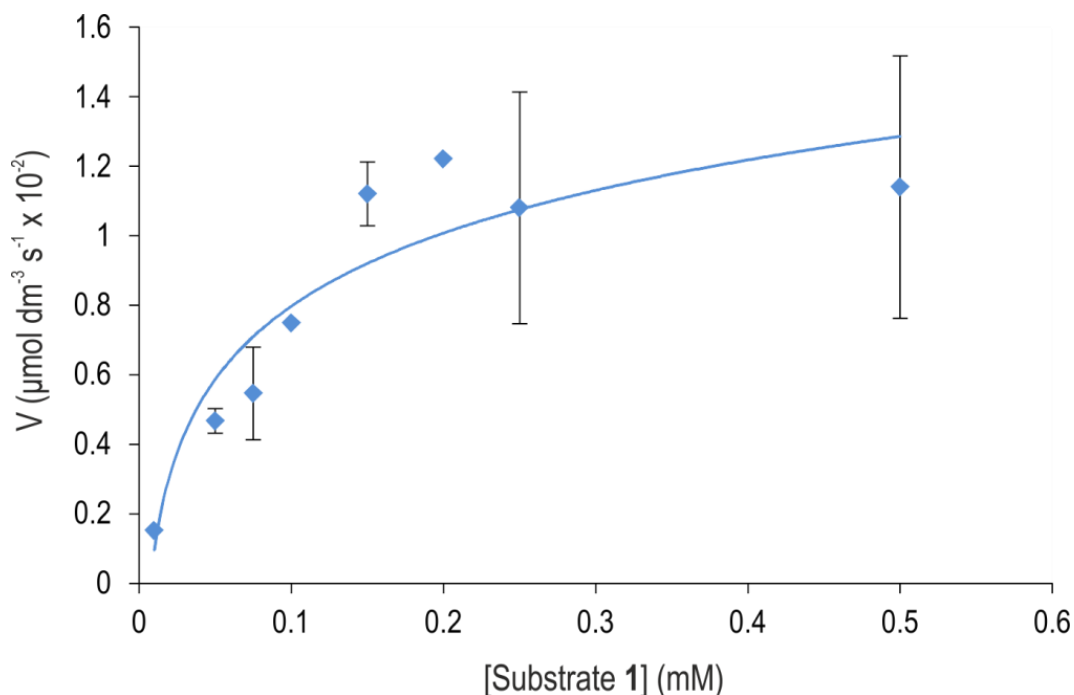
Figure 4.8 Kinetics of substrate **1** transformation by SMFMO using NADH as cofactor.
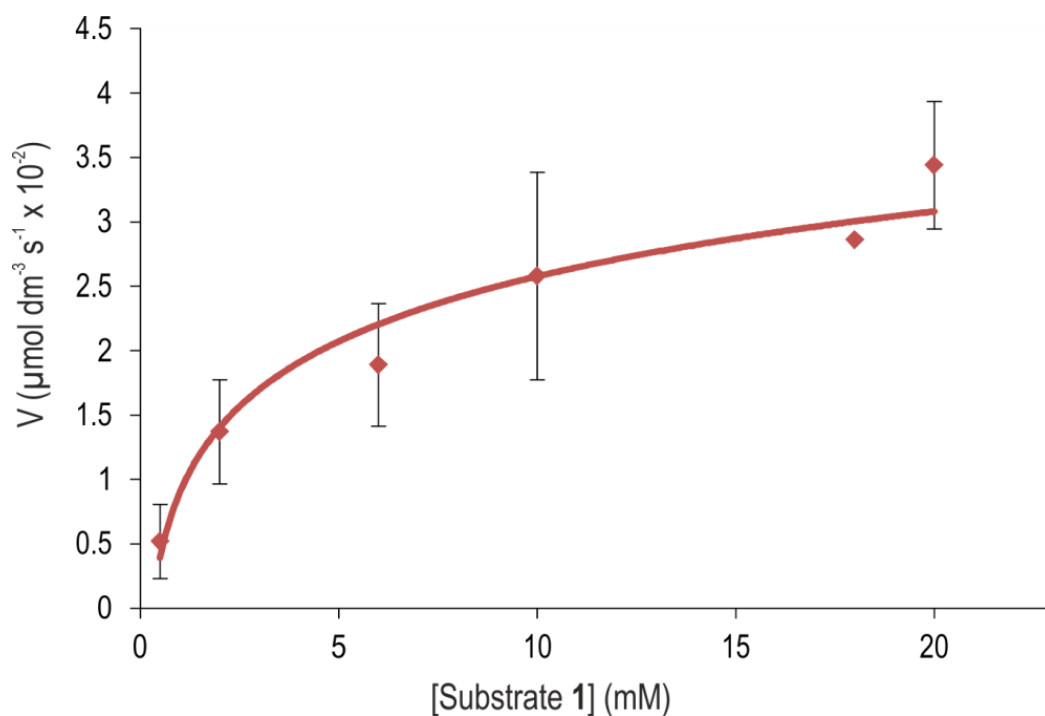
Figure 4.9 Kinetics of substrate **1** transformation by SMFMO using NADPH as cofactor.

Kinetic constants for **1** oxidation were calculated, shown in Table 4.1.

Michaelis-Menten behaviour was observed for both NADH and NADPH assays with respect to **1**. The K_M for the ketone substrate in the presence of NADH was calculated to be 70 µM, approximately 42-fold lower than that observed for NADPH (K_M 3.2 mM). The k_{cat}/K_M for NADH was 51 M^{-1} s^{-1}, 17-fold higher than that calculated for NADPH (3 M^{-1} s^{-1}). Such results lead to the assumption that the binding of NADH assists the transformation of the ketone substrate to a greater extent, and thus the protein has a greater affinity for **1** when NADH is utilized.

Table 4.1 Kinetic constants for SMFMO using NADH or NADPH as cofactor, with **1** as substrate.

Assay	K_M (µM)	V_{max} (µM s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
NADH (100 µM **1**)	23.7 ± 9.1	$11.2 \pm 1.5 \times 10^{-2}$	0.029	1223
NADPH (100 µM **1**)	27.3 ± 5.3	$8.4 \pm 0.6 \times 10^{-2}$	0.022	806
1 (250 µM NADH)	70.0 ± 30.0	$1.4 \pm 0.2 \times 10^{-2}$	3.6×10^{-3}	51
1 (250 µM NADPH)	3200 ± 900	$3.3 \pm 0.3 \times 10^{-2}$	8.5×10^{-3}	3

4.4.3 Substrate selectivity and enantioselectivity

A range of class B FPMO ketones (Figure 4.10) was tested as substrates for SMFMO, using NADH or NADPH as nicotinamide cofactor with the appropriate recycling system, described in Section 3.4.4. The reactions were analysed using gas chromatography described in Section 3.4.5.

Controls were set up, involving the same assays tested but in the absence of enzyme. For each of the samples, NADH and NADPH, no lactone was produced thus indicating that SMFMO catalyses the conversion of **1** to its corresponding lactone.
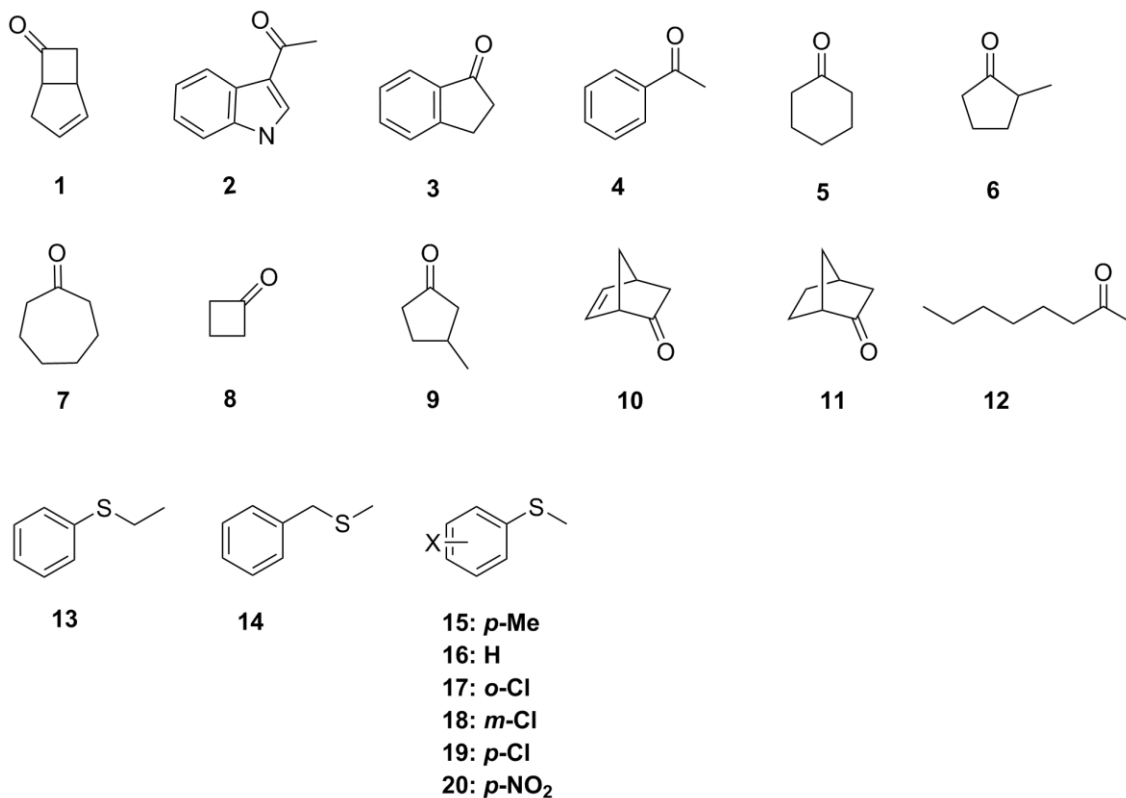
Figure 4.10 Range of ketone and sulfide substrates used in SMFMO characterisation.

SMFMO Baeyer-Villiger activity was found to be restricted to the strained, fused bicyclic system of **1**, with the remaining more simple aliphatic and alicyclic substrates such as **5**, **8** and **12** not transformed with either cofactor. Conversion of **1** was notably higher when NADH was utilised as cofactor with 93 % conversion observed compared to 15 % with NADPH over 24 hours (Figure 4.11).
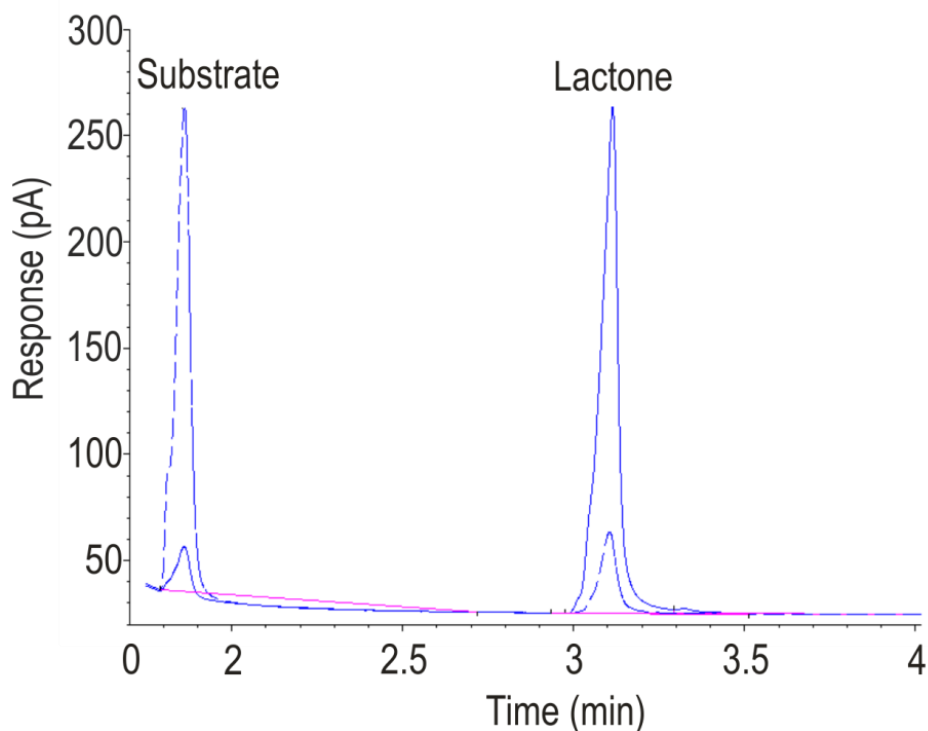
Figure 4.11 GC trace for t=24h for NADH and NADPH biotransformation. Utilising **1** as substrate. NADH is represented by the solid blue line and NADPH by the dashed blue line. Substrate peak ~1.8 min and lactone product peak ~3 min.

The NADH-dependent biotransformation of **1** was regioselective, giving a yield of regioisomers in the ratio 5:1, favouring the expected 2-oxa lactone over the 3-oxa lactone. Enantioselectivity was poor, but was measured, with 8 % *ee* for the (1*R*,5*S*) lactone and 36 % *ee* for the (1*S*,5*R*) lactone (Figure 4.12). Chiral GC analysis of the residual ketone substrate showed it to be racemic.
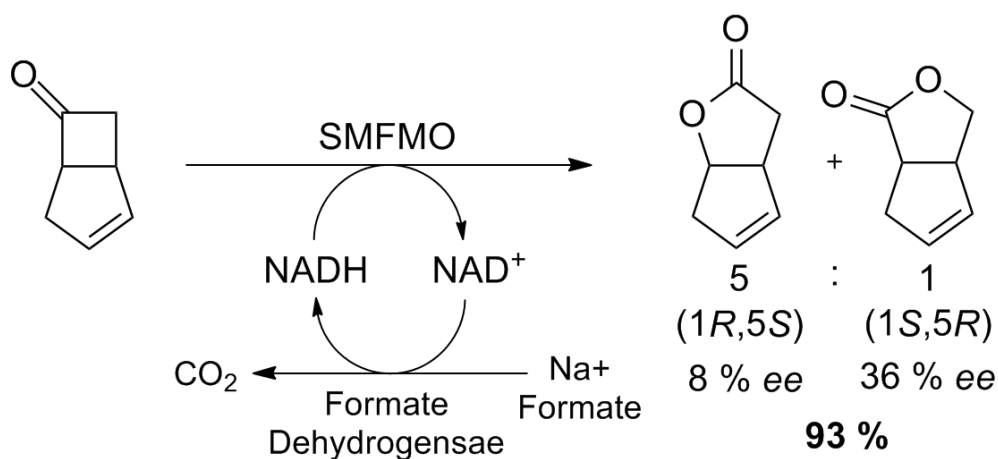
Figure 4.12 NADH-dependent transformation of **1**.
For NADPH, using glucose-6-phosphate dehydrogenase and glucose-6-phosphate as the recycling system, the conversion of the substrate was 15 %.

A range of prochiral sulfides were tested (Figure 4.10) as substrates for SMFMO as described in Section 3.4.4. SMFMO displayed more promiscuous substrate selectivity with the prochiral sulfide substrates and conversions were observed for the majority investigated, as shown in Table 4.2. Substrate **20** was not transformed. The sulfoxidation reactions by SMFMO were additionally confirmed by GC-MS, example shown in Figure 4.13.
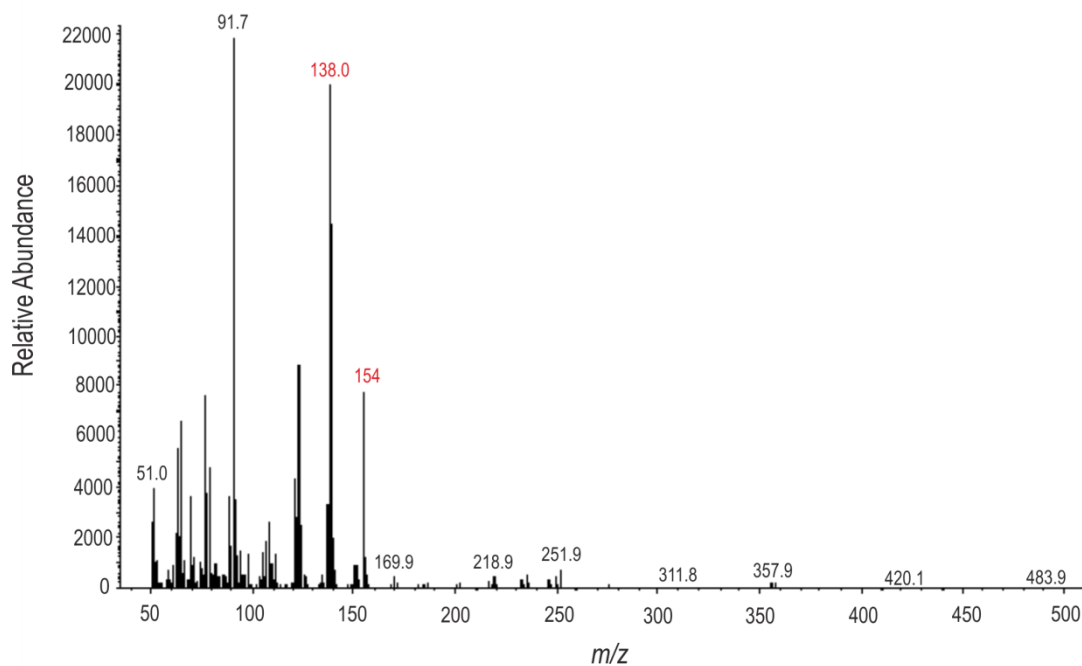
Figure 4.13 GC-MS spectra for NADH t=24 h with substrate **15**.
The spectra shows *m/z* peaks for both substrate **15** (138) and corresponding sulfoxide product (154).

The *m/z* peak for substrate **15** was present at 138 and an *m/z* peak at 154 for the sulfoxide product peak, verifying that the substrate was actually converted to the sulfoxide. For all reactions the *m/z* peak of the substrate and product could be seen.

For each of the successful transformations, greater conversions were observed when NADH was employed as cofactor, example shown in Figure 4.14.
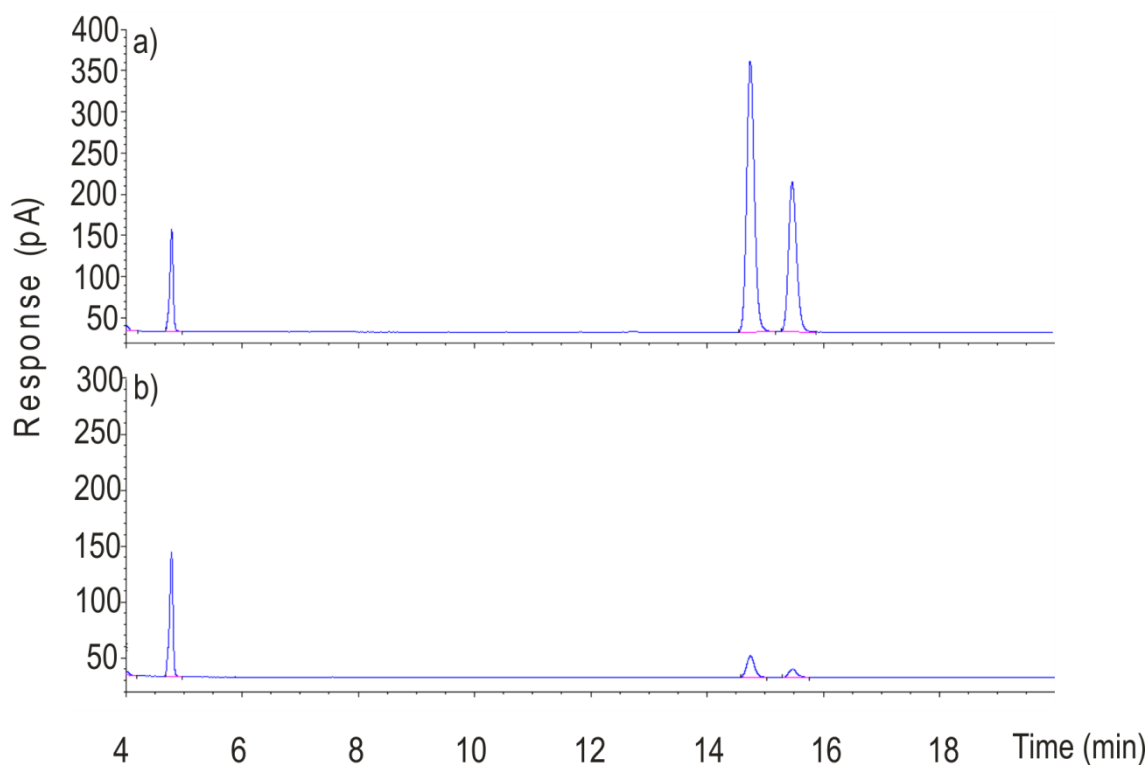
Figure 4.14 Chiral GC trace at t= 24 h for the oxidation of substrate **15** by SMFMO. NADH is shown in graph a), NADPH in graph b).

The GC trace also indicates that the enzyme favours one enantiomer over the other with an *ee* of 25 % for NADH and 44 % for NADPH. This is interesting as the results suggest that the enzyme had a "greater preference" for one enantiomer with NADPH than NADH, therefore indicating that the substrate binding is different depending on which cofactor is used.

However, an argument against such findings may lie in the accuracy of the chiral GC measurements particularly where the *ee*s for each cofactor are concerned as the results were not repeated. This is particularly the case for substrate **16** when NADPH is employed as the conversion is only 1 %, and without repeat results it's difficult to determine if this conversion of sulfide to sulfoxide is due to SMFMO or actually present within the sulfide starting material.

Table 4.2 Results of biotransformations of prochiral thioether substrates by SMFMO.

Substrate	Conversion NADH %	Conversion NADPH %	ee NADH %	ee NADPH %	Sulfoxide Configuration NADH	Sulfoxide Configuration NADPH
13	27	2	71	57	*R*	*R*
14	32	10	24	38	*R*	*R*
15	90	33	25	44	*R*	*R*
16	8	1	21	34	*R*	*n.d.*
17	6	0.2	15	0	*S*	*n.d.*
18	18	2	30	25	*n.d.*	*n.d.*
19	40	9	80	82	*R*	*R*
20	0	0	0	0	-	-

Absolute configurations were assigned by comparison with the sulfoxide products produced using CHMO.[23a, 98] For **18** a configuration could not be assigned.

4.5 Discussion

The flavoprotein monooxygenase SMFMO catalyses the asymmetric oxidation of thioethers and is unusual amongst FPMOs as it has the ability to use either nicotinamide cofactor, NADH or NADPH, in order to reduce the bound flavin. The K_M, V_{max} and k_{cat} values were comparable for both NADH and NADPH, with K_M in the region of 20-30 µM, indicating that the enzyme had the ability to use either cofactor with equal capability, and are comparable with the NADPH oxidase activity of the flavo monooxygenase mFMO from *Methylophaga* (13 µM, 0.06 s^{-1}).[67] However, for the ketone substrate in the presence of 250 µM cofactor Michaelis-Menten behaviour was observed, with the K_M value for the ketone substrate in the presence of NADH approximately 42 times lower, and the k_{cat}/K_M 17 times greater than for NADPH. This suggests that the transformation from ketone to lactone is variable depending on what

cofactor is used. The V_{max} value for substrate **1** with NADH is noticeably smaller than that usually observed for BVMOs acting on ketones but is in the range of the FPMO MtmOIV, when conveyed in the same terms (5 nmol min mg^{-1} SMFMO versus 147 nmol min mg^{-1} MtmOIV),[35a] and in this case MtmOIV was acting on its natural substrate.

The enzyme catalysed the Baeyer-Villiger conversion (93 %) of 5 mM **1** in the presence of sub-stoichiometric concentrations of NADH with the appropriate cofactor recycling system, to give the 2-oxa and 3-oxa lactone products in a ratio of 5:1, however with poor enantioselectivity. The conversion with NADPH was 15 %. Recently a group of FMOs related to SMFMO have been explored, encoded in the genome of *Rhodococcus jostii* RHA1.[18] The study reveals a set of enzymes that have the ability to use either NADH or NADPH as cofactor. FMO-E, FMO-F and FMO-G were found to catalyse the oxidation of **1**, with conversions similar to that of SMFMO. As with SMFMO, the enzymes are equally not very enantioselective in the formation of the 2-oxa lactone, but FMO-F and FMO-G yield >65 % *ee* for the 3-oxa lactone. SMFMO also catalysed the NADH-dependent transformation of prochiral aromatic thioethers, with preference for the (*R*) enantiomer. The type II FMOs from *R. jostii* RHA1 converted substrate **16** with similar enantioselectivity to SMFMO. SMFMO also catalysed the NADH-dependent transformation of prochiral aromatic thioethers, with preference for the (*R*)- enantiomer and the best case being the conversion of **19** with 80 % *ee*. In most cases, specifically **16**, the enantioselectivity was poorer than for CHMO[23a, 98] or mFMO[73], however, SMFMO had a much greater *ee* for substrate **19** and in the opposite enantiomeric series. The preferred absolute configuration of sulfoxide products by SMFMO was the (*R*)-enantiomer, similar in most cases to that encountered for chloroperoxidase. Interestingly, the *R. jostii* RHA1 enzymes had almost identical conversions and enantioselectivities of substrates when either NADH or NADPH was used as cofactor. SMFMO however, had much higher conversions for all substrates when NADH was employed albeit with poorer *ee*s than for NADPH in most cases. The higher *ee*s when NADPH is employed may suggest that binding of the substrate is different depending on which cofactor is used. Thus, NADPH allows for the substrate to bind in a specific way which is preferable within the reaction therefore yielding a higher *ee*.

Together with the type II FMOs from *R. jostii* RHA1, SMFMO represents a new class of class B FPMOs that can employ both NADH and NADPH. Further comparison between the two similar enzymes will be discussed in subsequent chapters.

4.6 Conclusion

SMFMO is found to accept both NADH and NADPH as coenzyme, and has identified a single-component FPMO capable of NADH-dependent oxygenation reactions. This is in divergence to that of usual BVMOs and FMOs that typically prefer NADPH as nicotinamide cofactor. The unusual relaxed coenzyme specificity of SMFMO is attractive for biocatalytic applications as NADH, is much cheaper than its phosphorylated neighbour.

In conclusion, further investigation into SMFMO could highlight a possible platform for developing NADH-dependent FPMOs for asymmetric oxygenation reactions.

Chapter 5: Structure determination of SMFMO

5.1 Introduction

In an effort to understand why SMFMO has the ability to use either NADPH or NADH as nicotinamide cofactor, and thus determine possible avenues for developing NADH-dependent FPMOs for asymmetric oxygenation reactions, the structure of SMFMO had to be solved.

Within the flavoprotein monooxygenase family the BVMO phenlyacetone monooxygenase (PAMO) from the thermophile *Thermobifida fusca*, complexed with FAD, was the first to have its structure determined (Figure 5.1).[61a]
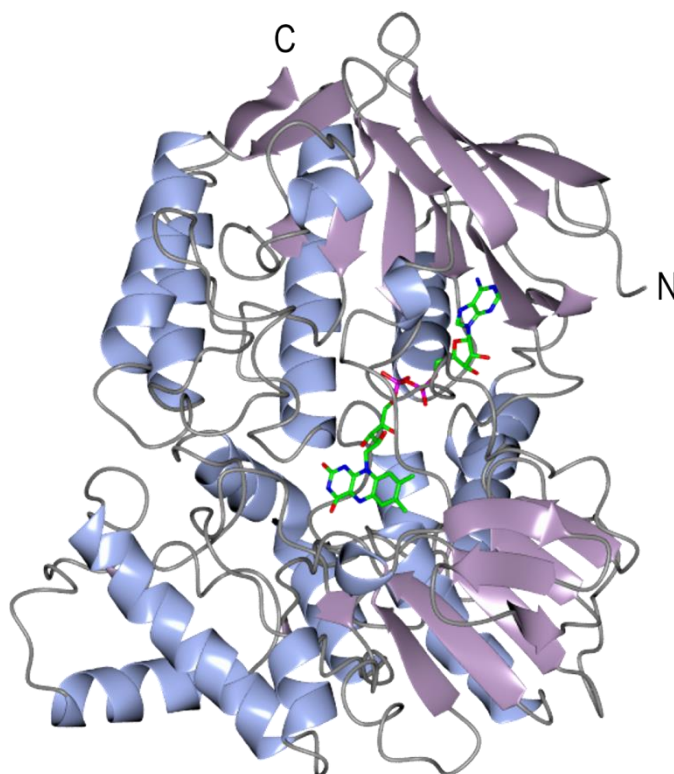
Figure 5.1 Tertiary structure of PAMO complexed with FAD.
Protein backbone is shown in ribbon format, with β-strands in purple and α-helices in blue. The FAD molecule is shown in cylinder format with carbon atoms in green.

The PAMO structure revealed the overall fold of a BVMO for the first time and illustrated the position and conformation of FAD, which remains present throughout

catalysis. The structure of a second BVMO, CHMO from a mesophilic species of *Rhodococcus* sp., complexed with NADP[+] and FAD was also determined (Figure 5.2).[59]
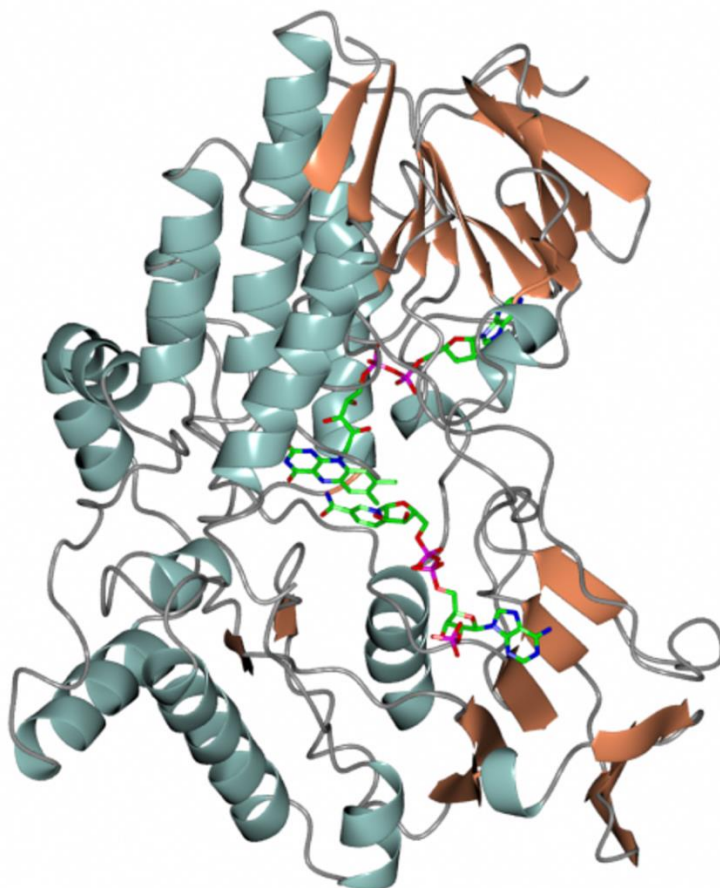
Figure 5.2 Tertiary structure of CHMO complexed with FAD and NADP[+].
The protein backbone is shown in ribbon format, with β-strands in orange and α-helices in blue. The FAD and NADP[+] molecules are shown in cylinder format with carbon atoms in green.

The structure was shown to consist of two domains, including both FAD and NADPH binding domains, with the active site present in a cleft at the domain interface. In addition, it was found that the conserved BVMO motif is on the surface of the enzyme, away from the implied active site. It was suggested that this fingerprint sequence may be involved in conformational changes of the protein during the catalytic cycle.[59] Recently a structure of CHMO complexed with FAD, NADP[+] and its substrate, cyclohexanone has been determined and highlights how NADP[+] has the ability to rotate in the active site in order to allow substrate access to the reactive flavin peroxyanion.[104]

Structure determination is an important piece of the puzzle when it comes to characterising proteins. Obtaining structures of enzymes allows for their unique features to be identified visually and therefore the reasons behind their enzymatic characteristics, for example, can be explored. By comparing defined structures of proteins, such as PAMO, with new unique enzymes such as SMFMO allows, for example, the rational design of mutants to investigate the residues responsible for enzymatic activity. Therefore, in order to obtain a structure for SMFMO and thus investigate the possible molecular determinants for cofactor promiscuity, SMFMO was subject to crystallisation.

5.2 Aims

The aim of this chapter was to further investigate the unique characteristics of SMFMO using structural approaches.

5.3 Materials and methods

5.3.1 Calibrated size exclusion

A Superdex® 75 gel filtration column (10 mm × 30 cm) from GE Healthcare was employed.

Preparation of gel column

The column was equilibrated with filtered H_2O (60 mL) followed by buffer A (60 mL).

Sample loading and gel filtration

A protein injection loop (0.5 mL) was attached to the AKTA-FPLC and washed with H_2O (5 mL) and buffer A (5 mL) using a syringe. Purified SMFMO (0.5 mL) was injected on to the loop and loaded onto the column. The FPLC was set up to run buffer A (60 mL) at a flow rate of 0.5 mL min^{-1}.

5.3.2 Liquid chromatography – mass spectrometry (LC-MS)

LC-MS analysis was carried out by Karl Heaton (Chemistry Department, University of York) using boiled extracts of SMFMO against standard samples of FAD and FMN, following an established procedure.[105]

5.3.3 Crystallisation studies

SMFMO was subjected to crystallisation using conditions described in Section 3.7. Data processing and structure solution for SMFMO was conducted by Dr Gideon Grogan. Collection and refinement statistics are shown in Table 5.1.

Table 5.1 Data collection and refinement statistics for SMFMO complexed with FAD and sulfate (figures in brackets refer to the highest resolution shell).

	SMFMO complexed with FAD and sulphate Collected 3rd March 2010
Beamline	ESRF Id142
Wavelength (Å)	0.933
Space group	$P3_2$
Unit cell (Å)	a = 84.002; b = 84.002; c = 103.324
Resolution (Å)	72.75-2.72 (2.79-2.72)
Unique reflections	20682 (1548)
Completeness (%)	100 (100)
R_{merge}	0.13 (0.55)
Multiplicity	3.8 (3.3)
$<I/\sigma(I)>$	7.5 (2.4)
Protein atoms	5056
R_{cryst}/R_{free} **(%)**	18.6 /23.9
r.m.s.d. 1-2 bonds (Å)	0.017
r.m.s.d. 1-3 bonds (°)	1.963
Average B factor (Å^2)	36

5.4 Results

5.4.1 Calibrated size exclusion chromatography

The SMFMO protein was subjected to size exclusion chromatography (Section 5.3.1). Once eluted a graph of appropriate molecular weights (logMW) was plotted against appropriate elution volumes (GE Healthcare Life Sciences) (Figure 5.3). Taking the

elution value of SMFMO to be ~9 mL, the molecular weight was calculated to be 75,683 Da, suggesting the enzyme is a dimer as a SMFMO monomer is 38 kDa in size.
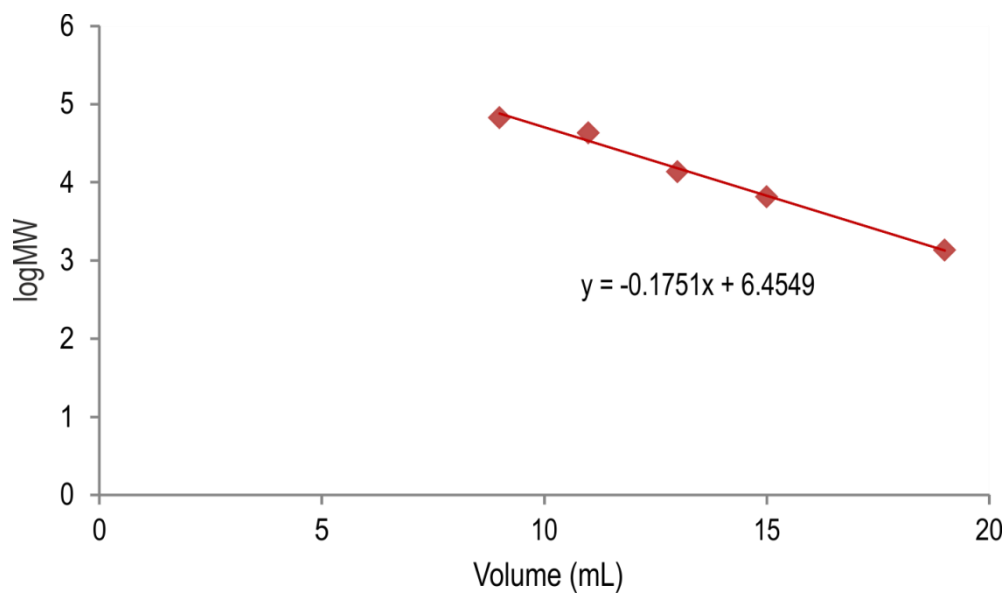
Figure 5.3 Calibrated size exclusion chromatography logMW against elution volume graph.

5.4.2 LC-MS

Using boiled extracts of SMFMO against standards of possible flavin cofactors, FMN and FAD (Section 5.3.2), LC-MS data was obtained which is shown in Figure 5.4.
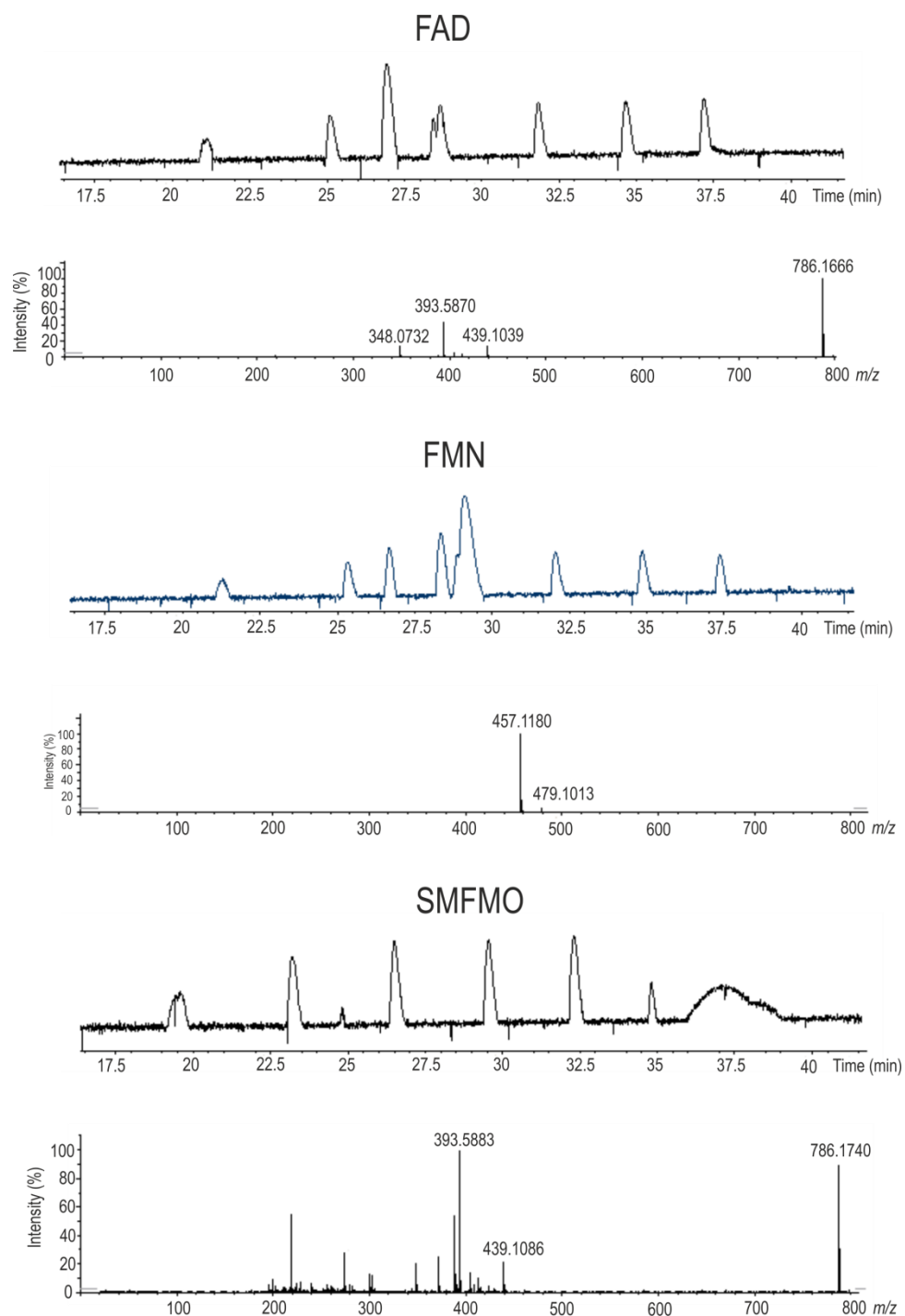
Figure 5.4 LC-MS data for SMFMO and flavin standards FMN and FAD.
The FAD has a *m/z* peak at 786 and FMN gives a *m/z* peak at 457. SMFMO produces a *m/z* peak at 786 also, indicating that FAD is present in SMFMO.

The LC-MS data for the boiled standard sample for FAD gave a *m/z* peak at 786 with high intensity which was expected as the molecular weight of FAD is 785.55 g mol[-1]. For the standard FMN sample a *m/z* peak at 457 was present which again was expected as the molecular weight of FMN is 456 g mol[-1]. For the boiled extract of SMFMO a *m/z*

peak at 786 was present with similar intensity to that obtained for the standard FAD sample. In addition, *m/z* peaks at 393 and 439 were also present in the LC-MS traces for both SMFMO and FAD. However, no *m/z* peaks present in the standard FMN trace were present in the SMFMO trace.

The LC-MS analysis data of boiled enzyme extract against boiled, standard samples of FAD and FMN validated the presence of FAD in SMFMO. Thus confirming, SMFMO is yellow in colour due to the bound flavin.

5.4.3 Structure determination of SMFMO

To investigate the molecular determinants of cofactor promiscuity in SMFMO, the enzyme was subjected to crystallisation. The screens were successful and yellow needle like crystals were produced (Figure 5.5), similar to those formed in the preliminary trials. Interestingly, crystals only appeared in the wells containing sulfate ions. The crystals were subjected to X-ray diffraction and a 2.6 Å dataset was collected as described in Section 3.8. The previous 2.7 Å together with the 2.6 Å dataset allowed for the structure of SMFMO to be solved, built and refined to a resolution of 2.72 Å (Section 5.3.3).
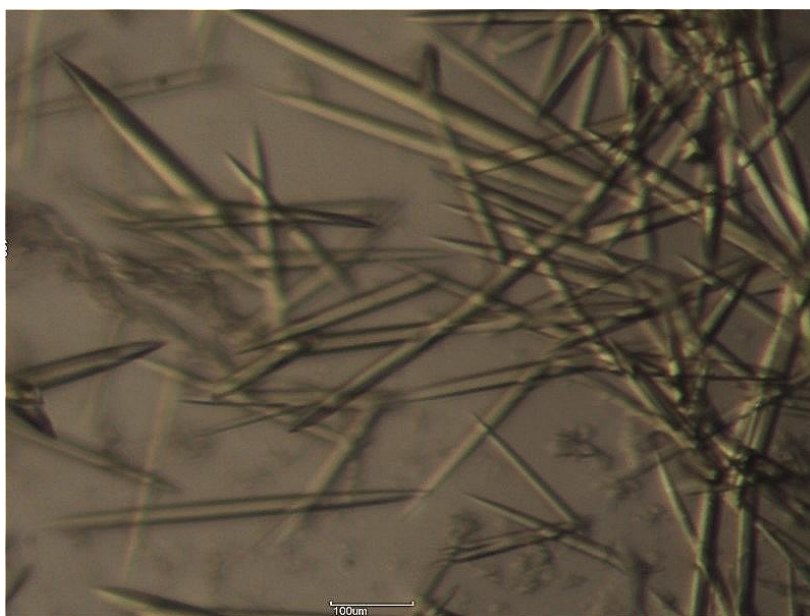
Figure 5.5 Native SMFMO crystals obtained from preliminary Li$_2$SO$_4$ conditions. Conditions: [Li$_2$SO$_4$]: 0.9 M, Buffer: btp pH5.6.

Overall structure of SMFMO

Crystals of SMFMO grew in the $P3_2$ space group, with two molecules, 'A' and 'B', representing one dimer in the asymmetric unit (Figure 5.6), which is in accordance with the SEC results. There was electron density for most amino acids in each monomer from Met1 to Ala352 with a stretch of missing density corresponding to twenty two amino acids from positions Ala212 to Gly233 (ATERWKAQQEGREPDLPPGGFG) that could not be modelled. It is thought that the missing residues may be owing to the flexibility of a loop that sits over the active site containing the bound flavin (Figure 5.7). The SMFMO dimer was made up of two monomers, sharing an interfacial area of approximately 680 Å^2. Analysis of SMFMO using PISA[106] found that the interactions that stabilise the dimer included four hydrogen bonds between the backbone carbonyl of residue Ala146(A) and the NH of Arg101(B) (3.48 Å) and the NH_2 of Arg110(B) (2.46 Å), the main chain carbonyl of Gly147(A) interacts with the NH_2 of Arg101(B) (3.28 Å) and between NH_2 of residue Arg101(A) and backbone carbonyl of Gly47(B) (3.85 Å).
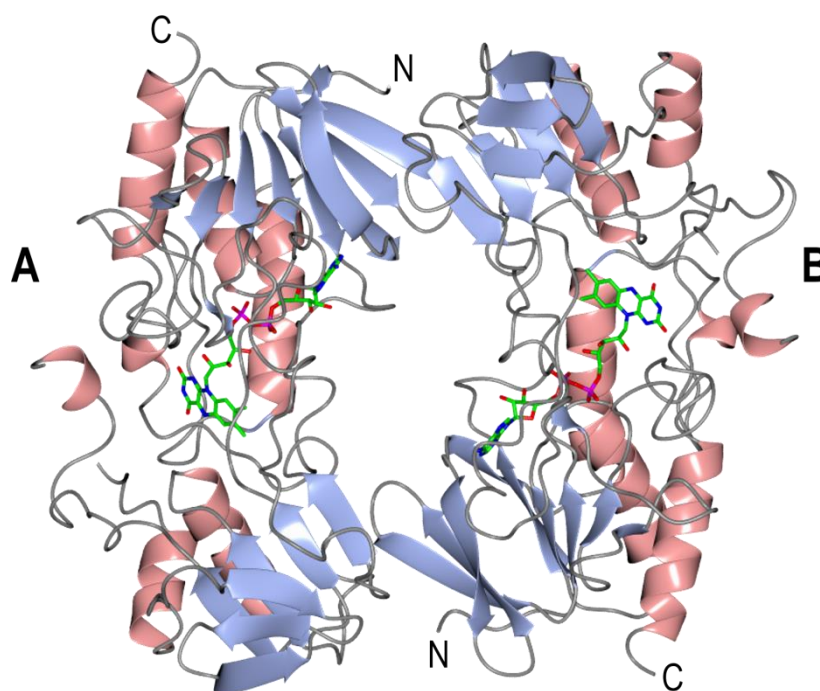
Figure 5.6 Quaternary structure of SMFMO showing the two subunits A and B. Protein backbone is shown in ribbon format, β strands in blue and α helices in pink. The FAD molecules shown in each subunit shown in cylinder format, with carbon atoms in green. The N and C termini for both subunits are also illustrated.

Each SMFMO monomer consisted of fourteen alpha helices: α1 (residues Gln13-Arg24), α2 (Gly40-His44), α3 (Ala55-Trp57), α4 (Arg75-Lys88), α5 (Trp131-Glu133), α6 (Leu142-Ser144), α7 (Ser152-His154), α8 (Ala159-Phe161), α9 (Asn173-Val182), α10 (Arg206-Phe209), α11 (Pro239-Ala246), α12 (Ser288-Leu290), α13 (Asp324-Asn326) and α14 (Val336-Tyr350) and fifteen beta strands: β1 (residues Asp2-Ile9), β2 (Tyr29-Leu32), β3 (Val93-Leu94), β4 (Val99-Phe105), β5(Arg108-Ala113), β6 (Gln118-Ser126), β7 (Ile148-His151), β8 (Arg165-Ile169), β9 (Glu187-Ile191), β10 (Ala252-Val253), β11 (Arg258-Ser260), β12 (Gly263-Gln265), β13 (Glu271-Ala273), β14 (Ala276-Trp279) and β15 (Val317-Leu319). The secondary elements for SMFMO are summarised in Figure 5.8. The β-strands form three distinct beta sheet motifs A (β1-β6, β15), B (β7- β 10, β14) and C (β11- β13). At the N terminus of SMFMO there is a high β-strand presence with the largest of the β – sheets (A) surrounding the terminus. The C terminus on the other hand is highly helical, as shown by the tertiary structure in Figure 5.7. The secondary structure of SMFMO also illustrates the missing structure between helices α10 and α11, representing the missing residues corresponding to the flexible loop amino acids 212 to 233.
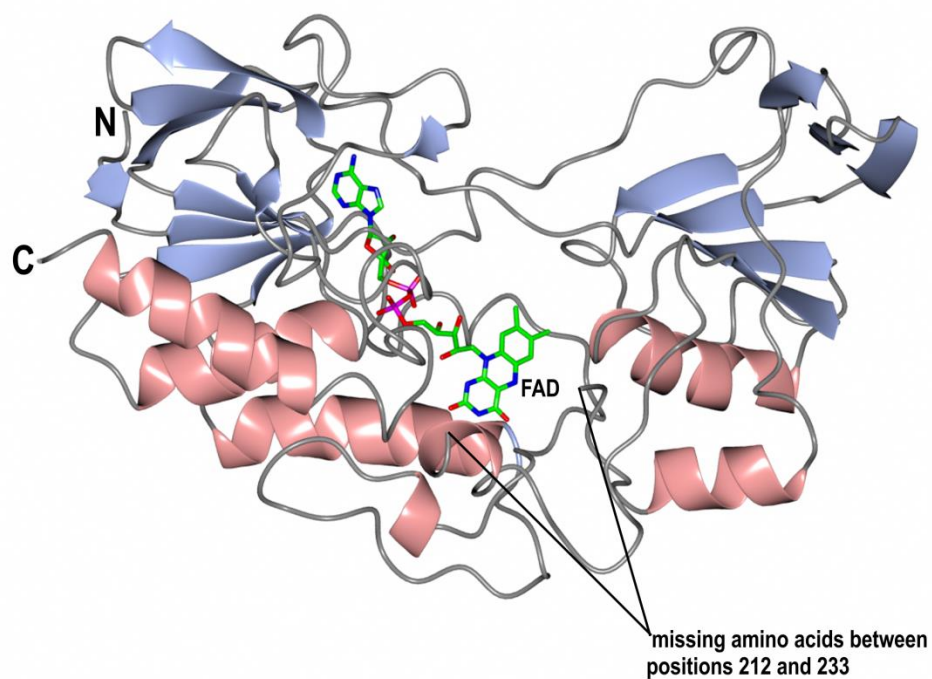
Figure 5.7 Tertiary structure of SMFMO monomer A with complexed FAD.
The protein backbone is shown in ribbon format with β-strands in blue and α-helices in pink. The FAD molecule is shown in cylinder format with the carbon atoms in green. The N and C termini are also illustrated.
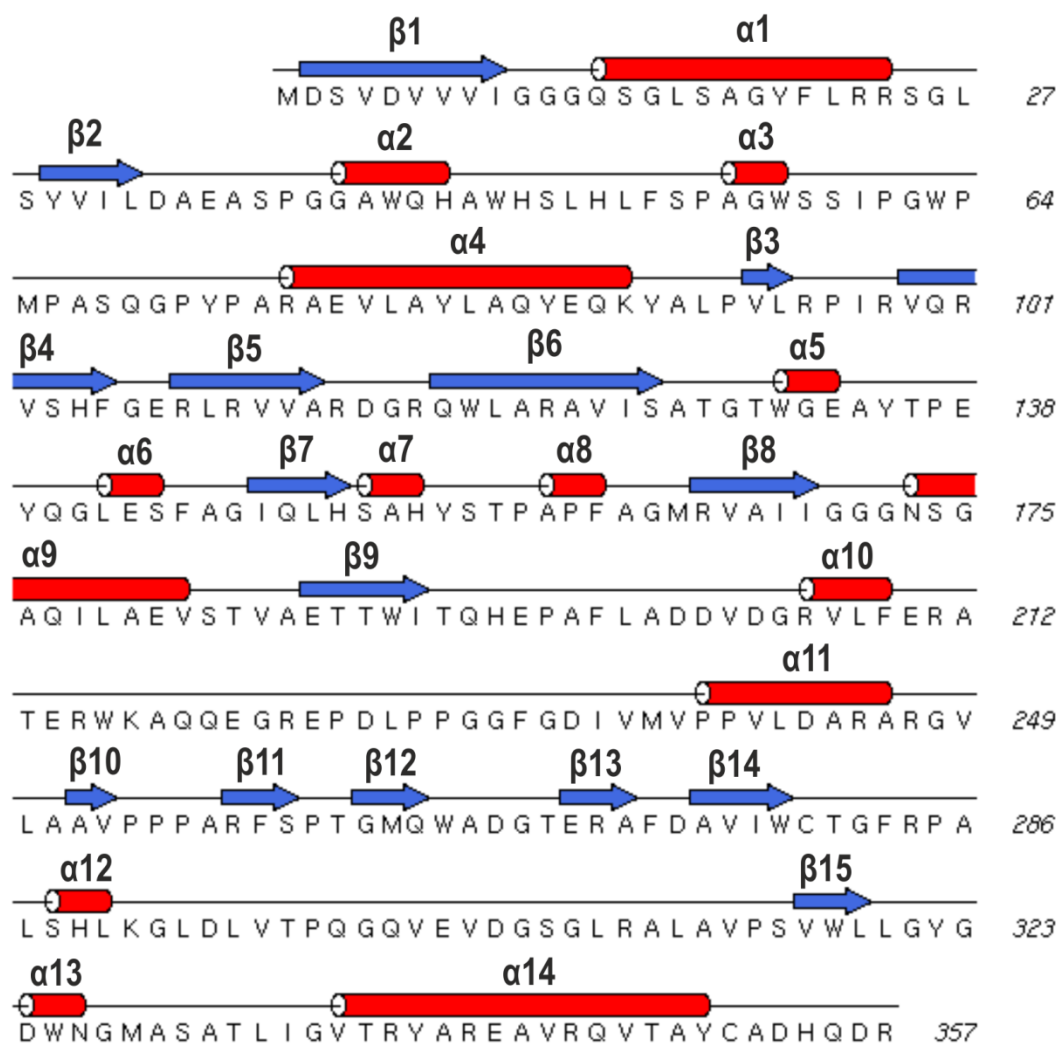
Figure 5.8 Diagram illustrating secondary structure of SMFMO.
The α-helices are shown as red cylinders and the β-strands are shown as blue arrows.

Active site of SMFMO

A monomer of SMFMO consists of two domains, a FAD binding domain and what is thought to be the substrate binding domain. Each monomer of SMFMO has a molecule of FAD present within the active site. The FAD is bound in the catalytic site by 11 hydrogen bonds, shown in Figure 5.9. The side chain N-H of Gln13 interacts with the oxygen (O2', 2.8Å) closest to the tricyclic ring present in the riboflavin moiety. Gln13 also hydrogen bonds to the phosphate oxygen (O1P, 3.05 Å). Hydrogen bond interactions between the gamma oxygen and main chain nitrogen of residue Ser14 and the phosphate carbonyl and hydroxyl oxygens (O1P and O2P, 2.66 Å) also occur. Residues Gln13 and Ser14 are part of the Rossman motif in SMFMO, present in α1, and are responsible for the binding of the ADP motif in FAD. Other hydrogen bonding

interactions involved in securing FAD in the active site are the backbone N-H of Ala41 to the phosphate carbonyl (O1A, 2.65 Å), side chain oxygens of Glu35 to ribose O-H (O2B, 3.75 Å), the backbone N-H and carbonyl oxygen of Val99 to the nitrogen and N-H on the adenine moiety (N6A, 3.08 Å), the main chain N-H of Leu333 to the carbonyl on the tricyclic ring (O2, 2.72 Å) and the backbone N-H of residue Phe52 (O4, 2.92 Å).
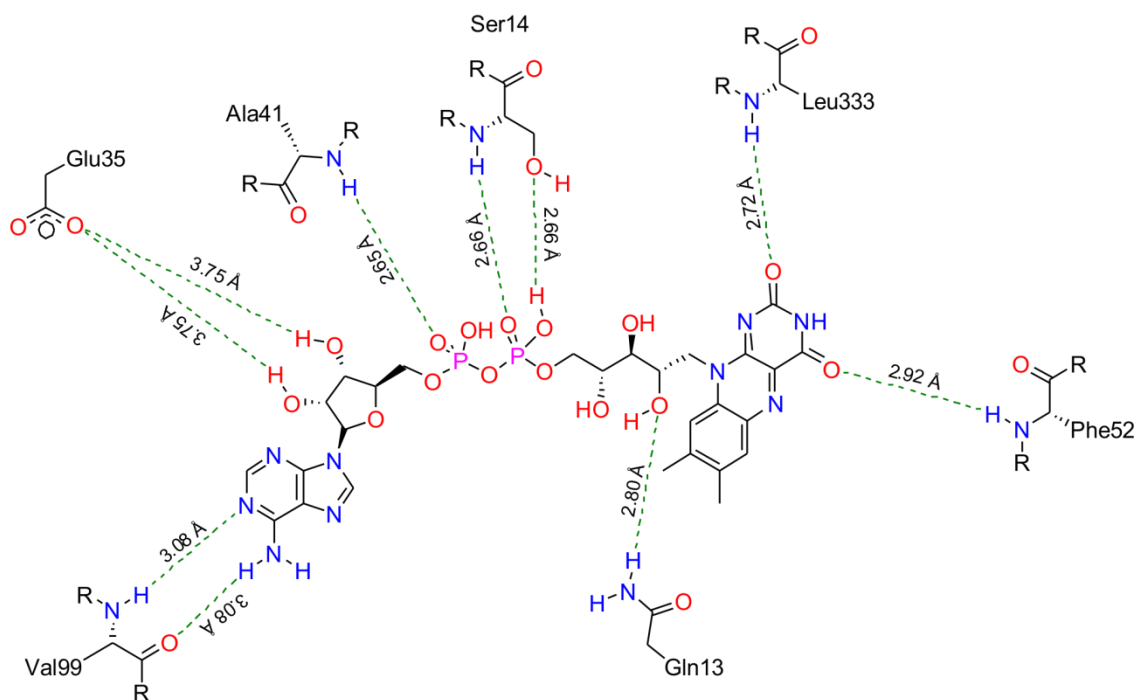
Figure 5.9 Illustrating the residues involved in binding FAD in the active site of SMFMO.

Within FAD the moiety of most interest is that of the tricyclic ring, as it possesses the C_4 carbon atom that reacts with molecular oxygen to become the active catalyst in the oxidation reactions. This prompted the search for residues within 4 Å of this moiety in order to find what other residues are present within the active site, Figure 5.10.
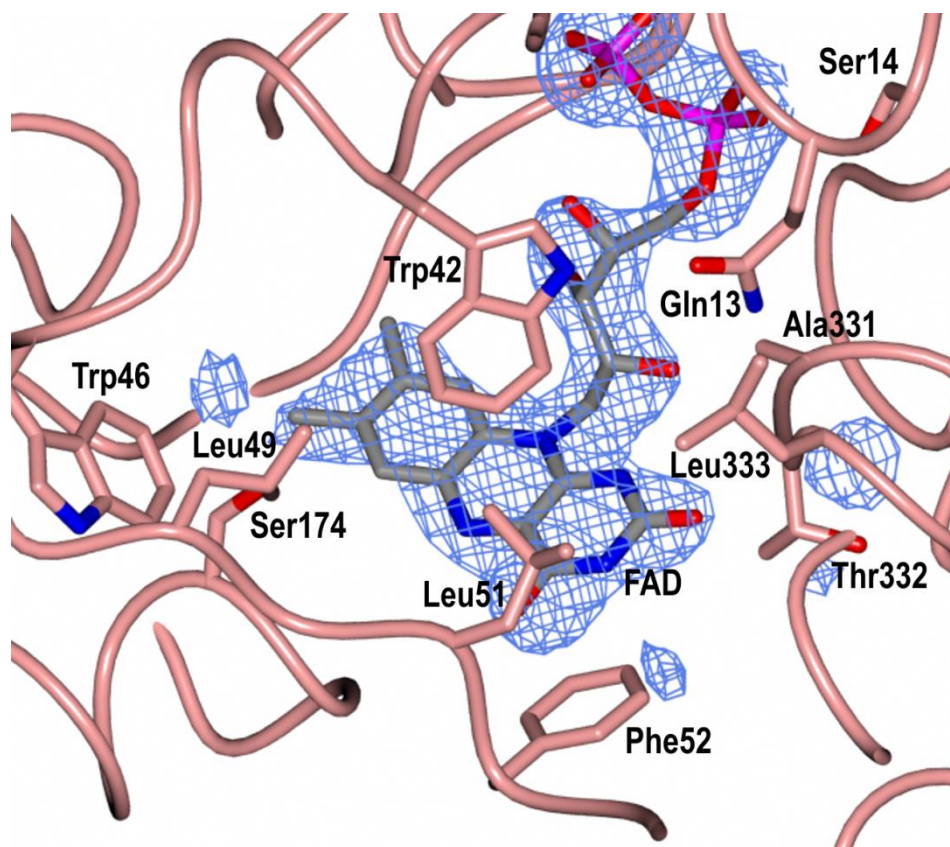
Figure 5.10 The FAD environment within the active site of SMFMO.
The peptide backbone is shown in pink. Amino acid residues within a 4 Å distance of
the FAD are shown in cylinder format with the carbon atoms in pink. The FAD is
shown cylinder format with carbon atoms in grey, surrounded by the F_o-F_c density map
in blue contoured at a level of 3σ, generated in the absence of FAD.

It was found that active site of SMFMO is predominantly hydrophobic with residues
such as tryptophan (Trp46, Trp42), leucine (Leu49, Leu51) and phenylalanine (Phe52)
surrounding the tricyclic ring, with the residues Leu5l and Phe52 sitting closest to the
C_4 carbon atom. The tryptophan residue in position 42 sits close to the dimethylbenzene
ring of FAD which may help to stabilise the FAD through π-π interactions.

5.5 Discussion

The structure of SMFMO was analysed by the DALI server[107] and showed that
SMFMOs monomeric tertiary structure of SMFMO is most similar to that of the
putative FMO 3D1C from *Staphylococcus aureus* MU50 (16% sequence identity; rmsd
2.4 Å over 358 Cα atoms), the BVMO cyclohexanone monooxygenase (3GWF[59], 20 %,
rmsd 3.7 Å over 316 Cα atoms) and the thioredoxin reductase enzymes from *Thermus*

thermophilus (2ZBW, 22 %, rmsd 7.1 Å over 289 Cα atoms). SMFMO also shares structural motifs with known FMOs such as mFMO from the methylotrophic bacterium *Methylophaga* sp. strain SK1[103] (2XLT, 19 %, rmsd 2.5 Å over 293 Cα atoms; 2XLP 19 %, rmsd 2.5 Å over 292 Cα atoms), shown in Figure 5.11. mFMO has been employed to catalyse the asymmetric oxidation of a range of prochiral sulfides to give the majority of (*S*)-enantiomers with moderate to high *ee*, in contrast to SMFMO. Structural studies of mFMO have revealed the molecular determinants responsible for the binding of the 2' ribose phosphate of NADPH, and those discriminating between NADPH and NADH (2XLT/P[103], 2XVJ[74]), and these structures will be used in comparison with SMFMO to further investigate the cofactor promiscuity in SMFMO.
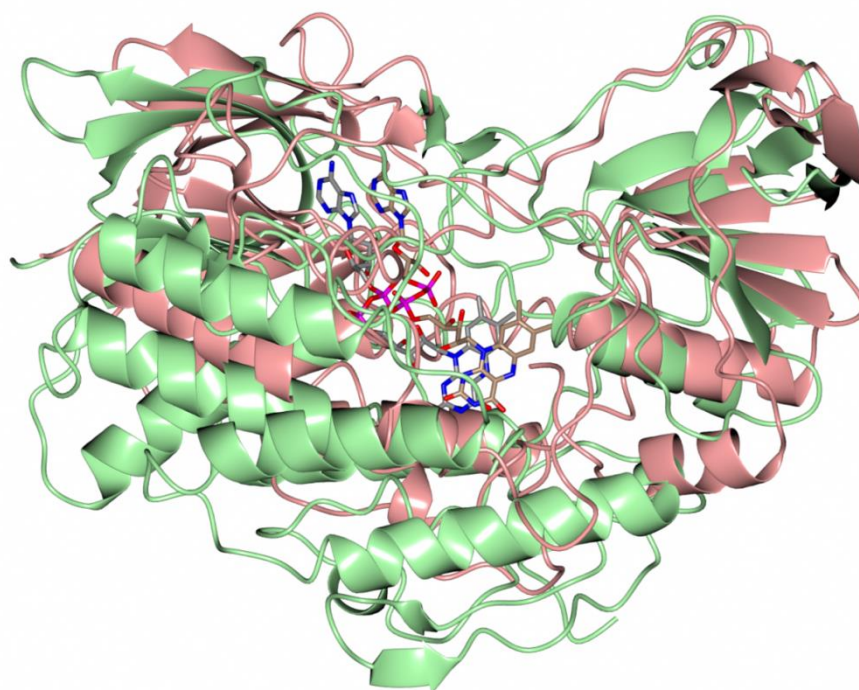
Figure 5.11 Tertiary structure of SMFMO monomer A superimposed with mFMO monomer B.
Protein backbone is shown in ribbon format. SMFMO is shown in pink with the FAD carbon atoms in gold. mFMO is shown in green with the FAD carbon atoms in grey. The superimposition of SMFMO and mFMO was achieved using the ccp4mg programme.

The superimposition of SMFMO with mFMO indicates that the active site of both flavoproteins are in similar positions and some helices and beta sheets are conserved within both structures, owing to the conserved FMO motif and two Rossman motifs responsible for binding the ADP moiety of the FAD and NADPH, typically observed in

class B FMOs. Figure 5.11 also confirms that SMFMO is significantly smaller than the class B bacterial FMO[11], with mFMO being noticeably larger towards the C terminus than SMFMO.

Each SMFMO monomer is complexed with FAD, however, the structure of SMFMO complexed with NAD(P)H in the active site could not be determined. Interestingly, the putative binding sites of the nicotinamide cofactor phosphates, including the 2-hydroxyl ribose phosphate, distinguishing NADPH from NADH, are occupied with sulfate, present in the crystallisation conditions. Superimposition of SMFMO with the NADPH dependent mFMO structure (2XLT[103]) complexed with FAD and NADPH (Figure 5.12), found that the residues apparently responsible for phosphate binding in mFMO Arg234 (and Thr235) are replaced by glutamine Gln193 (and His194) in SMFMO. The relaxation in nicotinamide cofactor in SMFMO may be attributed to the removal of the interactions between the positively charged arginine and negatively charged oxygens on the phosphate.
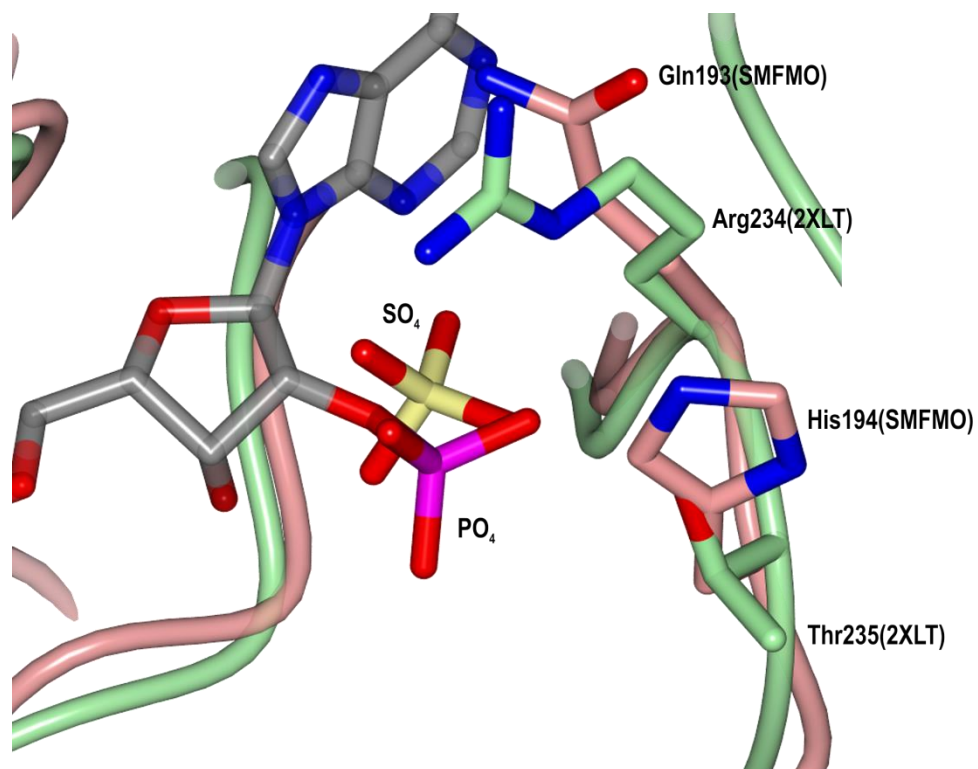
Figure 5.12 Superimposition of NAD(P)H binding site of mFMO(2XLT) and SMFMO. NADP ribose 2'-phosphate recognition site in mFMO (2XLT: back-bone, side chains shown in green and NADP carbon atoms in grey), superimposed with structurally homologous regions of SMFMO (pink) bound to sulfate, illustrating replacement of Arg234 and Thr235 in NADPH-dependent mFMO with Gln193 and His194 in SMFMO.

Within the FPMO family NADP$^+$ has been found to stabilise the flavin peroxidate species formed after flavin reduction and thus reaction with oxygen.[103] In the previous findings SMFMO favours NADH for catalysis rather than NADPH indicating that such stabilisation may be better achieved by the nonphosphorylated cofactor. In mFMO the oxygenating species is reportedly stabilised by Asn78, and activity is removed upon mutation to serine.[103] Within SMFMO this position is occupied by the large, non-polar Phe52 (Figure 5.10). However, reports suggest that hydrophobic residues in this region can be tolerated by flavin-dependent oxygenases and activation depends on the context of the active site.[108]

It is more difficult to explain why SMFMO is able to catalyse the Baeyer-Villiger oxidation of substrate **1**. The hydroperoxidate known to be the oxidising species in BV reactions is thought to be stabilised by an arginine residue that is conserved in all BVMOs, Arg337 in PAMO, but this is not present in SMFMO.

5.6 Conclusion

The structure of SMFMO was solved and analysis of the structure allowed a basis for cofactor promiscuity in SMFMO to be proposed. In addition, the residues immediately surrounding the flavin coenzyme were identified, therefore allowing the determinants of enantioselectivity of sulfoxidation by SMFMO to be investigated.

Chapter 6: Investigation into the cofactor promiscuity of SMFMO

6.1 Introduction

The ability of SMFMO to catalyse asymmetric sulfoxidation reactions and the BV oxidation of substrate **1,** using either NADH or NADPH as cofactor, prompted the investigation into the structure of SMFMO. The structure of SMFMO was found to exist as a dimer, with one molecule of FAD in each subunit. The structure was observed to be similar to mFMO, a monooxygenase which also had the ability to catalyse sulfoxidation reactions.[103] NAD(P)H could not be observed in the active site of SMFMO, however sulfate ions from the crystallisation conditions were bound to the NAD(P)H phosphate binding sites, identified by superimposition of SMFMO with the mFMO-NADPH complex (2XLT[103])(Figure 5.12). The superimposition highlighted that the cofactor promiscuity in SMFMO may be due to the substitution of two residues in mFMO, Arg234 and Thr235 (responsible for binding the 2' phosphate on the NADPH ribose) for a glutamine (Gln193) and histidine (His194) in SMFMO. In addition, the structures of phenylacetone monooxygenase (PAMO[61a]) and cyclohexanone monooxygenase (CHMO) from *Rhodococcus*[59] revealed that both had an arginine/threonine couple (PAMO: Arg217, Thr218; CHMO Arg209, Thr210) present in the cofactor binding pocket that are essential for binding the NADP ribose 2'-phosphate. Therefore mutants Gln193Arg, His194Thr and the double mutant Gln193Arg/His194Thr were constructed in order to investigate if these residues were in fact responsible for phosphate binding and thus responsible for cofactor specificity in SMFMO.

In addition, there is evidence to suggest that glutamate is partly responsible for NADH specificity in FDH enzyme and upon mutation to a glutamine this specificity is altered to also accept NADPH. Therefore the mutant Gln193Glu was constructed in an effort to produce a NADH-specific enzyme.

6.2 Aims

The aim of this chapter was to use site-directed mutagenesis to obtain the SMFMO mutants Gln193Arg, His194Thr Gln193Arg/His194Thr, express the mutant genes, and purify the enzymes, in order to characterise their activity and substrate specificity by biochemical techniques used in the characterisation of SMFMO.

6.3 Materials and Methods

6.3.1 Site directed mutagenesis

Site-directed mutagenesis reactions were carried out with two sets of primers, forward and reverse, using the method described in Section 2.5.1. Primers for the SMFMO mutants Gln193Arg, His194Thr, Gln193Arg/His194Thr and Gln193Glu can be found in Table 3.15. The double mutant was obtained by site-directed mutagenesis on the Gln193Arg SMFMO mutant.

6.3.2 Expression and purification

The SMFMO mutants were expressed in *E. coli* strain BL21 (DE3) as described in Section 3.2.5. SMFMO was purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography described in Section 3.3.

6.3.3 Enzyme assays

Kinetic assays

Kinetic studies of mutants Gln193Arg, His194Thr and Gln193Arg/His194Thr were carried out in accordance to Section 3.4.2. Kinetic parameters were calculated using the GraFit data analysis software.

Biotransformations with SMFMO

Biotransformations were carried out as described in Section 3.4.4 using substrates **1** and **13 – 19**. Biotransformations were analysed by GC and GC-MS described in Sections 3.4.5.

6.3.4 Crystallisation studies

Pure SMFMO mutant Gln193Arg/His194Thr was crystallised using conditions described in Section 3.7. Data processing and structure solution for SMFMO was conducted by Dr Gideon Grogan. Collection and refinement statistics are shown in Table 6.1.

Table 6.1 Data collection and refinement statistics for Q193R/H194T double mutant of SMFMO (figures in brackets correspond to data corresponding to the highest resolution shell).

	Q193R/H194T double mutant of SMFMO
	Collected 7th June 2011
Beamline	Diamond IO3
Wavelength (Å)	0.97630
Resolution (Å)	59.78-2.60 (2.65-2.60)
Space group	$P3_2$
Unit cell (Å)	a = 170.54; b = 170.54; c = 101.80
	$\alpha = \beta = 90.00; \gamma = 120$
No. Molecules in the asymmetric unit	8
Unique reflections	102,301 (6014)
Completeness (%)	97.8 (99.4)
R_{merge} (%)	0.15 (0.56)
$R_{p.i.m}$	0.12 (0.44)
Multiplicity	4.8 (4.6)
$<I/\sigma(I)>$	4.6 (2.1)
$CC_{1/2}$	0.99 (0.71)
Overall B factor from Wilson plot (Å^2)	43
R_{cryst}/R_{free} (%)	25.5/29.5
r.m.s.d. 1-2 bonds (Å)	0.011
r.m.s.d. 1-3 bonds (^{o})	1.675
Average main chain B (Å^2)	50
Average side chain B (Å^2)	51
Average water B (Å^2)	34

6.4 Results

6.4.1 Site-directed mutagenesis

The selected residues in SMFMO, Gln193 and His194, were individually mutated to the residues responsible for phosphate binding in mFMO, Arg234 and Thr235, to yield individual SMFMO genes with the mutations Gln193Arg, His194Thr and Gln193Arg/His194Thr. Residue Gln193 was also successfully mutated to glutamate.

Site-directed mutagenesis was carried out as described in Section 3.5.1 and 6.3.1. A sample from the reactions was run on an Agarose gel (Figure 6.1).
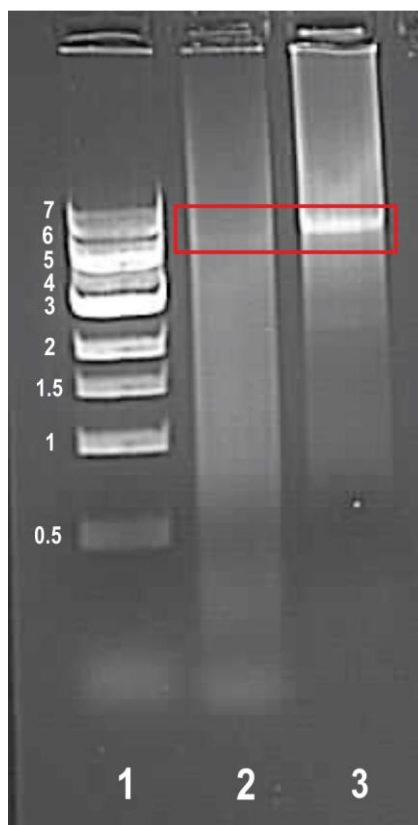
Figure 6.1 An agarose gel of SDM using SMFMO.
Lane 1: NEB 1 kb ladder, lane 2: SDM reaction for SMFMO mutant Gln193Arg, lane 3: SDM reaction for mutant His194Thr. A clear band can be seen at ~6 kb for the His194Thr reaction and a much fainter band for Gln193Arg, outlined in red.

For SDM reactions Gln193Arg, His194Thr, Gln193Arg/His194Thr and Gln193Glu, a clear band of amplified DNA can be seen at ~6 kb (double mutant and Gln193Glu not

shown), illustrating that the reactions were a success. The reactions were digested with DpnI and transformed as described in Section 3.5.1.

The resulting colonies were used to inoculate cultures and the DNA was extracted using a mini-prep kit. Samples of the extracted DNA were sequenced in order to confirm the SMFMO mutation was present and that the rest of the gene sequence was conserved. All three SMFMO mutants were obtained successfully.

6.4.2 Expression and purification of SMFMO mutants

The SMFMO gene, containing the desired mutation, in the pET-YSBLIC-3C vector was expressed in *E. coli* BL21 (DE3) cells as described in Section 3.2.5 with a total volume of 1L LB media. After expression the protein was purified. All three mutants Gln193Arg, His194Thr and Gln193Arg/His194Thr were purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography (Figure 6.2) which was successful in yielding pure yellow protein with a molecular weight of ~38 kDa. The expression of the mutants proved similar to that of the wild type SMFMO yielding ~3 mg of protein per litre of medium.
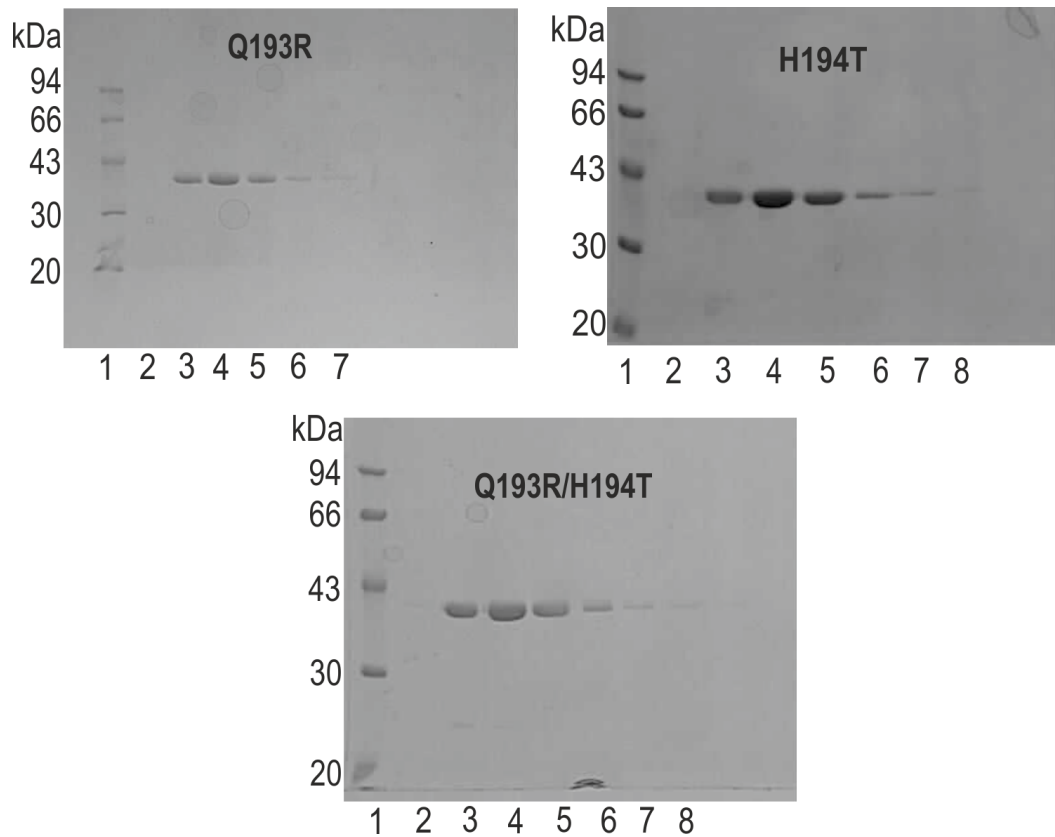
Figure 6.2 SDS-PAGE gel of purification by size exclusion chromatography for SMFMO variants Gln193Arg, His194Thr and Gln193Arg/His194Thr.
Lane 1: Bio-Rad low weight molecular marker, lane 2-8: collected fractions after gel filtration. The mutant proteins can be seen at ~38 kDa

The SMFMO variant Gln193Glu resulted in a mutant that could not be expressed in the soluble fraction of the *E. coli* BL21 (DE3) cells. A band at ~38 kDa can be seen in the insoluble fraction on the SDS-PAGE gel but the band is absent in the soluble fraction (Figure 6.3). The expression of the Gln193Glu mutant was also poor compared to that of the wild type and Gln193Arg, His194Thr and Gln193Arg/His194Thr mutants.
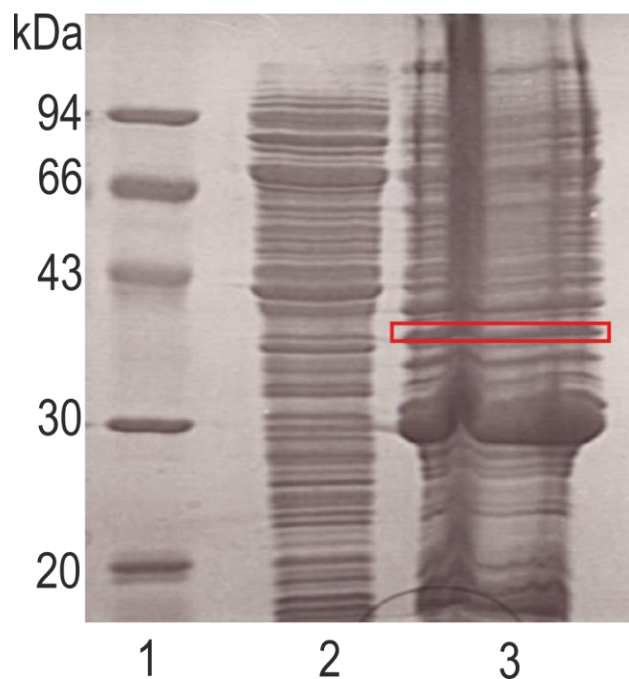
Figure 6.3 SDS-PAGE gel of expression in *E. coli* BL21 (DE3) cells for SMFMO variant Gln193Glu.
Lane 1: Bio-Rad low weight molecular marker, lane 2: soluble fraction, lane 3: insoluble fraction. The mutant protein can be seen at ~38 kDa in the insoluble fraction but the band is not present in the soluble fraction

6.4.3 Kinetic assays with SMFMO mutants

It has been found that SMFMO has the ability to use either NADH or NADPH as nicotinamide cofactor to reduce flavin with equal capability. For the NADH assay, the K_M was 23 μM and Vmax 11.2 $\times 10^{-2}$ μM s^{-1}. Similarly for NADPH assay, the K_M was calculated to be 27 μM and Vmax 8.4 $\times 10^{-2}$ μM s^{-1}. The results were comparable, indicating that the enzyme was able to utilize either cofactor to reduce the bound flavin (Figure 4.7).

To investigate the ability of the mutant enzymes to use either NADH or NADPH to reduce FAD, UV- spectrophotometry assays were carried out as described in Section 3.4.2, using increasing concentrations of either cofactor.

The single mutations Gln193Arg and His194Thr were investigated first. For the Gln193Arg mutation a preference for NADH as cofactor over NADPH was very clear (Figure 6.4). The graph shows that this particular mutation decreases the ability of SMFMO to use NADPH.
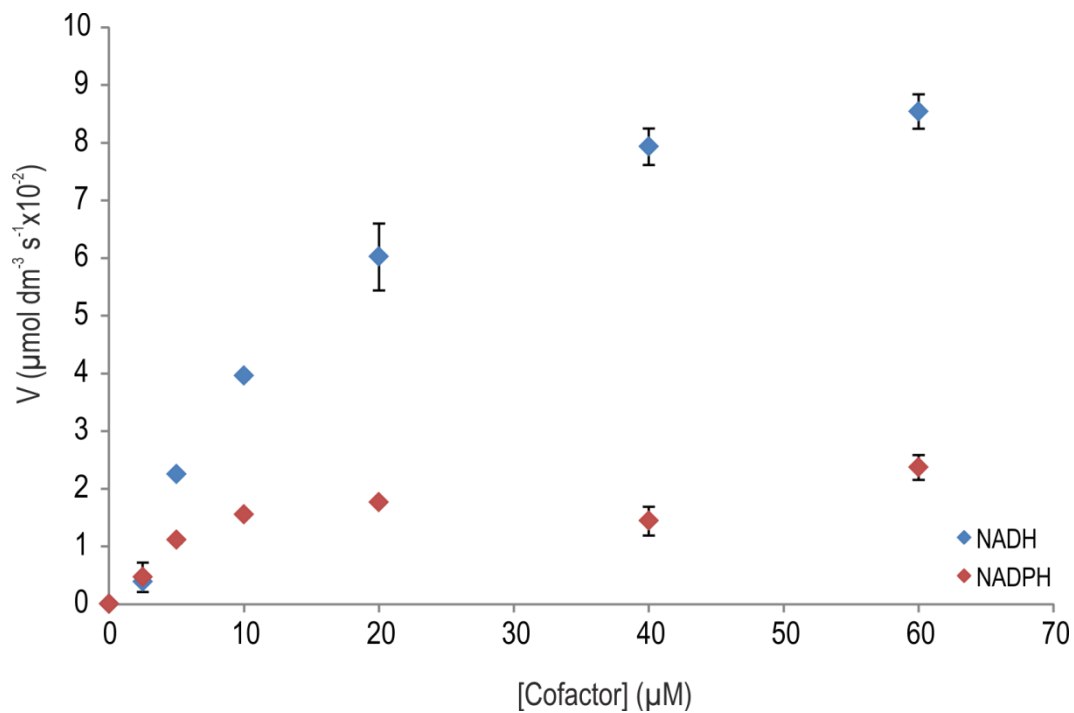
Figure 6.4 Kinetics of cofactor (NAD(P)H) utilisation by SMFMO mutant Gln193Arg.

For the His194Thr mutation, however, the graph indicated a similar trend to that of the wild type SMFMO with both cofactor curves lying in a comparable fashion (Figure 6.5). The mutant seems to mimic that of the wild type in which both cofactors can be utilized equally.
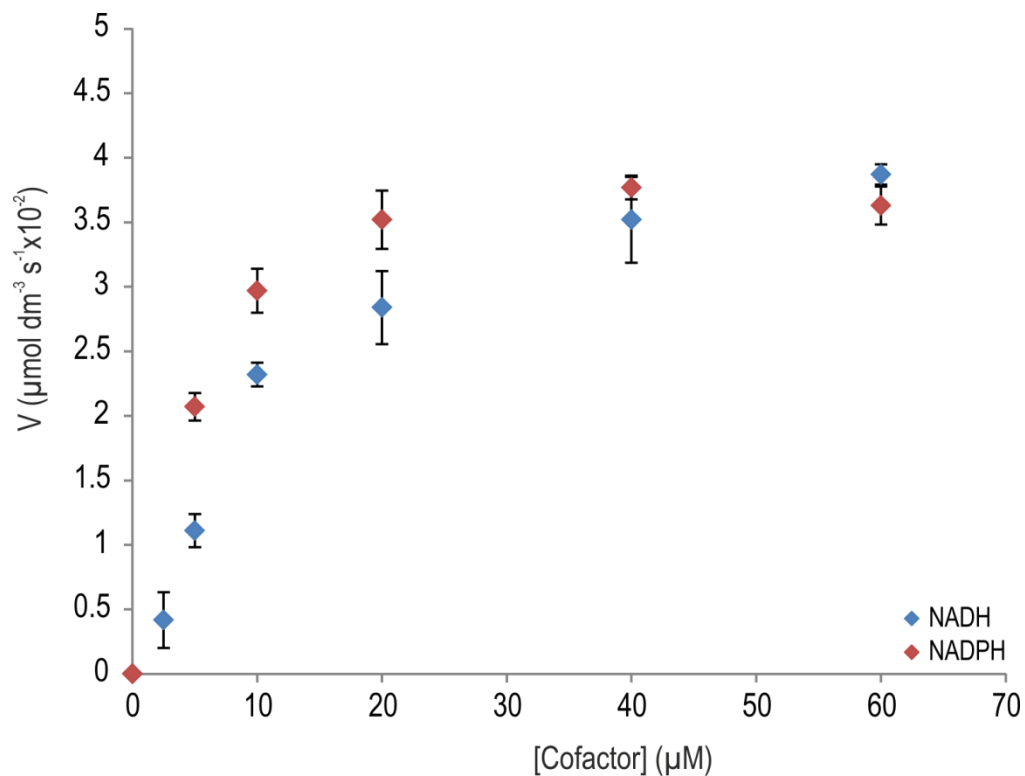
Figure 6.5 Kinetics of cofactor (NAD(P)H) utilisation by SMFMO mutant His194Thr.

To investigate these finding further the double mutation Gln193Arg/ His194Thr was introduced. It can be seen from the data that the mutants seem to coincide with one another with the double mutant's preference for NADH to reduce the flavin over NADPH (Figure 6.6).
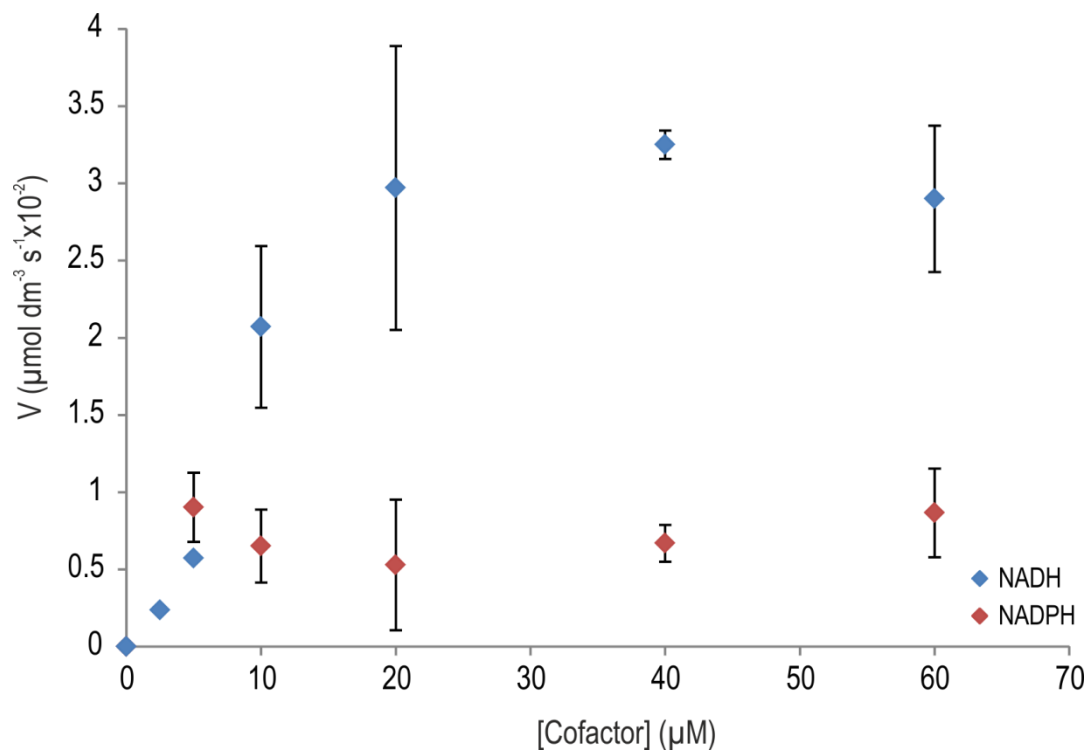
Figure 6.6 Kinetics of cofactor (NAD(P)H) utilisation by double mutant Gln193Arg/ His194Thr.

For each mutant the kinetic graphs demonstrated Michaelis-Menten behaviour, similar to that of the wild type, for both NADH and NADPH assays the kinetic parameters were calculated in order to explore the effects each SMFMO variant had on cofactor oxidation, the results are shown in Table 6.2.

Table 6.2 Kinetic parameters of nicotinamide cofactor oxidation by SMFMO and SMFMO variants Gln193Arg, His194Thr and Gln193Arg/His194Thr.

SMFMO variant	K_M NADH (µM)	V_{max} (µM s^{-1})	k_{cat} NADH (s^{-1})	k_{cat}/K_M NADH (M^{-1} s^{-1})	K_M NADPH (µM)	V_{max} (µM s^{-1})	k_{cat} NADPH (s^{-1})	k_{cat}/K_M NADPH (M^{-1} s^{-1})
Wild-type	23.7 ± 9.1	11.2 ± 1.5 × 10^{-2}	0.029±0.003	1223	27.3±5.3	11.2 ± 1.5 × 10^{-2}	0.022±0.002	805
Gln193Arg	21.2± 4.3	11.9 ± 1.0 × 10^{-2}	0.031±0.003	1476	8.3±1.6	2.65 ± 0.2 × 10^{-2}	0.007±0.0005	875
His194Thr	14.1 ± 2.9	4.82 ± 0.4 × 10^{-2}	0.012±0.0008	857	4.2±0.7	4.09 ± 0.1 × 10^{-2}	0.012±0.0004	3000
Gln193Arg/ His194Thr	12.8 ± 6.9	4.05 ± 0.7 × 10^{-2}	0.01±0.002	769	4.1±2.8	0.8 ± 0.1× 10^{-2}	0.003±0.0003	743

The mutation of Gln193 to arginine resulted in an SMFMO variant with a fourfold reduced K_M for NADPH (8.3 µM compared to 27.3 µM). The turnover, k_{cat} (0.007 s^{-1}), was lower than that of the wild type when NADPH was employed (0.022 s^{-1}), which resulted in the catalytic efficiency (k_{cat}/K_M, 875 M^{-1} s^{-1}) of Gln193Arg to be comparable with that of the wild type. When NADH was employed as cofactor for the Gln193Arg variant the binding (K_M, 21.2 µM) and efficiency (k_{cat}/K_M, 1476 M^{-1} s^{-1}) was similar to that of the wild type (23.7 µM, 1223 M^{-1} s^{-1}). Variant His194 to threonine resulted in a mutant with a sevenfold reduced K_M for NADPH (4.2 µM), compared to the wild type, which resulted in an increased catalytic efficiency (k_{cat}/K_M, 3000 M^{-1} s^{-1}) for this enzyme with NADPH. Compared to SMFMO, binding of NADH for His194Thr was also marginally improved (K_M, 14.1 µM), however, the mutation led to a drop in turnover rates. Therefore, the His194Thr variant led to reduced catalytic efficiency with NADH and increased catalytic efficiency with NADPH, resulting in an overall change of ratio (k_{cat}/K_M NADH/ k_{cat}/K_M NADPH) from 1.5:1 to 1:3.5. Interestingly, the double mutant Gln193Arg/His194Thr gave a combination of effects, in which the lower K_M and lower k_{cat} of mutants Gln193Arg and His194Thr combine to give a mutant of lower

catalytic efficiency overall with both cofactors (NADH k_{cat}/K_M, 769 M^{-1} s^{-1}, NADPH 743 M^{-1} s^{-1}).

6.4.4 Structure determination of double mutant Gln193Arg/His194Thr

The mutants Gln193Arg, His194Thr and Gln193Arg/His194Thr were subject to crystallisation in order to obtain a structure and thus, eliminate the possibility that the mutation of Gln193 to arginine and His194 to threonine had caused any major structural changes in the active site. The crystallisation screens for mutants Gln193Arg and His194Thr were unsuccessful. However, for the double mutant Gln193Arg/His194Thr the screens were successful and yellow needle like crystals were produced (Figure 6.7), similar to those produced for the wild type and also in the same conditions ([Li$_2$SO$_4$]: 0.9 M, Buffer: bis-tris propane pH5.6).
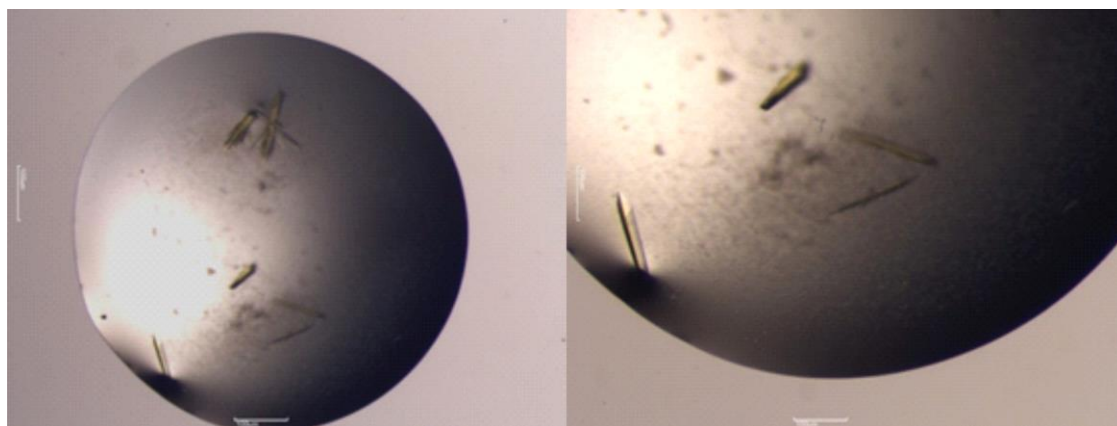
Figure 6.7 Crystals obtained for SMFMO variant Gln193Arg/His194Thr. Crystallisation conditions were the same as for SMFMO: ([Li$_2$SO$_4$]: 0.9 M, Buffer: btp pH5.6.

The crystals were subjected to X-ray diffraction and the structure of the mutant Gln193Arg/His194Thr was solved and refined to a resolution of 2.6 Å, described in Section 6.3.4. The native SMFMO structure was found to exist as a dimer, with two molecules in the asymmetric unit, the double mutant however revealed eight monomers, arranged as four dimers in the unit cell (Figure 6.8). The difference in structures could be credited to less symmetry within the crystal thus leading to a difference in packing within the unit cell. Therefore, for the native structure the symmetry may be much

higher than that of the mutant, leading to only two subunits in the unit cell rather than eight.
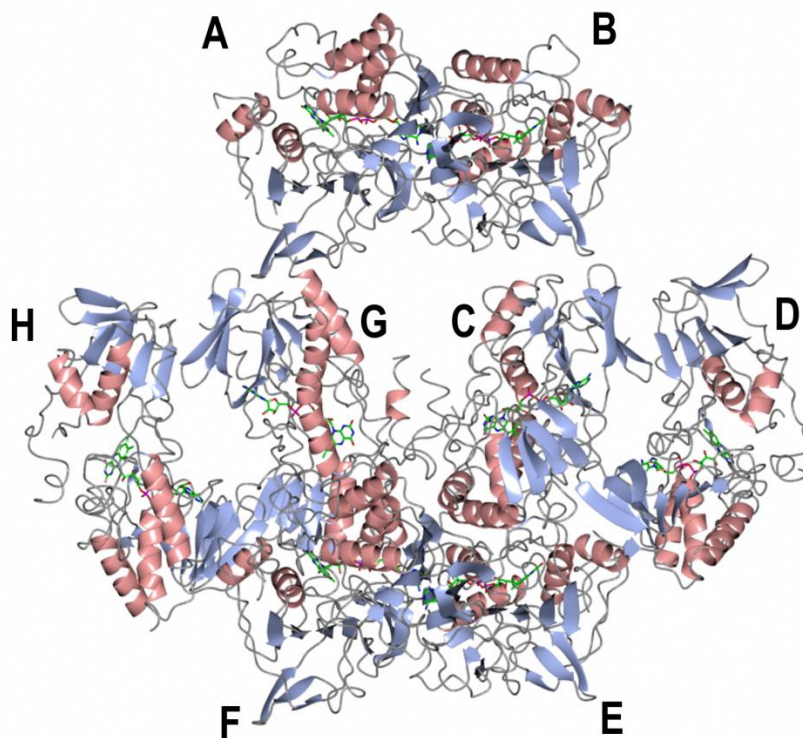
Figure 6.8 Structure of SMFMO mutant Gln193Arg/His194Thr.
Protein backbone is shown in ribbon format, beta sheets in blue and alpha helices in pink. FAD molecules shown in each subunit shown in stick format.

Similar to the wild type, omit maps after refinement in the absence of flavin clearly revealed residual density for the FAD in the active sites of each mutant monomer. As with SMFMO, each subunit was largely complete apart from stretch of missing density corresponding to twenty four amino acids from positions Glu210 to Asp234, which could not be modelled. In order to investigate if the mutations of the Gln193 and His194 residues had affected the cofactor 2' phosphate binding region of the enzyme in any way, subunit 'A' from the wild type and subunit 'H' from the Gln193Arg/His194Thr mutant structure were superimposed (Figure 6.9).
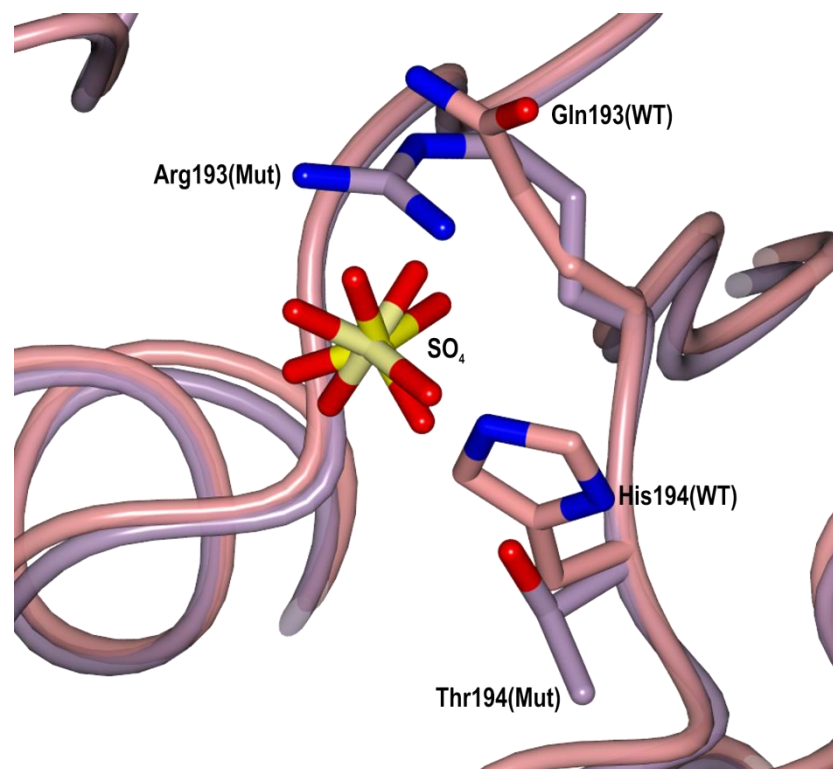
Figure 6.9 Superimposition of nicotinamide cofactor binding site in SMFMO and mutant Gln193Arg/His194Thr structures.
The wild type SMFMO is shown in worm format in pink. The double mutant is shown in worm format in purple. The sulfate ions, for both wild type and mutant, are illustrated in yellow.

The superimposition of the cofactor binding region of the wild type and the mutant revealed very little movement of the backbone and the position of the sulfate ion is conserved in both wild-type and double mutant (rmsd 0.42 Å over 330 Cα atoms). Interestingly, similar to the wild type SMFMO, the two mutant residues Arg193 and Thr194 are pointing inwards towards the sulphate ion. The arginine and threonine of the double mutant are closer to the sulfate ion compared to the glutamine and histadine of the wild type. The Gln193Arg/His194Thr mutant has bond distances of 3.36 and 3.89 Å from the NH of the arginine side chain to the oxygen on the SO_4 and 4.20 Å from the threonine side chain hydroxyl. The wild type however has bond distances of 4.20 Å from the NH on the glutamine side chain to the sulfate ion oxygen and 4.77 and 5.37 Å from the NH on the histidine side chain. Although the bond distances to the sulfate ion of the mutated residues arginine and threonine are nearer than those of the wild type, they are not significantly closer to generate a NADPH-specific mutant.

6.5 Discussion

The SMFMO variants demonstrated that the strength of NADPH binding can be attributed to the nature of the residue at position 193 and 194. However, a straight mutation of the Gln193 and His194 to those found in NADPH dependent FMOs (Arg234 and Thr235 in mFMO) does not engineer absolute NADPH specificity in SMFMO. This may be due to a steric effect in the cofactor binding site, as residual space is still available and can still accommodate the smaller NADH ribose 2' hydroxyl. However, in mFMO the residual space is also present so it is difficult to determine why NADH binding is disfavoured and NADPH binding is favoured in this enzyme. To our knowledge, the activity of mFMO with NADH has not been reported as of yet, so it may be possible that some activity of mFMO with NADH exists. In addition, the arginine/threonine residues in the cofactor binding site in mFMO are conserved throughout other NADPH-dependent flavoproteins, for example the BVMOs PAMO[61a] and CHMO.[59] A study into the nicotinamide cofactor binding site of PAMO revealed that the residue Arg217 was essential for NADPH binding as the mutations to alanine and leucine reduced the catalytic efficiency dramatically. The mutation of the residue Thr218 to alanine, however, did not have a significant effect.[87a] The substitution of His220 to glutamine in PAMO resulted in a variant with an increased catalytic efficiency with NADH but the mutant Gln220 residue is not structurally homologous with Gln193 in SMFMO. While the SMFMO variant Gln193Arg did not dramatically alter cofactor specificity, there is widespread evidence that glutamine is a determinant in cofactor promiscuity. For example, formate dehydrogenases (FDHs) are mostly NAD$^+$ dependent, however mutational[109] and genomic[110] studies have revealed enzymes that are cofactor promiscuous or NADPH-specific which is partly due to the mutation of a glutamate or aspartate residue in NADH-specific enzymes that repel negatively charged phosphate, for a glutamine, which has no charge. This finding prompted the Gln193 residue in SMFMO to be exchanged with a glutamate residue in an effort to produce a strict NADH-specific enzyme. Unfortunately the mutation Gln193Glu in SMFMO resulted in a variant which could not be expressed within the soluble fraction of the *E. coli* cells (shown in Section 6.4.2). The His194Thr variant had a much greater effect on NADPH binding in SMFMO. The increased contribution of the threonine residue to NADPH binding could be owing to the bond distances from the Thr235 side chain to the ribose 2' phosphate oxygen in mFMO, for example, in the 2XLT structure the

Thr235 is closer (2.6 Å) compared to those for the Arg234 side-chain NH atoms with the corresponding phosphate oxygens (3.2-3.5 Å). The sequence alignment of the FMO motif regions of SMFMO, mFMO[103] and PAMO[61a] with the group of cofactor promiscuous *Rhodococcus jostii* RHA1 FMOs[18] using ClustalW2[113] revealed possible residues responsible for the cofactor promiscuity in such enzymes (Figure 6.10).

```
SMFMO    FAGIQLHSAHY-STPAPFAGMRVAIIGGGNSGAQILAEVSTV-AETTWITQHEPAFLADD
mFMO     FGGRILHAHDF-RDALEFKDKTVLLVGSSYSAEDIGSQCYKYGAKKLISCYRTAPMG--Y
PAMO     FAGNLYHTGNWPHEPVDFSGQRVGVIGTGSSGIQVSPQIAKQAAELFVFQRT-PHFAVPA
FMO-A    WRGALVHSSQY-RNPSQYNGKRVLVVGAGCSGMEIAYDLATGGAAKVWLSARTPPNIMLR
FMO-C    PRIRQLHSSDY-RNPSQLQDGPVLVVGCSHSGADIALEASRSHRTTICGPVR-GEVPFDI
FMO-D    FTGRQVHTKDY-SSAQDFAGQHVLVVGGGISAVQLLDEISRV-TTTTWVTRRPPEFRDEP
FMO-E    FAGRTLHSSEY-DDANDFAGQRVVIIGTGNSAHDVAQDLHAHGIDVTMVQRS-STTIVSV
FMO-F    FRGEQHHSSRH-PGPDAYVGKRVVVVGANNSAHDICKALFENGADVTMLQRS-STHIVKS
FMO-G    FEGTIEHSSWF-VGGREMQGKKALVVGCCNSGHDIAQELNEQGADVTILQRS-STYVMSS
```

Figure 6.10 Sequence alignment of flavin containing enzymes with SMFMO. SMFMO = *Stenotrophomonas maltophilia* Flavin-containing Monooxygenase; mFMO from *Methylophaga aminisulfidivirans*; PAMO = phenylacetone monooxygenase from *Thermobifida fusca*; FMO A-G = *Rhodococcus jostii* RHA1 FMOs. The residues responsible for phosphate binding in SMFMO, mFMO and PAMO are highlighted in yellow. The residues that may be responsible for cofactor phosphate recognition for the group of FMOs from *Rhodococcus jostii* RHA1 are highlighted in yellow.

Structural alignment of SMFMO, PAMO and mFMO revealed that the residues responsible for cofactor binding in SMFMO were Gln193 and His194 (an Arg/Thr couple in mFMO and PAMO). Sequence alignment of the *Rhodococcus jostii* RHA1 FMOs, that can utilise both NADH and NADPH as cofactor giving identical conversions and enantioselectivities of substrates, found that FMO E, G and F had an arginine and serine couple (FMO E: Arg370, Ser371; FMO F: Arg369, Ser370; FMO G: Arg379, Ser380), which is not too dissimilar to the arginine and threonine residues responsible for phosphate binding in mFMO and PAMO. SMFMO however is much smaller than the FMOs E, F and G, with the three being closer in sequence to BVMOs. FMO D is similar in size to SMFMO, and has two arginine residues that may be responsible for phosphate binding (Arg206 and Arg207), however it is difficult to pinpoint the positions in the *Rhodococcus* FMOs without a structural alignment. Thus, further structural and mechanistic studies of SMFMO and the group of FMOs identified by Riebel *et al*. could help determine further cofactor specificity determinants in this group of enzymes.

6.6 Conclusion

The SMFMO mutants Gln193Arg, His194Thr and Gln193Arg/His194Thr allowed for the cofactor promiscuity of SMFMO to be investigated. The mutation of the histidine residue to threonine proved to have the greater effect in producing an NADPH-specific enzyme, however, SMFMO was still able to utilise NADH. A structure of SMFMO complexed with NAD(P)H could help shed some light on the enzymes ability to use both cofactors.

Chapter 7: Investigation into the enantioselectivity of SMFMO

7.1 Introduction

Mutagenesis studies towards the active site of BVMOs like PAMO and CHMO and FMOs such as mFMO has highlighted residues that have effects on the activity and enantioselectivity of these enzymes.[58b, 111] SMFMO was found to oxidise substrates with moderate to low enantioselectivity, the best case being that of substrate **19** with 80 % -(*R*) *ee* with NADH and 82 % -(*R*) with NADPH. The structure obtained for SMFMO revealed the environment surrounding the flavin, particularly the tricyclic ring environment which possesses the C4 carbon atom that reacts with molecular oxygen to give the reactive intermediate, (hydro)peroxy flavin, that goes on to oxidise substrates. The structure of SMFMO allowed for a structure-informed approach to the analysis of the active site, in order to improve or alter SMFMOs enantioselectivity, when NADH was employed as cofactor. Cho and co-workers revealed a structure that highlighted the active site of the FMO mFMO. The mFMO structure had been determined in complex with the substrate indole, which sits close to the FAD, in order to undergo an oxygenation reaction to generate indoxyl.[74] To investigate the active site determinants for enantioselectivity in SMFMO, the active site was superimposed with flavin-binding site of the indole bound mFMO structure (2XVJ[74]), Figure 7.1.
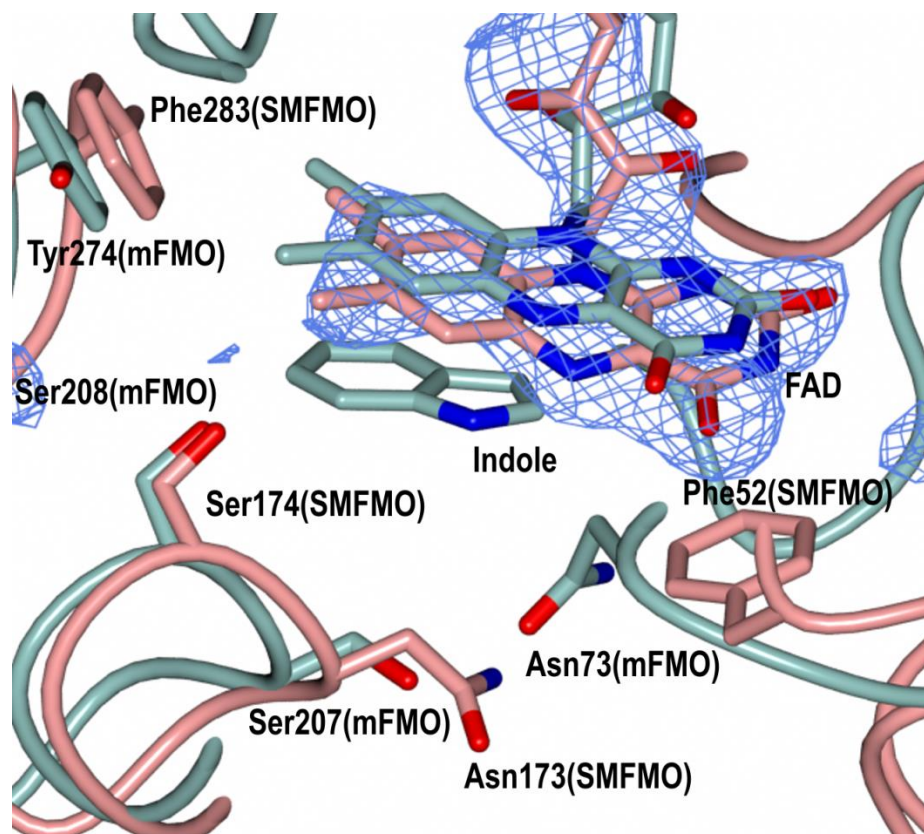
Figure7.1 Superimposition of the flavin binding site of SMFMO with the flavin binding site of mFMO in complex with indole (2XVJ) illustrating the sites picked for mutation to investigate the enantioselectivity of SMFMO.
The protein backbone is shown in worm format with SMFMO in pink and mFMO in green. The FAD for SMFMO is shown in cylinder format with carbon atoms in pink. The FAD and indole for mFMO are shown in cylinder format with carbon atoms in green. The FAD is surrounded by the F_o-F_c density map in blue contoured at a level of 3σ, generated in the absence of FAD.

The selected residues for mutagenesis were Phe52, Asn173, Ser174 and Phe283 as these residues were close to the indole ligand (within 4 Å) and also the C4 position in the tricyclic ring of the flavin which is the site of formation of the reactive hydroperoxy flavin intermediate. In mFMO residue Asn73 is thought to stabilise the flavin hydroperoxide, as mutation results in an inactive enzyme. Therefore residue Phe52 in SMFMO was also mutated to asparagine in an effort to see what effect this would have on the activity of SMFMO. Substrates chosen for the investigation of enantioselectivity in SMFMO were substrates **13** and **15**. Substrate **13** was selected as the conversion with SMFMO and NADH was 27 % with 71 % -(*R*) *ee*, whereas **15** had a conversion of 90 % with 25 % -(*R*) *ee*, thus allowing for an improvement in enantioselectivity to be

monitored. Substrate **19** was not chosen as the *ee* for this substrate when NADH was employed was already high making it difficult for an increase in *ee* to be observed.

7.2 Aims

The aim of this chapter was to use mutagenesis techniques, SSM and SDM, to obtain mutants of the selected residues in order to investigate the effects such mutations had on the enantioselectivity of SMFMO when NADH is employed as nicotinamide cofactor.

7.3 Materials and methods

7.3.1 Mutagenesis

SDM

SDM reactions were carried out with two sets of primers, forward and reverse, using the method described in Section 3.5.1. Primers for the SMFMO mutants Phe52Asn, Phe52Ala, Asn173Phe, Asn173His and Asn173Tyr can be found in Table 3.15.

Site saturation mutagenesis

SSM reactions were carried out with the NDT primer for each specific residue, using the method described in Section 3.5.2. Primers for the SMFMO mutants Phe52X, Asp173X, Ser174X and Phe283X can be found in Table 3.18.

7.3.2 Expression and purification

The SMFMO mutants were expressed in *E. coli* strain BL21 (DE3) as described in Section 3.2.5. Mutants Phe52Asn and Phe52Val were purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography described in Section 3.3.

7.3.3 Enzyme assays

Biotransformations with SMFMO variants

The expressed mutants were sonicated (3×30 s) and the cell lysate (1 mL) was used in the biotransformations of sulfides **13** and **15,** carried out as described in Section 3.4.4. For the mutant Phe52Val, isolated protein was used in the biotransformations (Section

3.4.4) of sulfides **13**–**19** (Figure 4.10). Biotransformations were analysed by GC described in Section 3.4.5. For Phe52Asn isolated protein was used for the biotransformations of substrates **1** and **15**.

7.4 Results

7.4.1 Mutagenesis

The selected residues in SMFMO, Phe52X, Asn173X, Ser174X and Phe283X, were individually mutated using primers with an NDT codon, yielding possible variants with the mutations Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly.[99]

Site-saturation mutagenesis was carried out as described in Section 3.5.2 and 7.3.1. A sample from the reactions was run on an Agarose gel to illustrate that the reactions were a success. The reactions were digested with DpnI and transformed as described in Section 3.5.1.

The resulting colonies were used to inoculate cultures and the DNA was extracted using a mini-prep kit. Samples of the extracted DNA were subjected to restriction digest (Section 3.1.7) to confirm insert was present (Figure 7.2) and sequenced in order to confirm the SMFMO mutation was present and that the rest of the gene sequence was reserved (Figure 7.3). Colonies were continuously picked until many of the mutations were obtained.
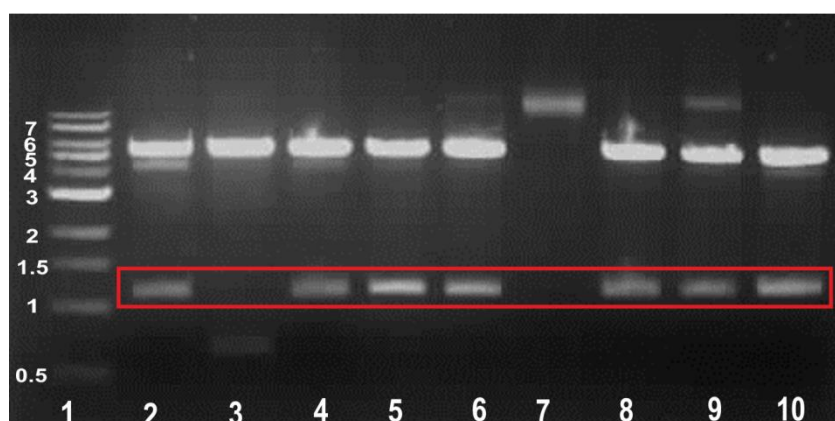
Figure 7.2 Agarose gel of restriction digest for Asp173X DNA samples.
Lane 1: NEB 1 kb ladder, lane 2-10: Restriction digest reactions for SMFMO mutant Asn173X. A clear band can be seen at ~6 kb for the pET-YSBLIC-3C vector and a band at ~1.1 kb for the mutant SMFMO insert, outlined in red.

For SSM reactions Phe52X, Asn173X, Ser174X, clear bands at ~1.1 kb and ~6 kb were present, illustrating that a DNA insert was present within the complete pET-YSBLIC-3C vector.

```
PML_BVMO_B2FR      1 -------------------------------MDSVDVVVIGGGQSGLSA      18
                                                    |||||||||||||||||||
F52N_4.2_f         1 FCLTLRRRYTMGSSHHHHHHSSGLEVLFQGPAMDSVDVVVIGGGQSGLSA     50

PML_BVMO_B2FR     19 GYFLRRSGLSYVILDAEASPGGAWQHAWHSLHLFSPAGWSSIPGWPMPAS     68
                     |||||||||||||||||||||||||||||||||.|||||||||||||||||
F52N_4.2_f        51 GYFLRRSGLSYVILDAEASPGGAWQHAWHSLHLNSPAGWSSIPGWPMPAS    100

PML_BVMO_B2FR     69 QGPYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERLRVVARDGRQ    118
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_f       101 QGPYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERLRVVARDGRQ    150

PML_BVMO_B2FR    119 WLARAVISATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAI    168
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_f       151 WLARAVISATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAI    200

PML_BVMO_B2FR    169 IGGGNSGAQILAEVSTVAETTWITQHEPAFLADDVDGRVLFERATERWKA    218
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_f       201 IGGGNSGAQILAEVSTVAETTWITQHEPAFLADDVDGRVLFERATERWKA    250

PML_BVMO_B2FR    219 QQEGREPDLPPGGFGDIVMVPPVLDARARGVLAAVPPPARFSPTGMQWAD    268
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_f       251 QQEGREPDLPPGGFGDIVMVPPVLDARARGVLAAVPPPARFSPTGMQWAD    300

PML_BVMO_B2FR    269 GTERAFDAVIWCTGFRPALSHLKGLDLVTPQGQVEVDGSGLRALAVPSVW    318
                     ||||||||||||||||||||||||||.      .|.....|.|
F52N_4.2_f       301 GTERAFDAVIWCTGFRPALSHLKGLG----SGYSAGPGGG----------    336

PML_BVMO_B2FR    319 LLGYGDWNGMASATLIGVTRYAREAVRQVTAYCADHQDR       357

F52N_4.2_f       337 --------------------------------------       336


PML_BVMO_B2FR      1 MDSVDVVVIGGGQSGLSAGYFLRRSGLSYVILDAEASPGGAWQHAWHSLH     50

F52N_4.2_r         1 -------------------------------------------------      0

PML_BVMO_B2FR     51 LFSPAGWSSIPGWPMPASQGPYPARAEVLAYLAQYEQKYALPVLRPIRVQ    100
                                                     .:......||||||||
F52N_4.2_r         1 --------------------------------TEIRAACLRPIRVQ       14

PML_BVMO_B2FR    101 RVSHFGERLRVVARDGRQWLARAVISATGTWGEAYTPEYQGLESFAGIQL    150
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_r        15 RVSHFGERLRVVARDGRQWLARAVISATGTWGEAYTPEYQGLESFAGIQL     64

PML_BVMO_B2FR    151 HSAHYSTPAPFAGMRVAIIGGGNSGAQILAEVSTVAETTWITQHEPAFLA    200
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_r        65 HSAHYSTPAPFAGMRVAIIGGGNSGAQILAEVSTVAETTWITQHEPAFLA    114

PML_BVMO_B2FR    201 DDVDGRVLFERATERWKAQQEGREPDLPPGGFGDIVMVPPVLDARARGVL    250
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_r       115 DDVDGRVLFERATERWKAQQEGREPDLPPGGFGDIVMVPPVLDARARGVL    164

PML_BVMO_B2FR    251 AAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSHLKGLDLVTPQG    300
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_r       165 AAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSHLKGLDLVTPQG    214

PML_BVMO_B2FR    301 QVEVDGSGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTAY    350
                     |||||||||||||||||||||||||||||||||||||||||||||||||
F52N_4.2_r       215 QVEVDGSGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTAY    264

PML_BVMO_B2FR    351 CADHQDR       357
                     |||||||
F52N_4.2_r       265 CADHQDR       271
```

Figure 7.3 Sequence alignment of wild type SMFMO and Phe52Asn mutant. Phe52Asn mutation is highlighted in yellow.

134

The SMFMO variants obtained successfully are shown in Table 7.1.

Table 7.1 Table of NDT mutations. *Obtained by site-directed mutagenesis. Green indicates successful mutations, red indicates unsuccessful mutations, purple indicates native amino acid.

NDT Library	Phe52X	Asn173X	Ser174X	Phe283X
Phe	Phe	Phe	Phe	Phe
Leu	Leu	Leu	Leu	Leu
Ile	Ile	Ile	Ile	Ile
Val	Val	Val	Val	Val
Tyr	Tyr	Tyr	Tyr	Tyr
His	His	His*	His	His
Asn	Asn*	Asn	Asn	Asn
Asp	Asp	Asp	Asp	Asp
Cys	Cys	Cys	Cys	Cys
Arg	Arg	Arg	Arg (Q118R)	Arg
Ser	Ser	Ser	Ser	Ser

For the Phe52X mutant SSM gave only four of the eleven available mutations. The Phe52Asn mutant was obtained by SDM. For the Asn173X variants eight mutants were obtained. The variant Asn173His was obtained by SDM. However, the Asn173Tyr and Asn173Phe mutants could not be obtained by SSM or SDM. For the point mutation Phe283X SSM reactions proved unsuccessful.

7.4.2 Expression and purification of SMFMO variants

Phe52X, Asp173X and Ser174X expression

The SMFMO gene, containing the desired mutation, in the pET-YSBLIC-3C vector was expressed in BL21 (DE3) cells as described in Section 3.2.5 with a total volume of 100 mL LB media. The expressed mutants were sonicated and the lysate was analysed by SDS-PAGE in order to find if the expressed mutants were soluble (Figure 7.4).
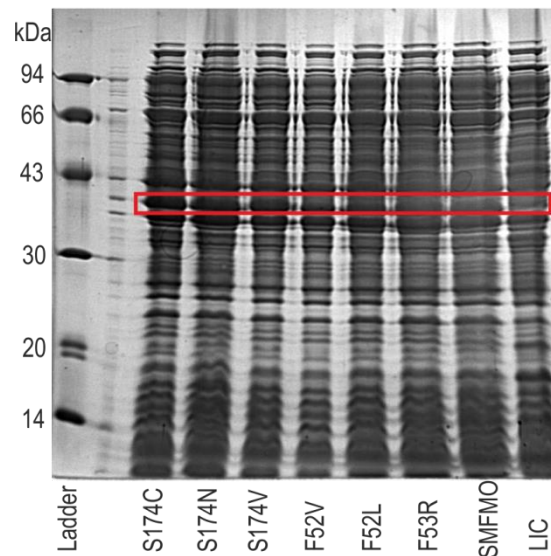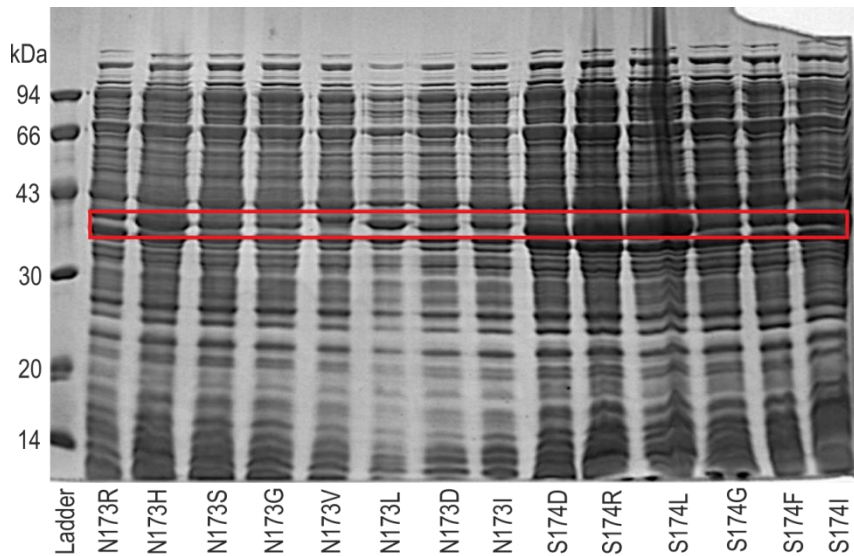
Figure 7.4 SDS-PAGE gel of soluble fractions for SMFMO mutants Phe52X, Asn173X and Ser174X.
Clear expression of mutants can be seen at ~38 kDa for mutants and SMFMO but not for the LIC control.

For each SMFMO mutant soluble expression could be seen at ~38 kDa. A similar band was seen for the SMFMO control but was absent in the soluble fraction of LIC expression.

Large scale Phe52Asn and Phe52Val expression and purification

The SMFMO gene, containing the desired mutation, in the pET-YSBLIC-3C vector was expressed in BL21 (DE3) cells as described in Section 3.2.5 with a total volume of 2 L LB media. After expression the protein was purified.
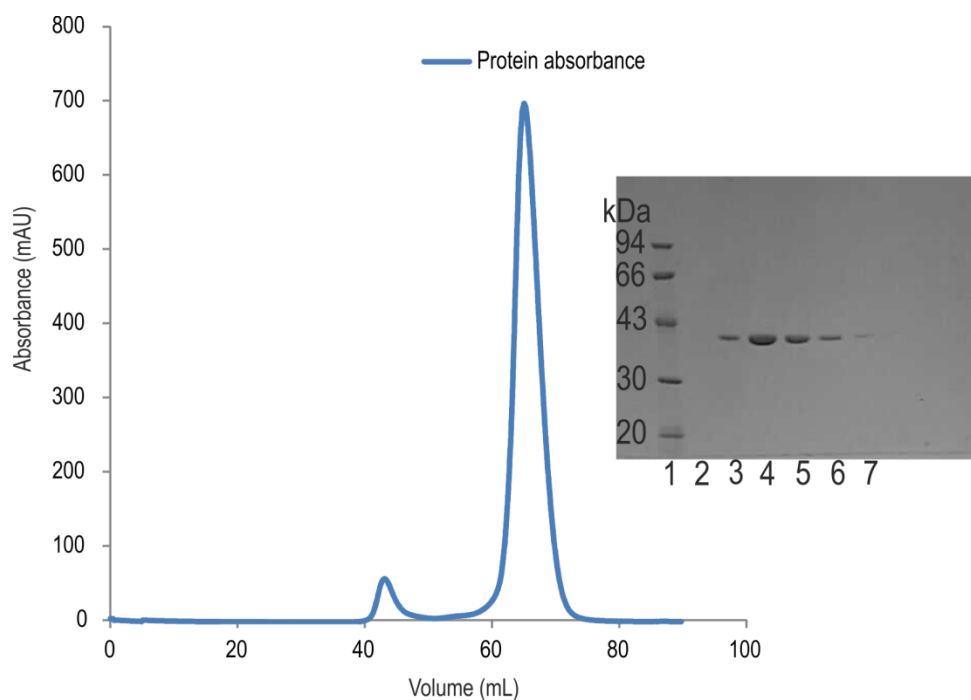
Figure 7.5 Chromatogram and SDS-PAGE gel of SMFMO mutant Phe52Val purification by size exclusion chromatography.
Protein absorbance at 280 nm is indicated by the blue trace. Lane 1: Bio-Rad low weight molecular marker, lane 2-7: collected fractions after FPLC run. The protein can be seen at 38 kDa.

Both mutants Phe52Asn and Phe52Val were purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography (Figure 7.5) which was successful in yielding pure yellow protein with a molecular weight of ~38 kDa, as shown on the SDS-PAGE gel for mutant Phe52Val. The mutants eluted between 55-70 mL in accordance with the wild type indicating the mutants are dimers. The expression of the mutants proved similar to that of the wild type SMFMO yielding ~3 mg of protein per litre of medium.

The mutation Phe52Ala in SMFMO resulted in a variant which could not be expressed within the soluble fraction of the *E. coli* cells (Figure 7.6).
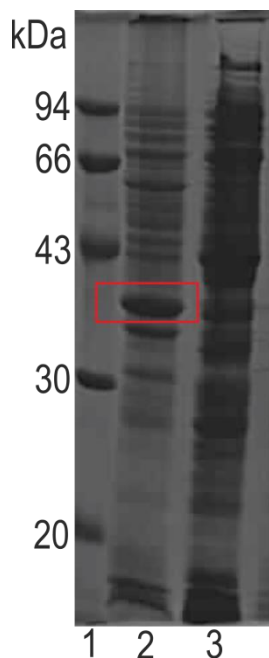
Figure 7.6 SDS-PAGE gel of expression for SMFMO mutant Phe52Ala.
Lane 1: Bio-Rad low weight molecular marker, lane 2: insoluble fraction of Phe52Ala expression, lane 3: soluble fraction of Phe52Ala expression. The protein can be seen at ~38 kDa in the insoluble fraction.

A band at ~38 kDa can be seen in the insoluble fraction on the SDS-PAGE gel but the band is absent in the soluble fraction. In an effort to salvage any soluble expression Ni^{2+} affinity chromatography purification was attempted. Unfortunately, no protein peak was present on the chromatogram.

7.4.3 Biotransformations with SMFMO variants Phe52X, Asn173X and Ser174X

It was found that SMFMO had the ability to convert substrates **13** and **15** with respective conversions and *ee*s of 27 % (71 % -(*R*)) and 90% (25 % -(*R*)) when NADH was employed as cofactor. The expressed mutants Phe52X, Asn173X and Ser174X were sonicated and the lysate was used in the biotransformations of substrates **13** and **15** in order to investigate the mutants' ability to oxidise the substrates and also their

enantioselectivity towards them, when NADH was used as cofactor. The results are shown in Table 7.2.

Table 7.2 Conversions and enantioselectivity of SMFMO mutants Phe52X, Asn173X and Ser174X towards substrates **13** and **15**.

SMFMO Variant	Conversion of 13 (%)	*ee* for sulfoxide product of 13 (%)	Conversion of 15 (%)	*ee* for sulfoxide product of 15 (%)
SMFMO	27	71 –(*R*)	90	25 –(*R*)
LIC	0	-	0	-
N173G	0	-	4	12 –(*S*)
N173S	0	-	3	8 –(*R*)
N173R	0	-	1	3 –(*R*)
N173V	0	-	2	-
N173L	0	-	2	-
N173I	0	-	2	4 –(*R*)
N173H	0	-	1	6 –(*R*)
N173D	0	-	1	3 –(*R*)
S174R	0	-	4	-
S174G	0	-	0	-
S174I	0	-	1	11 –(*R*)
S174N	0	-	0.5	16 –(*R*)
S174C	3	48 –(*R*)	14	37 –(*R*)
S174D	0	-	2	12 –(*R*)
S174F	0	-	1	1 –(*R*)
S174L	0	-	0	-
S174V	0	-	1	12 –(*R*)
F52V	28	racemic	17	32 –(*S*)
F52R	1	46 –(*S*)	4	38 –(*R*)
F52L	24	67 –(*R*)	14	24 –(*R*)

The mutations of Asn173 and Ser174 resulted largely in variants of low activity. The conversions of substrate **13** to its sulfoxide were zero in most cases. When **15** was

utilised as substrate the conversions were much lower than those observed for the wild type when NADH is used as cofactor. However, for the Ser174Cys mutation higher conversions could be seen and the *ee* for the sulfoxide for substrate **15** was similar to that of the wild type (37 % versus 25 %), which can be expected as the mutation was conservative, both similar polar amino acids. The Phe52 mutants were found to be more active, with the mutation to hydrophobic leucine producing a mutant with increased conversions which was also reasonably enantioselective with both substrates (67 % and 24%), with the same configuration –(*R*) as that of the wild type. The mutant Phe52Asn was found to have similar conversions for substrate **15** that of wild type SMFMO (86 %) but diminished the enantioselectivity for both NADH and NADPH causing the mutated enzyme to give a racemic product (not shown in Table 7.2). Interestingly, the Phe52 mutant with the smaller hydrophobic group, valine, gave an oxidised product for substrate **15** with inverted absolute configuration (-(*S*)) compared to the wild type enzyme (-(*R*)). This observation insinuated that when less sterically hindered amino acids are utilized the active site maybe be more available in some way. In order to demonstrate this effect, Phe52 was mutated to the much smaller alanine, but unfortunately the mutant was not expressed well, and the cell extracts that contained the mutant did not appear yellow, suggesting that a side chain of minimum size is needed in this position to maintain flavin binding, in addition to the flavin hydrogen bonding to the main chain NH of residue Phe52.

Further investigation into the phenyalanine52 residue

Using the Phe52Val mutant the inversion of stereochemistry when compared to the wild type was investigated. Isolated Phe52Val mutant protein was used in the biotransformations of the sulfide library (substrates **13**-**19**, Figure 4.10) used in the characterisation of SMFMO, with NADH as cofactor. The results can be seen in Table 7.3.

Table 7.3 Biotransformations of sulfides by SMFMO mutant Phe52Val compared to biotransformations of equivalent substrates using wild type SMFMO.

Substrate	Conversion (%) WT SMFMO	ee (%) WT SMFMO	Conversion (%) Phe52Val	ee (%) Phe52Val
13	27	71 –(*R*)	79	11 –(*S*)
14	32	24 –(*R*)	91	7 –(*S*)
15	90	25 –(*R*)	77	32 –(*S*)
16	8	21 –(*R*)	69	8 –(*S*)
17	6	15 –(*S*)	69	19 –(*S*)
18	18	30 –(*R*)	20	41 –(*S*)
19	40	80 –(*R*)	24	28 –(*R*)

With the Phe52Val mutant it was found that the conversions of the sulfides into their corresponding sulfoxides were increased save for substrate **15**. Whilst the enantioselectivity in most cases was poor the absolute configurations of products for the mutant ((*S*)-enantiomer) was opposite to those obtained for the wild-type ((*R*)-enantiomer). However, for substrate **19** the (*R*)-selectivity was maintained although significantly reduced compared to that obtained for the wild type (28 % *ee* compared to 80 % *ee*). In addition, for substrate **17** the (*S*)-selectivity was maintained with similar ee to that of the wild type.

7.5 Discussion

It was found that mutations at residues Asn173 and Ser174 produced variants with little activity when tested for their ability to oxidise substrates **13** and **15**. Neither of these residues interact with the flavin however superimposition of mFMO complexed with NADPH (2XLT) indicates that residues Asn173 and Ser174 may be in proximity to the nicotinamide ring of NADPH. Interestingly, position 173 is occupied by Tyr212 in

mFMO (Figure 7.7) which is thought to act as a 'backdoor' residue in order to protect the active site from bulk solvent.[103] The mutation of Asn173 to a tyrosine residue however could not be obtained.
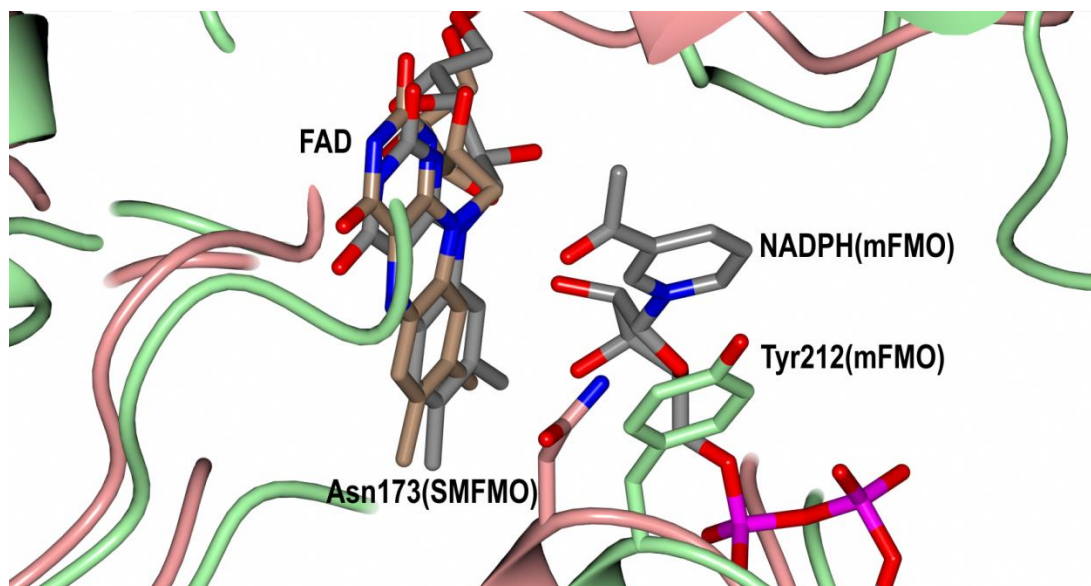
Figure 7.7 Superimposition of SMFMO with the flavin binding site of mFMO (2XLT) illustrating that the 'backdoor' residue Tyr212 of mFMO is occupied by Asn173 in SMFMO.
Protein backbone is shown in ribbon format with SMFMO in pink and mFMO in green. The FAD for SMFMO is shown in cylinder format with carbon atoms in gold. The FAD and NADPH for mFMO are shown in cylinder format with carbon atoms in grey.

The mutations at residue Phe52 proved more successful, more specifically the mutation to valine in which the absolute configuration of sulfoxide products was inverted in most cases compared to those obtained with the wild type. The role of Phe52 in SMFMO is somewhat different when compared to structurally homologous residues in other FPMOs. In mFMO an asparagine residue is present[73] and mutation of this residue, even to a conservative serine, produced an inactive enzyme. It has been proposed that Asn73 in mFMO is involved in the stabilisation of the flavin hydroperoxide, as mutation of this residue knocks out the enzymes ability to oxygenate substrates.[103] However in other cases, mutation of the residue at this position has had little or no effect on the activity of the enzyme.[108] In SMFMO, Phe52 was found to have an effect on substrate binding, as the mutation to a smaller residue, Phe52Val, resulted in a variant with altered enantioselectivity. In other FPMOs the selectivity in regards to sulfoxidation has also

been altered by mutagenesis. For example, the structure guided mutagenesis of PAMO allowed for the Met446 active site residue mutant, Met446Gly, which catalysed the oxidation of methyl-p-tolyl sulfide to its (*R*)-enantiomer with 92 % ee compared to 6% for the wild type.[111] A non-structure based investigation of CHMO from *Acinetobacter calcoaceticus* was performed by Reetz *et al.* in which directed evolution was used to generate mutants which were enantiocomplementary when methyl-p-methyl benzyl thioether was employed as substrate. For example, the CHMO variant Phe432Ser allowed for the *(R)*-sulfoxide enantiomer with 98.7 % *ee* and the double mutant Phe16Leu/Phe227Ser variant gave the (S)-product with 95.2 % *ee*.[58b] It can be seen that a number of different residues, in the proximity of the active site or remote, in FPMOs have been valuable in regards to mutations to alter or improve enantioselectivity.

7.6 Conclusion

The mutagenesis study of SMFMO found that residue Phe52 had no absolute requirement but however had an interesting effect on the active sites recognition of substrates when mutated to a much smaller amino acid, thus, resulting in inverted enantioselectivity. The mutation of residues Asn173 and Ser174 resulted in enzymes with low activity indicating that both residues are important within the active site. The role residues Phe52, Asn173 and Ser174 play in SMFMO and structurally homologous residues in other FPMOs is still under investigation.

Chapter 8: Investigation into putative NADH specific enzymes

8.1 Introduction

SMFMO is interesting amongst FPMOs as it is one of the only enzymes, in addition to the group of enzymes from *Rhodococcus jostii* RHA1[18] and the BVMO mekA from *Pseudomonas veronii* MEK700[112], that can utilise the cheaper nicotinamide cofactor NADH to oxidise prochiral sulfides, albeit with poor enantioselectivity. The residues Gln193 and His194 were found to be the determinants of cofactor promiscuity in SMFMO, in particular residue His194. In an effort to find novel FPMOs with the potential to be NADH specific a BLAST search was conducted, looking for naturally occurring mutants of SMFMO in positions Gln193 and His194. From the search two enzymes were selected, a monooxygenase from *Pseudomonas stutzeri* NF13 (PFMO) and pyridine nucleotide-disulfide oxidoreductase from *Cellvibrio sp*. BR (CFMO) (Figure 8.1).

```
SMFMO    ----------MDSVDVVVIGGGQSGLSAGYFLRRSGLSYVILDAEASPGGAWQHAWHSLH    50
PFMO     ---------MPPILDVIVIGGGQAALTTAYFLRRTSLSYLLLDEQPGPGGAWLHAWDSLR    51
CFMO     MDTPVMDSTDTDSTDIVIIGGGQAALSVAYYLRRSKYSFVMLDAEQTPGGAWLHGWDSLR    60

SMFMO    LFSPAGWSSIPGWPMPASQG-PYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERL    109
PFMO     LFSPAAWSSIAGWPMPSPTEPGNPTRNDVIDYLRRYEDRYQFPIQRPVRVDTVTRLDDLW    111
CFMO     LFSPSTWSSLSGWQMPPT-GETYPSRDQVVDYLRHYESRYEFPVQRPVWVSAVNNLGDRL    119

SMFMO    RVVARDGRQWLARAVISATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAII    169
PFMO     RVQAG-DQQWLARAVISATGTWSKPFIPPYEGRELFQGAQIHSAHYRTPAPFAGKRVMVV    170
CFMO     EVVSE-RQQWRARVVISATGTWRNPFIPAYPGADLFQGAQLHSAHYQSPAPFAGQKVLVV    178

SMFMO    GGGNSGAQILAEVSTVAETTWITQHEPAFLADDVDGRVLFERATERWKAQQEGREPDLPP    229
PFMO     GGGNSGAQVLAELSSVSETLWITQEPPAFLPDEVDGRVLFERATARWKAQQEGRSIDEPA    230
CFMO     GGGNSGAQILAEVSRVADCTWVTTSEPIFLPDDVDGRVLFQRATDRWKAAQEGREIEQPV    238

SMFMO    GGFGDIVMVPPVLDARARGVLAAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSH    289
PFMO     GGFGDIVMVPPVREARERGVLVAERPFARFTETGVEWADGRRENLDAVIWCSGFRPALDH    290
CFMO     GGLGDVVMVPPVKEARERGALHAVRPFTRFTANGVVWADGTGSAVDAVIWCTGFRPALAH    298

SMFMO    LKGLDLVTPQGQVEVDGSGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTA    349
PFMO     LRELGVVEADGKVQVED--TRVVKQPNLWLVGYGDWTGMASATLIGVTRTARSTADQVVQ    348
CFMO     LQSLGVINPDGKVDLAG--TRSLQEPRLWLLGYGEWTGLASATLIGVGRSARATAEEIIQ    356

SMFMO    YCADHQDR--      357
PFMO     ALTATPSRRP      358
CFMO     YLDSA-----      361
```

Figure 8.1 Sequence alignment of SMFMO, PFMO and CFMO highlighting the residues in line with residues Gln193 and His194 of SMFMO. PFMO= monooxygenase from *Pseudomonas stutzeri* NF13; CFMO= pyridine nucleotide-disulfide oxidoreductase from *Cellvibrio sp*. BR. Rossman motifs are shown in red and FMO motif shown in red.

As with SMFMO, proteins PFMO and CFMO contain two Rossman motifs for the binding of the ADP moiety in FAD and NAD(P)H, and an amino acid motif typical of a FMO.[11] Both PFMO and CFMO are of similar size to SMFMO, with PFMO being 358 aa (39.6 kDa) in length and CFMO 361 aa (39.4 kDa) compared to the 357 aa (38.6 kDa) of SMFMO. A sequence alignment obtained using ClustalW2[113] revealed a 60.78 % sequence identity between PFMO and SMFMO and a 58.26 % sequence identity between CFMO and SMFMO.

PFMO and CFMO proved interesting as they did not have the usual arginine/threonine couple present in other FPMOs which is responsible for phosphate binding.[59, 61a, 103] PFMO was selected as it was found to possess a glutamine (Gln194) and glutamate (Glu195) at the position where SMFMO has the cofactor determinant residues Gln193 and His194. The PFMO residues sparked interest due to the glutamate residue as the Gln193Glu mutant could not be obtained for SMFMO. PFMO is a variant of SMFMO that could be a strict NADH-dependent enzyme because of the glutamate residues ability to repel negatively charged phosphate. CFMO was found to be also a variant of SMFMO at positions 193 and 194, with CFMO possessing polar residues threonine

(Thr202) and serine (Ser203). The selected targets are different compared to the NADP ribose 2'-phosphate recognition site in SMFMO (when superimposed with mFMO, residues Gln193 and His194) and the natural mutants could prove to be novel NADH-dependent FMOs.

8.2 Aims

The aim of this chapter is to clone, express and characterise the selected targets PFMO and CFMO in order to identify if the variants obtain NADH-specific activity.

8.3 Materials and methods

8.3.1 Obtaining of PFMO and CFMO genes

The genes coding a monooxygenase from *Pseudomonas stutzeri* NF13 (PFMO) and a pyridine nucleotide-disulfide oxidoreductase from *Cellvibrio sp*. BR (CFMO) were codon optimised for *E. coli* and purchased from Geneart in vector pMA-T. The gene sequences for the targets are shown in Figure 8.2.

```
PFMO 1077 BP;
ATGCCTCCGA TTCTGGATGT TATTGTTATT GGTGGTGGTC AGGCAGCACT GACAACCGCA        60
TATTTTCTGC GTCGTACCAG CCTGAGCTAT CTGCTGCTGG ATGAACAGCC TGGTCCGGGT       120
GGTGCATGGC TGCATGCATG GGATAGCCTG CGTCTGTTTA GTCCGGCAGC ATGGTCAAGC       180
ATTGCAGGTT GGCCGATGCC GAGCCCGACC GAACCGGGTA ATCCGACCCG TAATGATGTG       240
ATTGATTATC TGCGTCGCTA TGAAGATCGT TATCAGTTTC CGATTCAGCG TCCGGTTCGT       300
GTTGATACCG TTACCCGTCT GGATGATCTG TGGCGTGTTC AGGCAGGCGA TCAGCAGTGG       360
CTGGCACGTG CAGTTATTAG CGCAACCGGC ACCTGGTCAA AACCGTTTAT TCCGCCTTAT       420
GAAGGTCGTG AACTGTTTCA GGGTGCACAG ATTCATAGCG CACATTATCG TACACCGGCA       480
CCGTTTGCAG GTAAACGTGT TATGGTTGTG GGTGGTGGCA ATAGCGGTGC ACAGGTTCTG       540
GCAGAACTGA GCAGCGTTAG CGAAACCCTG TGGATTACCC AAGAACCGCC TGCATTTCTG       600
CCGGATGAAG TTGATGGTCG TGTTCTGTTT GAACGTGCAA CCGCACGTTG GAAAGCACAG       660
CAAGAGGGTC GTAGCATTGA TGAACCTGCC GGTGGTTTTG GTGATATTGT GATGGTTCCG       720
CCTGTTCGTG AAGCACGTGA ACGTGGTGTT CTGGTTGCAG AACGTCCGTT TGCCCGTTTT       780
ACCGAAACCG GTGTTGAATG GGCAGATGGT CGTCGTGAAA ATCTGGATGC CGTTATTTGG       840
TGTAGCGGTT TTCGTCCGGC ACTGGATCAT CTGCGTGAAC TGGGTGTTGT TGAAGCAGAT       900
GGTAAAGTTC AGGTTGAAGA TACCCGTGTT GTTAAACAGC CGAATCTGTG GCTGGTTGGT       960
TATGGTGATT GGACCGGTAT GGCAAGCGCA ACCCTGATTG GTGTTACCCG TACCGCACGT      1020
AGCACCGCAG ATCAGGTTGT TCAGGCCCTG ACCGCAACCC CGAGCCGTCG TCCGTAA        1077
```

```
CFMO 1086 BP;
ATGGATACAC CGGTTATGGA TAGCACCGAT ACCGATAGTA CCGATATTGT TATTATTGGT        60
GGTGGTCAGG CAGCACTGAG CGTTGCATAT TATCTGCGTC GTAGCAAATA TAGCTTTGTT       120
ATGCTGGATG CAGAACAGAC ACCGGGTGGT GCATGGCTGC ATGGTTGGGA TAGCCTGCGT       180
CTGTTTAGCC CGAGCACCTG GTCAAGCCTG AGCGGTTGGC AGATGCCTCC GACCGGTGAA       240
ACCTATCCGA GCCGTGATCA GGTTGTTGAT TACCTGCGTC ATTATGAAAG CCGTTATGAA       300
TTTCCGGTTC AGCGTCCGGT TTGGGTTAGC GCAGTTAATA ATCTGGGTGA TCGTCTGGAA       360
GTTGTTAGCG AACGTCAGCA GTGGCGTGCA CGTGTTGTTA TTAGCGCAAC CGGCACCTGG       420
CGTAATCCGT TTATTCCGGC ATATCCGGGT GCAGACCTGT TTCAGGGTGC ACAGCTGCAT       480
AGCGCACATT ATCAGAGTCC GGCACCGTTT GCAGGTCAGA AAGTTCTGGT TGTGGGTGGT       540
GGTAATAGCG GTGCACAGAT TCTGGCCGAA GTTAGCCGTG TTGCAGATTG TACCTGGGTT       600
ACCACCAGTG AACCGATTTT TCTGCCGGAT GATGTTGATG GTCGTGTTCT GTTCCAGCGT       660
GCAACCGATC GTTGGAAAGC AGCACAAGAA GGTCGCGAAA TTGAACAGCC GGTTGGTGGC       720
CTGGGTGATG TTGTTATGGT TCCGCCTGTT AAAGAAGCAC GCGAACGCGG AGCACTGCAT       780
GCCGTTCGTC CGTTTACCCG TTTTACCGCA AATGGTGTTG TTTGGGCAGA TGGCACCGGT       840
AGCGCAGTGG ATGCAGTTAT TTGGTGTACC GGTTTTCGTC CGGCACTGGC ACATCTGCAG       900
AGCCTGGGTG TTATTAATCC GGATGGTAAA GTTGATCTGG CAGGCACCCG TAGCCTGCAA       960
GAACCGCGTC TGTGGCTGCT GGGTTATGGT GAATGGACCG GTCTGGCAAG CGCAACCCTG      1020
ATTGGTGTTG TCGTAGCGC ACGTGCCACC GCAGAAGAAA TTATTCAGTA TCTGGATAGC      1080
GCCTAA                                                              1086
```

Figure 8.2 The codon optimised gene sequence for PFMO and CFMO.

8.3.2 Ligation independent cloning

Both the PFMO and CFMO genes were cloned into the pET-YSBLIC-3C vector using the LIC protocol described in Section 3.1.2. Each gene was amplified using the protocol in Section 3.1.1. The primers used for each gene in the PCR are shown in Table 8.1.

Table 8.1 Primers used in PCR for PFMO and CFMO gene amplification

Gene	Forward primer	Reverse primer
PFMO	CCAGGGACCAGCAATGCCTC CGATTCTGG	GAGGAGAAGGCGCGTTACGGAC GACGGCTCGG
CFMO	CCAGGGACCAGCAATGGATA CACCGGTTATGG	GAGGAGAAGGCGCGTTAGGCGCT ATCCAGATACTG

Pure T4 treated linearised pET-YSBLIC-3C vector was sourced from Laila Roper within the Grogan group. The insert was treated with T4 polymerase as detailed in Section 3.1.4 and the vector and insert were annealed together as described in Section 3.1.5.

8.3.3 Expression and purification

The target genes in the pET-YSBLIC-3C vector were expressed in *E. coli* BL21 (DE3) cells.

Expression tests

Small scale expression was carried out as described in section 3.2.3 and the soluble and insoluble fractions were analysed by SDS-PAGE (Section 3.2.4).

Large scale expression and purification

The target genes were expressed as described in Section 3.2.5 and the proteins were purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography described in Section 3.3.

8.3.4 Enzyme assays

Kinetic assays with PFMO and CFMO

Kinetic studies were carried out in accordance to Section 3.4.2. Kinetic parameters were calculated using the GraFit data analysis software.

Biotransformations with PFMO and CFMO

Biotransformations were carried out as described in Section 3.4.4 using a range of prochiral sulfides (Figure 4.8). Biotransformations were analysed by GC and GC-MS described in Sections 3.4.5 and 3.4.6.

8.3.5 His$_6$-tag cleavage of PFMO

The His$_6$-tag present in PFMO was cleaved using 3C protease as described in Section 3.3.4.

8.3.6 Crystallisation studies

Pure PFMO and CFMO was subjected to crystallisation trials using a range of commercially available screens in 96-well plates employing 300 nL drops at a range of protein concentrations (3, 10 and 20 mg mL^{-1}). The best crystals for PFMO were obtained using the Clear Strategy Screen conditions containing 35 % w/v tacsimate pH 7.0 and His$_6$-tag cleaved protein (20 mg mL^{-1}). The best crystals for CFMO were obtained using the Clear Strategy Screen conditions containing 1.5 M ammonium sulfate and protein (20 mg mL^{-1}).[101] Larger crystals for diffraction analysis using optimised conditions were prepared using the hanging-drop vapour diffusion method in 24-well plate Linbro dishes and using crystallisation drops of 2 mL with protein (20 mg mL^{-1}). For PFMO, the best crystals were obtained in crystal drops containing 35 % w/v tacsimate at pH 7.0. For CFMO, the best crystals were obtained in crystal drops containing 1.5 M ammonium sulfate and propan-2-ol (10 μL). Crystals were tested for diffraction using a Rigaku Micromax-007HF generator fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. Those crystals that diffracted to greater than 3 Å resolutions were flash-cooled in liquid nitrogen in a cryogenic solution containing the mother liquor containing also 10 % w/v glycerol and retained for data collection at the synchrotron. Data processing and structure solution

for PFMO and CFMO was conducted by Dr Gideon Grogan. Collection and refinement statistics are shown in Table 8.2.

Table 8.2 Data collection and refinement statistics for PFMO and CFMO

	CFMO	PFMO
Date	12-03-14	12-03-14
Beamline	Diamond i24	Diamond i24
Wavelength (Å)	0.96862	0.96862
Resolution (Å)	22.44 – 2.39 (2.45-2.39)	35.94-1.83 (1.88-1.83)
Space Group	$C2$	$P3_22_1$
Unit Cell	a = 115.41; b = 95.09; c = 92.37 $\alpha = \beta = 90.0$; $\gamma = 126.3$	a = b = 63.56; c = 189.82 $\alpha = \beta = 90.0$; $\gamma = 120.0$
No. of molecules in the asymmetric unit	2	1
Unique reflections	31344 (2329)	41893 (3176)
Completeness (%)	98.1 (98.4)	100 (99.9)
R_{merge} **(%)**	0.09 (0.50)	0.112 (0.62)
$R_{p.i.m.}$	0.09 (0.50)	0.054 (0.30)
Multiplicity	3.2 (2.9)	9.8 (9.9)
$<I/\sigma(I)>$	9.0 (2.1)	14.9 (3.8)
$CC_{1/2}$	0.99 (0.76)	100 (0.90)
Overall B **factor from Wilson plot (Å^2)**	25	13
R_{cryst}/R_{free} **(%)**	24.8/28.7	16.5 (20.5)
r.m.s.d. 1-2 bonds (Å)	0.014	0.02
r.m.s.d. 1-3 bonds (°)	1.659	2.26
Avge main chain B (Å^2)	38	19
Avge side chain B (Å^2)	40	22
Avge water B (Å^2)	31	24

8.4 Results

8.4.1 Ligation independent cloning

Amplification of target genes

LIC primers were obtained for both PFMO and CFMO and the genes were amplified by PCR. The PCR results were observed by agarose gel electrophoresis (Figure 8.3)
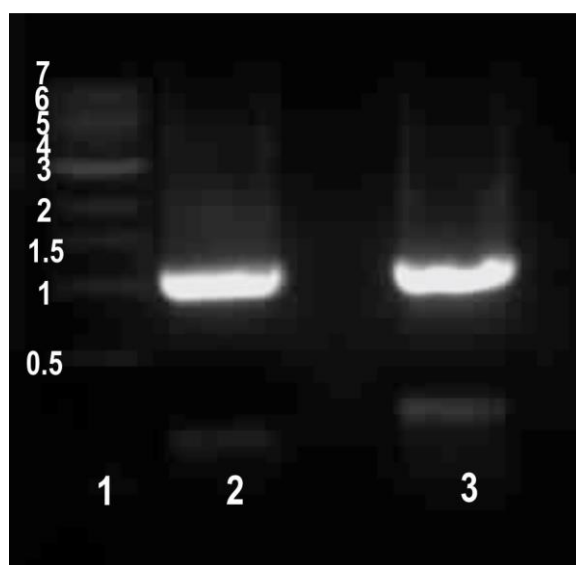
Figure 8.3 Agarose gel of gene amplification by PCR for PFMO and CFMO genes. Lane 1: NEB 1 kb ladder, lane 2: PCR product for CFMO gene, lane 3: PCR product for PFMO gene.

A bright band at 1 kb can be seen for the PFMO gene and the CFMO gene. The PCR products were cut out and purified by gel extraction yielding 5 μg of PCR product for PFMO and 5.5 μg for CFMO.

T4 polymerase reaction

Pure T4 treated linearised pET-YSBLIC-3C vector at a concentration of 5.4 μg was sourced from Laila Roper within the Grogan group. The gene inserts were treated with T4 polymerase as described in Section 3.1.4 and 8.3.2.

LIC annealing, transformation and mini-prep

The treated vector and insert for PFMO and CFMO were annealed together by incubation at room temperature as described in Section 3.1.5. The resulting reaction was transformed into *E. coli* XL10 Gold Ultra-competent cells and the next day 4 colonies were picked for each target gene and mini-prepped to recover the DNA.

Restriction digest

A sample for each of the DNA samples for PFMO and CFMO were digested with the restriction endonucleases NcoI and NdeI (Section 3.1.7) to confirm gene insertion. The results are shown in Figure 8.4.
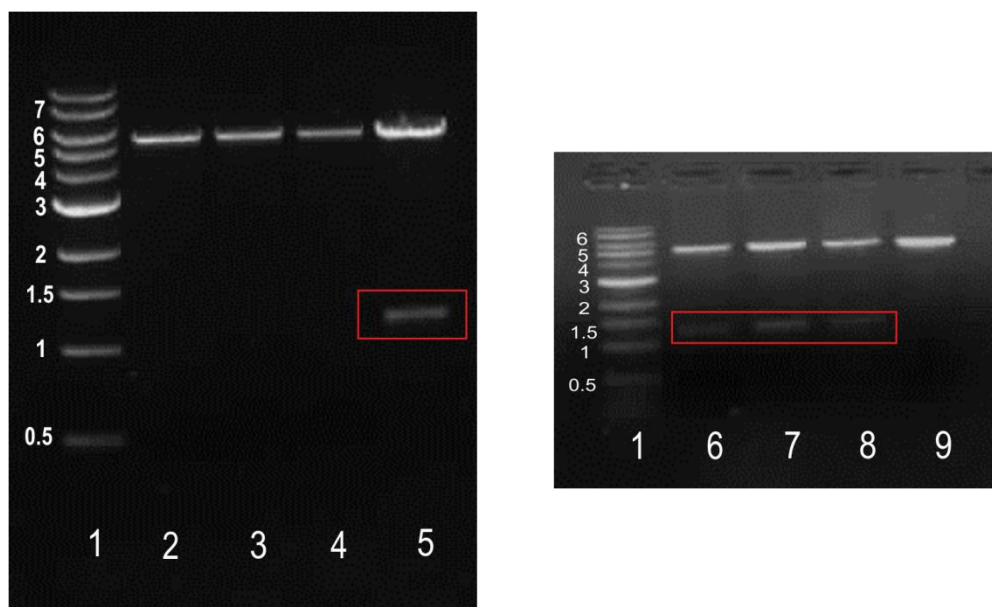
Figure 8.4 Agarose gel for PFMO and CFMO restriction digest experiments. Lane 1: NEB 1 kb ladder, lanes 2-5: restriction digest for the four colonies picked for CFMO gene, lane 6-9: restriction digest for the four colonies picked for PFMO gene. The insert for each successful reaction is boxed in red.

The restriction digest experiment for PFMO shows that in three of the cases the gene insert was present. The insert can be seen just above the 1 kb (boxed in red) with the remaining vector at ~6 kb. For CFMO one of the experiments had successfully inserted the gene insert with the band at ~1 kb (highlighted in red) and the vector at ~6 kb.

DNA sequencing

For PFMO reactions corresponding to lanes 6 and 7 in Figure 8.4 were submitted for sequencing. For CFMO the reaction corresponding to lane 5 was also sequenced. DNA sequencing confirmed that the gene sequences for PFMO (lane 6) and CFMO had 100 % sequence identity with their gene sequences. The PFMO reaction yielded a DNA concentration of 4.3 µg and the CFMO gave a DNA concentration of 4.1 µg.

8.4.2 Expression and purification of PFMO and CFMO

The target genes in the pET-YSBLIC-3C vector were expressed in *E. coli* BL21 (DE3) cells.

Expression testing

Expression testing at different temperatures after IPTG induction was carried out to determine the favourable conditions for gene expression of soluble recombinant protein. The expression cultures were incubated at $37^{o}C$ until OD_{600nm} reached approx. 0.6 and then induced with IPTG. The cultures were then incubated overnight at 16, 30 and $37^{o}C$ as described in Section 3.2.3.

The soluble and insoluble fractions were analysed using SDS-PAGE, Figure 8.5.
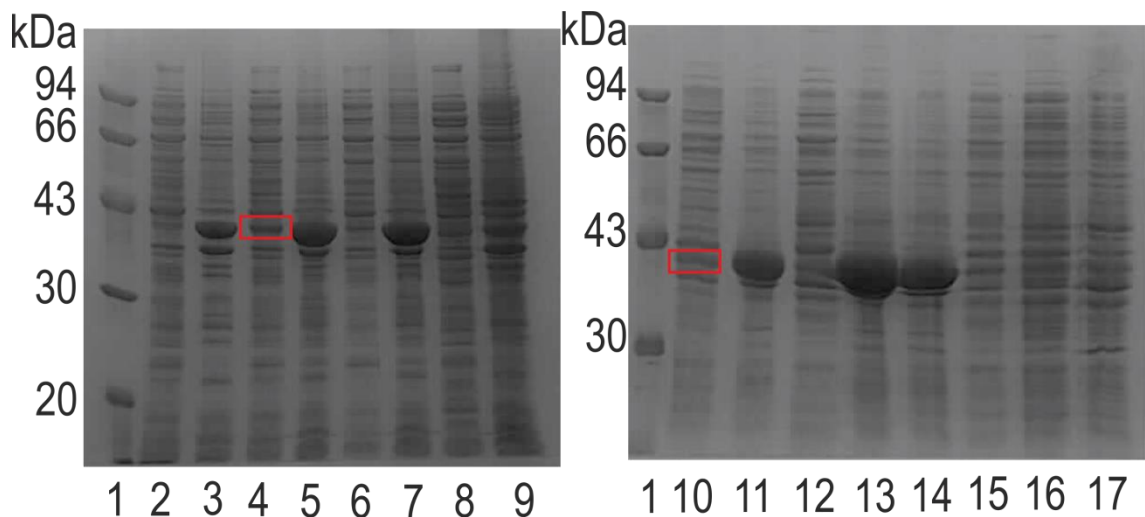
Figure 8.5 SDS-PAGE gels for expression testing for PFMO and CFMO.
Lane 1: Bio-Rad low weight molecular marker, lane 2: soluble PFMO fraction for 16°C incubation, lane 3: insoluble PFMO fraction for 16°C incubation, lane 4: soluble PFMO fraction for 30°C incubation, lane 5: insoluble PFMO fraction for 30°C incubation, lane 6: soluble PFMO fraction for 37°C incubation, lane 7: insoluble PFMO fraction for 37°C incubation, lane 8: soluble PFMO fraction for 16°C incubation without IPTG induction, lane 9: insoluble PFMO fraction for 16°C incubation without IPTG induction, lane 10: soluble CFMO fraction for 16°C incubation, lane 11: insoluble CFMO fraction for 16°C incubation, lane 12: soluble CFMO fraction for 30°C incubation, lane 13: insoluble CFMO fraction for 30°C incubation, lane 14: insoluble CFMO fraction for 37°C incubation, lane 15: soluble CFMO fraction for 37°C incubation, lane 16: soluble CFMO fraction for 16°C incubation without IPTG induction, lane 17: insoluble CFMO fraction for 16°C incubation without IPTG induction. Soluble expression for both PFMO and CFMO is outlined in red.

The expression testing revealed the favoured expression conditions for PFMO were expression at 37°C followed by incubation at 30°C after IPTG induction. This system gave the greatest soluble expression of PFMO as indicted by a band at ~39 kDa (boxed in red). For CFMO the favoured expression was similar to that of SMFMO, expression at 37°C followed by incubation at 16°C after IPTG induction, giving the greatest soluble expression of CFMO as indicted by a band at ~39 kDa (boxed in red). For both proteins the majority of the expression however was insoluble. Controls set up in the absence of IPTG induction confirmed over-expression did not occur without induction.

Large scale expression and purification

PFMO

The gene was expressed in a total volume of 2 L LB media using the expression conditions obtained from expression testing. The protein was purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography (Figure 8.6).
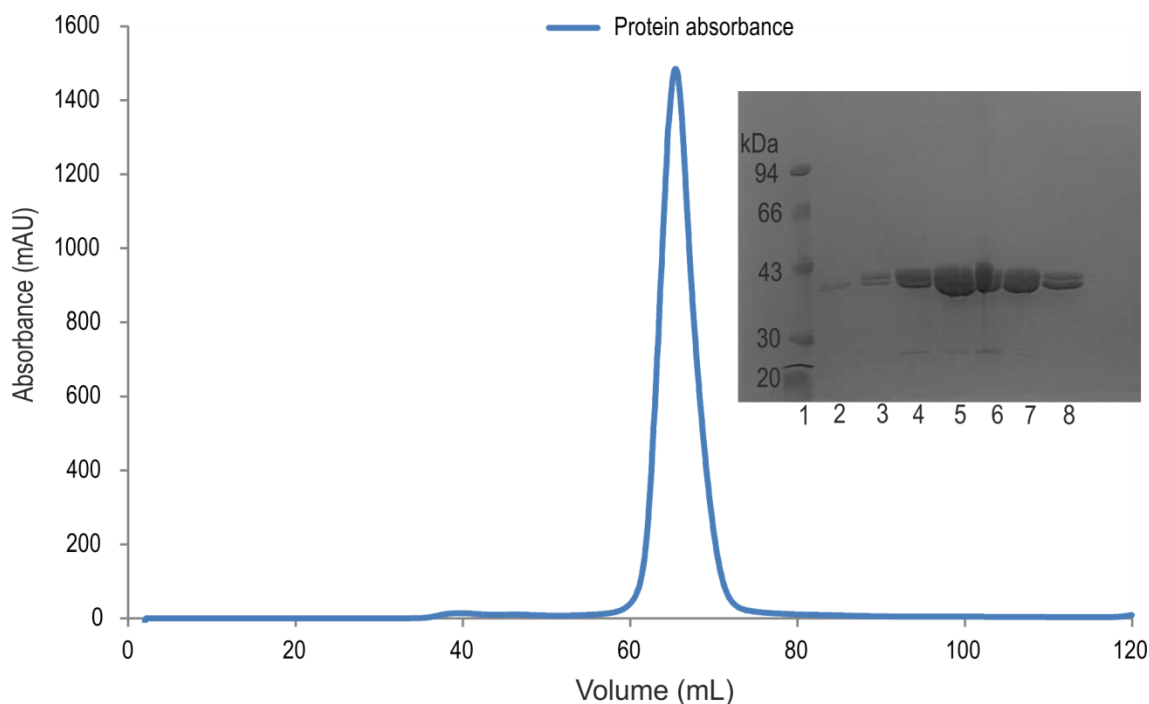
Figure 8.6 Chromatogram and SDS-PAGE gel of PFMO purification by size exclusion chromatography.
Protein absorbance at 280 nm is indicated by the blue trace. Lane 1: Bio-Rad low weight molecular marker, lane 2-8: collected fractions after FPLC run. The protein can be seen at ~39 kDa.

The protein eluted as a single peak (between 60-75 mL). Using a calibration curve provided by GE Healthcare, the position of the elution peak gave the molecular weight at ~40 kDa, indicating that PFMO is a monomer in solution. The fractions eluted were bright yellow in colour, indicating FAD was present within the purified protein and the fractions analysed by SDS-PAGE confirmed a protein band at ~39 kDa. The fractions were combined to give ~15 mg of pure protein.

CFMO

The CFMO gene was expressed in a total volume of 2 L LB media using the expression conditions obtained from expression testing. The protein was purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography (Figure 8.7).
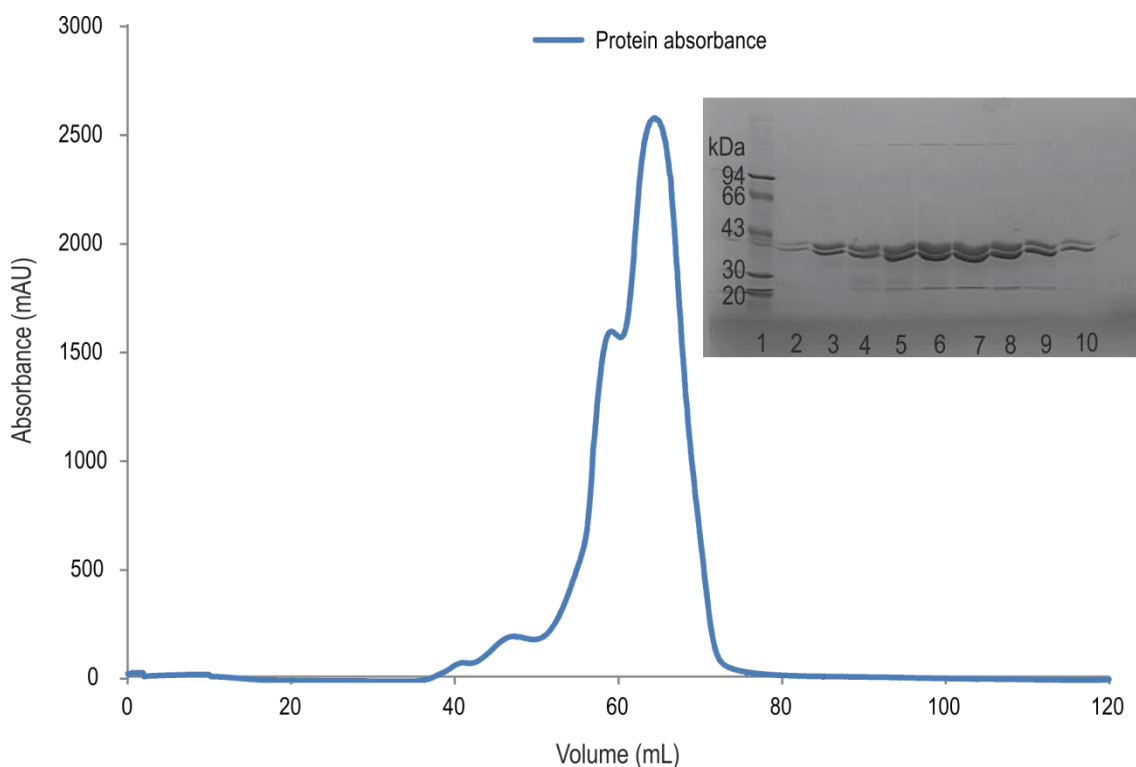
Figure 8.7 Chromatogram and SDS-PAGE gel of CFMO purification by size exclusion chromatography.
Protein absorbance at 280 nm is indicated by the blue trace. Lane 1: Bio-Rad low weight molecular marker, lane 2-4: collected fractions between 40-55 mL, lane 5-10: collected fractions between 60-75 mL. The protein can be seen at ~39 kDa.

CFMO eluted as a large peak (from 40-70 mL) with shouldering. Fractions collected between 40-60 mL were colourless; however SDS-PAGE gel indicated a protein band at ~39 kDa. The fractions collected between 60 -75 mL were bright yellow indicating FAD was present within the purified protein and the fractions analysed by SDS-PAGE confirmed a protein band at ~39 kDa. The shouldering effect indicates the protein exists in two different oligomeric states. The elution profiles of standard proteins run on the Superdex® 75 gel filtration column assigned the shouldering peak between 40-55 mL as a trimer and the main peak between 60-75 mL as a monomer. The yellow fractions were pooled to give ~40 mg of pure protein.

8.4.3 Kinetic assays

SMFMO has the ability to use either NADH or NADPH as nicotinamide cofactor with comparable capability.

To investigate the ability of the target enzymes, PFMO and CFMO, to use either NADH or NADPH to reduce FAD, UV- spectrophotometry assays were carried out as described in Section 3.4.2, using increasing concentrations of either cofactor.

Similar to SMFMO, PFMO could utilise both cofactors to reduce the bound flavin, illustrated in Figure 8.8.
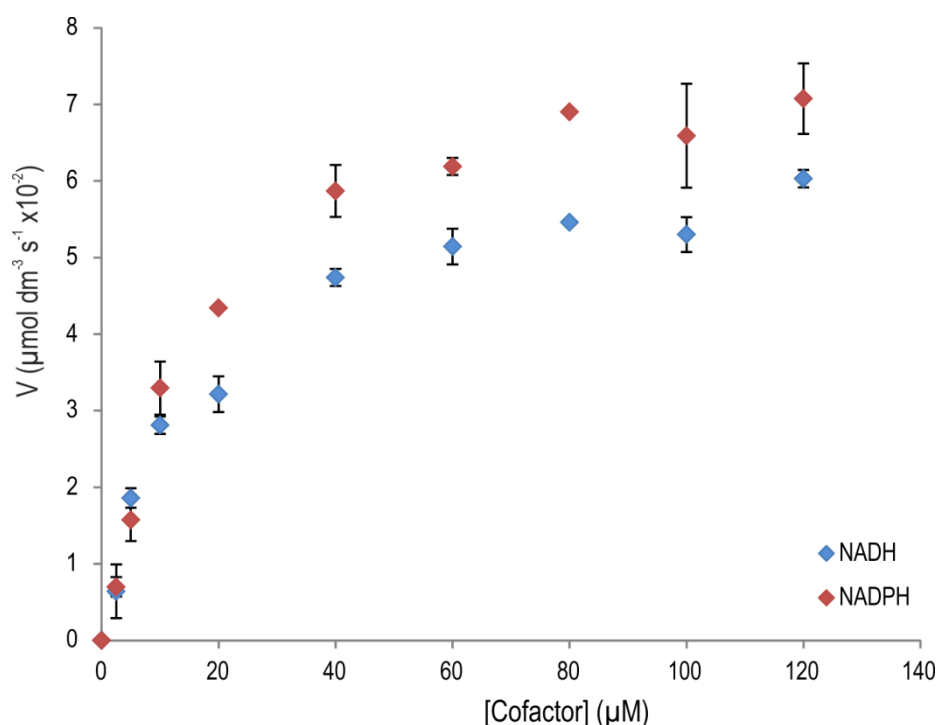
Figure 8.8 Kinetics of cofactor (NAD(P)H) utilisation by PFMO.

Similar to SMFMO and PFMO, CFMO was also able to use either cofactor to reduce the flavin (Figure 8.9).
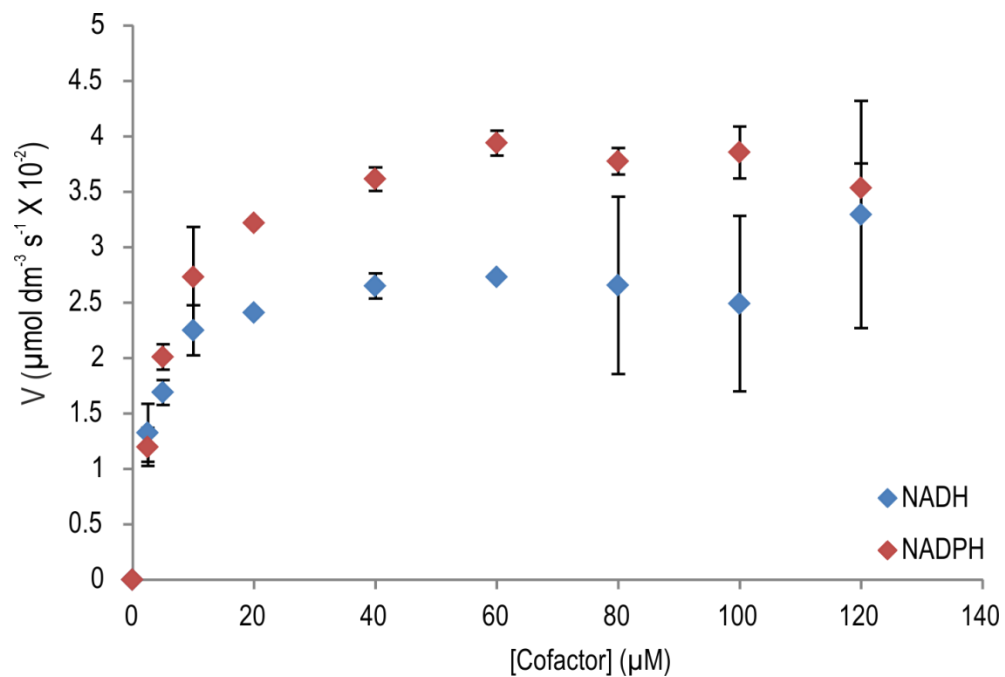
Figure 8.9 Kinetics of cofactor (NAD(P)H) utilisation by CFMO.

For both enzymes, PFMO and CFMO, the kinetic graphs demonstrated Michaelis-Menten behaviour, similar to that of SMFMO. For both NADH and NADPH assays the kinetic parameters were calculated in order to explore the effects PFMO and CFMO had on cofactor oxidation, the results are shown in Table 8.3.

Table 8.3 Kinetic constants for SMFMO, PFMO and CFMO using NADH or NADPH as cofactor.

Assay	K_M (μM)	V_{max} (μM s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
NADH (SMFMO)	23.7 ± 9.1	$11.2 \pm 1.5 \times 10^{-2}$	0.029	1223
NADPH (SMFMO)	27.3 ± 5.3	$8.4 \pm 0.6 \times 10^{-2}$	0.022	806
NADH (PFMO)	15.5 ± 2.3	$6.5 \pm 0.2 \times 10^{-2}$	0.0341	2273
NADPH (PFMO)	16.9 ± 1.8	$8.0 \pm 0.2 \times 10^{-2}$	0.0424	2510
NADH (CFMO)	3.22 ± 0.75	$2.9 \pm 0.1 \times 10^{-2}$	0.015	4711
NADPH (CFMO)	5.04 ± 0.57	$4.0 \pm 0.08 \times 10^{-2}$	0.0211	4184

For the enzyme PFMO the binding constant, K_M, was found to be comparable when either NADH (15.5 μM) or NADPH (16.9 μM) was used as nicotinamide cofactor. The turnover, k_{cat}, and catalytic efficiency, k_{cat}/K_M, when NADH was employed was also similar to that for NADPH (NADH: k_{cat} = 0.0341 s^{-1}, k_{cat}/K_M = 2273 M^{-1} s^{-1} versus NADPH: 0.0424 s^{-1}, k_{cat}/K_M = 2510 M^{-1} s^{-1}), indicating that the enzyme has the ability to utilise either cofactor with equal capability in order to reduce the bound flavin. The values observed for K_M with PFMO were noticeably smaller than those produced for SMFMO with both cofactors. The V_{max} values for PFMO were also lower than those for SMFMO, especially the value for the NADH assay, which was approximately 2 fold lower for PFMO (6.5×10^{-2} μM s^{-1} for PFMO compared to 11.2×10^{-2} μM s^{-1} for SMFMO).

The turnover value for PFMO (0.034 s^{-1}) when NADH was employed was higher than that observed for SMFMO (0.029 s^{-1}), thus, the higher turnover and lower K_M led to a 2 fold increase in catalytic efficiency for PFMO (2273 M^{-1} s^{-1}) compared to that observed for SMFMO (1223 M^{-1} s^{-1}). Similarly for the NADPH assay, the turnover value for PFMO (0.042 s^{-1}) was double that observed for SMFMO (0.022 s^{-1}), thus, the higher turnover and lower K_M led to a 3 fold increase in catalytic efficiency for PFMO (2510 M^{-1} s^{-1}) compared to that observed for SMFMO (806 M^{-1} s^{-1}). The results for PFMO suggest that the enzyme can utilise both cofactors for reduction of the flavin more efficiently than SMFMO.

The enzyme CFMO, had binding constants that were similar when both cofactors were used, with the greater binding for NADH (3.22 µM) than that for NADPH (5.04 µM). The turnover rate for the NADH assay (0.015 s^{-1}) was smaller than for NADPH (0.021 s^{-1}), thus giving a higher catalytic efficiency for CFMO when NADH (4711 M^{-1} s^{-1}) is employed compared to that for NADPH (4184 M^{-1} s^{-1}). However, the small difference between the catalytic parameters when either cofactor is employed suggests that CFMO can also utilise both cofactors with equal capability.

CFMO had the ability to bind NADH and NADPH with increased affinity compared to PFMO and SMFMO. When NADH was employed as cofactor the K_M values for CFMO were 5-fold lower than that observed for PFMO and had 8-fold decrease when compared with the binding constant for SMFMO. The turnover value for CFMO (0.015 s^{-1}) was half that observed for PFMO (0.034 s^{-1}) and SMFMO (0.029 s^{-1}), thus, the lower turnover and lower K_M led to an approximate 4 fold increase in catalytic efficiency for CFMO (4711 M^{-1} s^{-1}) compared to that observed for SMFMO (1223 M^{-1} s^{-1}) and an approximate 2-fold increase to that observed for PFMO (2273 M^{-1} s^{-1}), when NADH was utilised as cofactor. When NADPH was employed as cofactor the K_M values for CFMO were 3-fold lower than that observed for PFMO and had 5-fold decrease when compared with the binding constant for SMFMO. The turnover value for CFMO (0.021 s^{-1}) was half that observed for PFMO (0.042 s^{-1}) and lower than for SMFMO (0.029 s^{-1}), thus, the lower turnover and lower K_M led to an approximate 5-fold increase in catalytic efficiency for CFMO (4184 M^{-1} s^{-1}) compared to that observed for SMFMO (860 M^{-1} s^{-1}) and an approximate 1.5-fold increase to that observed for PFMO (2510 M^{-1} s^{-1}), when NADPH was utilised as cofactor. The results indicate that CFMO

can utilise both NADH and NADPH as nicotinamide cofactor for reduction of flavin more efficiently than SMFMO and PFMO.

8.4.4 Biotransformation of prochiral sulfides using PFMO and CFMO

SMFMO had the ability to catalyse the oxidation of a range of prochiral sulfides **13–19** to their corresponding sulfoxides, in the presence of NADH or NAPDH with the appropriate cofactor recycling system. To investigate PFMOs and CFMOs ability to catalyse the oxidation of substrates **13**-**19** and their enantioselectivity towards them with either NADH or NADPH as nicotinamide cofactor biotransformations were set up in the presence of NADH or NADPH (Section 3.4.4) and the extracted products were analysed by GC as detailed in Section 3.4.5. The results are shown in Table 8.4 and Table 8.5.

Table 8.4 Results of biotransformations of prochiral thioether substrates by PFMO.

Substrate	Conversion NADH %	Conversion NADPH %	*ee* NADH %	*ee* NADPH %	Sulfoxide Configuration NADH	Sulfoxide Configuration NADPH
13	61	28	30	24	*R*	*R*
14	-	-			-	-
15	97	78	47	32	*R*	*R*
16	99	73	14	4	*R*	*S*
17	13	-	10	-	*S*	-
18	10	2	18	11	*n.d.*	*n.d.*
19	50	22	85	57	*R*	*R*

Table 8.5 Results of biotransformations of prochiral thioether substrates by CFMO.

Substrate	Conversion NADH %	Conversion NADPH %	*ee* NADH %	*ee* NADPH %	Sulfoxide Configuration NADH	Sulfoxide Configuration NADPH
13	17	6	54	43	*R*	*S*
14	-	-	-	-	-	-
15	65	32	22	32	*R*	*R*
16	64	38	58	64	*R*	*R*
17	-	11	-	3	-	*R*
18	17	3	19	25	*n.d.*	*n.d.*
19	14	47	66	77	*R*	*R*

The results for PFMO revealed that the conversions of sulfides **13**-**19** to their corresponding sulfoxides and the enantiomeric excess of the products are higher when NADH is employed as nicotinamide cofactor compared to those obtained for the NADPH reactions. Substrate **14** was not converted with either cofactor. The highest conversion rate was seen for substrate **16**, 99 % for NADH and 73 % for NADPH, however *ee* was poor (14 % -(*R*) for NADH and 4 % -(*S*) for NADPH). The greatest ee values were found for substrate **19** with 85 % -(*R*) for NADH and 57 % -(*R*) for NADPH, albeit with moderate to low conversions (50 % for NADH and 22 % for NADPH, highlighted in Figure 8.10. PFMO gave the (*R*)-configuration for the majority of sulfoxide products except for substrate **16** when NADPH was employed as cofactor and **17** when NADH was used, which gave the (*S*)-enantiomers.
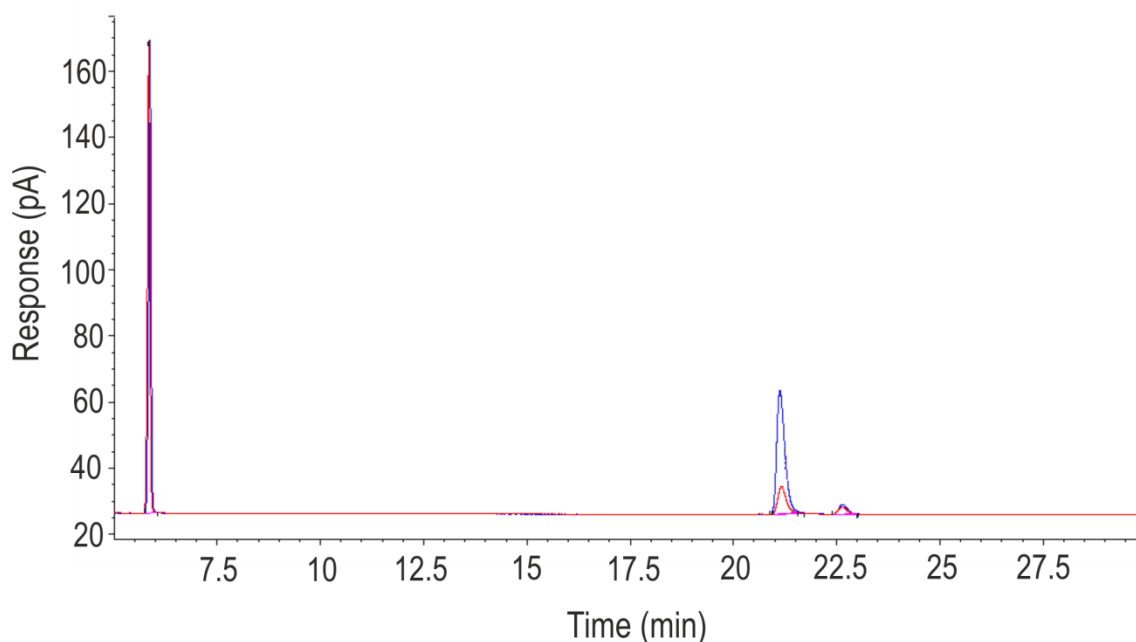
Figure 8.10 Chiral GC trace at t= 24 h for the oxidation of substrate **19** with PFMO. The trace for the NADH reaction is shown in blue and the trace for NADPH is shown in red.

In the majority of cases PFMO had greater conversions of sulfides to sulfoxides for each cofactor and poorer *ee* compared to those observed for SMFMO. For substrate **16** the conversions with PFMO (99 % for NADH and 73 % for NADPH) were considerably higher than the conversions for SMFMO (8 % for NADH and 1 % for NADPH), however the *ee*s for PFMO (14 % -(*R*) for NADH and 4 % -(*S*) for NADPH) were noticeably smaller than those produced when SMFMO was employed as enzyme (21 % -(*R*) for NADH and 34 % -(*n.d.*) for NADPH). In the case of substrate **18** the *ee* for PFMO with both cofactors were 2-fold lower than those observed for SMFMO. Interestingly for substrate **19**, the conversion rates for PFMO with both NADH and NADPH (50 % for NADH and 22 % for NADPH) were higher than that for SMFMO (40 % for NADH and 9 % for NADPH), with the *ee* in the case of PFMO and NADH (85 % -(*R*) for NADH) being higher than that observed for SMFMO (80 % -(*R*) for NADH). However the *ee* for PFMO and NADPH (57 % -(*R*) for NADPH) was poorer than that observed for SMFMO (82 % -(*R*) for NADPH). In addition, both PFMO and SMFMO produced sulfoxide products of the same (*R*)-configuration except for substrate **17**, in which the configurations was (*S*) with both enzymes.

164

Similar to PFMO, the results for CFMO revealed that the conversions of sulfides **13**-**18** to their corresponding sulfoxides and the enantiomeric excess of the products are higher when NADH is employed as nicotinamide cofactor compared to those obtained for the NADPH reactions. Substrate **14** was not converted with either cofactor. The highest conversion rate was seen for substrate **16**, 64 % for NADH and 38 % for NADPH, with moderate *ee* (58 % -(*R*) for NADH and 64 % -(*R*) for NADPH). As with PFMO, the greatest *ee* values for CFMO were found for substrate **19** with 66 % -(*R*) for NADH and 77 % -(*R*) for NADPH, again with low to moderate conversions (14 % for NADH and 47 % for NADPH, Figure 8.11). CFMO gave the (*R*)-configuration for the majority of sulfoxide products except for substrate **13** when NADPH was employed as cofactor which gave the (*S*)-enantiomers.
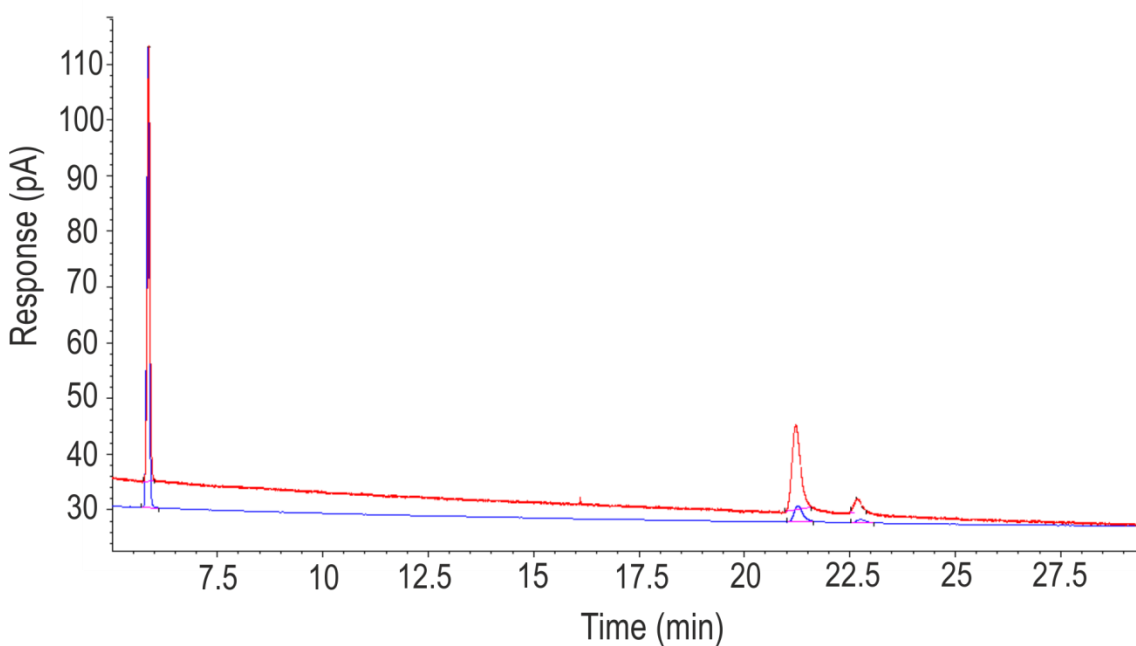
Figure 8.11 Chiral GC trace at t= 24 h for the oxidation of substrate **19** with CFMO. The trace for the NADH reaction is shown in blue and the trace for NADPH is shown in red.

When NADH was employed as nicotinamide cofactor CFMO gave lower conversion rates for the majority of sulfide substrates compared to SMFMO. In the case of substrate **16** however, CFMO had a conversion rate of 64 % which is considerably higher than that observed for SMFMO (8 % for NADH). In addition for **16** the *ee* was greater for CFMO (58 % -(*R*) for NADH) compared to that of SMFMO (21 % -(*R*) for NADH).

Interestingly, when NADPH was employed as nicotinamide cofactor CFMO gave higher conversion rates for the all the sulfide substrates compared to SMFMO, except **14**. A high conversion difference between CFMO and SMFMO was again observed for substrate **16** with 38 % for CFMO with NADPH and 1 % for SMFMO with NADPH. Similar to NADH, the *ee* when NADPH was employed as cofactor approximately doubled for CFMO (64 % -(*R*) for NADH) compared to the *ee* for SMFMO (34 % -(*R*) for NADH). SMFMO gave the highest *ee*s when **19** was employed as substrate (80 % -(*R*) for NADH and 82 % -(*R*) for NADPH), for CFMO however the *ee*s for **19** were noticeably poorer (66% -(*R*) for NADH and 77 % -(*R*) for NADPH), although from all the substrates **19** gave the highest *ee* for CFMO, PFMO and SMFMO. Interestingly, both CFMO and SMFMO produced sulfoxide products of the same (*R*)-configuration except for substrate **13**, in which CFMO gave the (*S*)-configuration which was opposite to that observed for SMFMO.

8.4.5 Structure determination

To investigate the ability of PFMO and CFMO to use either NADH or NADPH as cofactor to reduce the flavin, the enzymes were subject to crystallisation. The screens were successful for both PFMO and CFMO, producing yellow cubic crystals (Figure 8.12). The crystals were subject to X-ray diffraction allowing for a 1.83 Å dataset for PFMO and a 2.39 Å dataset for CFMO to be collected. The collected datasets allowed for the structures of both enzymes to be solved, built and refined to a resolution of 1.88 Å for PFMO and 2.45 Å for CFMO.
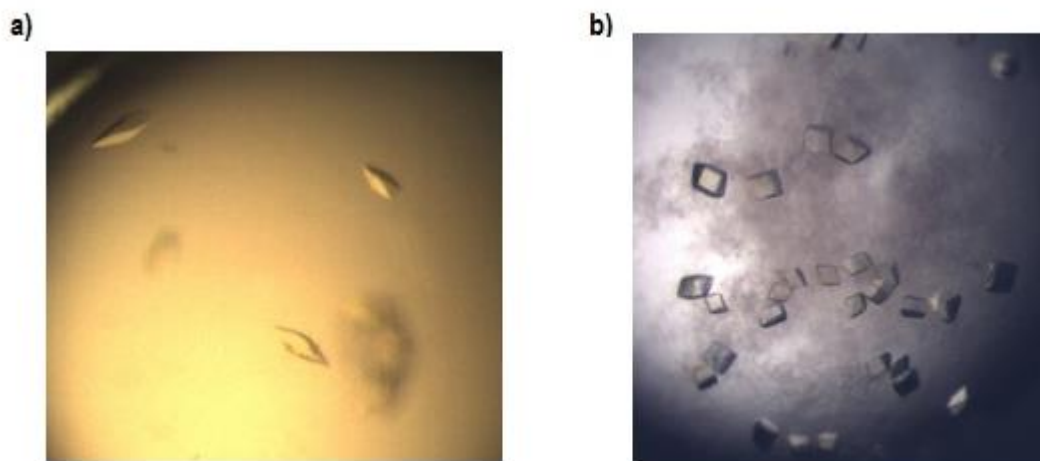
Figure 8.12 PFMO and CFMO crystals obtained from Clear Strategy Screen conditions. a) PFMO crystals with crystallisation conditions: 35 % w/v tacsimate at pH 7.0 b) CFMO crystals with crystallisation conditions: 1.5 M ammonium sulfate and propan-2-ol (10 μL).

Overall structure for PFMO

Crystals of PFMO grew in the $P3_22_1$ space group, with one molecule representing a monomer in the asymmetric unit (Figure 8.13). There was electron density for each amino acid in the monomer from residue Met1 to Pro354 with no chain breaks. Electron density was also present for residues Gly-2, Pro-1 and Ala0 at the N-terminal which are present after His_6-tag cleavage.
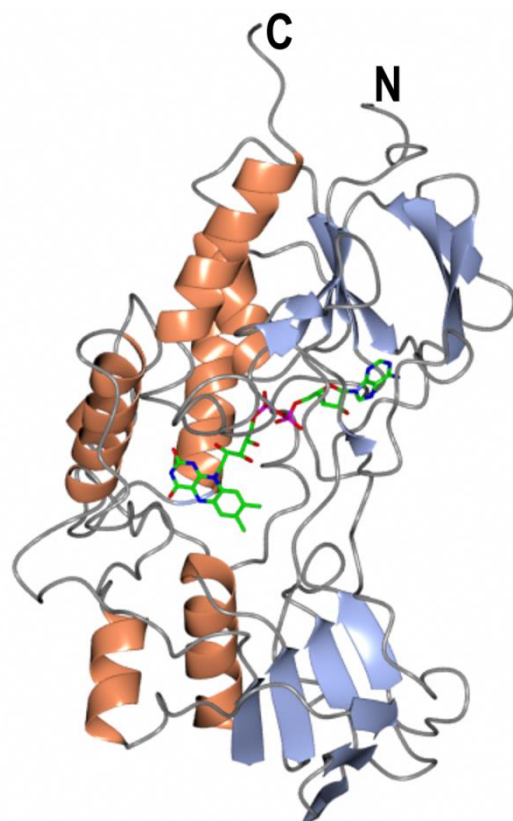
Figure 8.13 Structure of PFMO showing one molecule in the asymmetric unit.
Protein backbone is shown in ribbon format, β strands in blue and α helices in coral.
The FAD molecule is shown in cylinder format, with carbon atoms in green. The N and
C terminus for the subunit is also illustrated.

The PFMO monomer consisted of thirteen alpha helices: α1 (residues Gln14-Leu23), α2
(Gly41-His45), α3 (Ala56-Trp58), α4 (Arg77-Tyr91), α5 (Arg143-Leu145), α6
(Ser153-His155), α7 (Ala160-Phe162), α8 (Gly173-Ser184), α9 (Arg207-Ala219), α10
(Phe240-Glu247), α11 (Asp289-Leu291), α12 (Asp323-Thr325), α13 (Leu332-Leu350)
and sixteen beta strands: β1 (residues Leu5-Ile10), β2 (Tyr30-Leu33), β3 (Ile95-Gln96),
β4 (Thr103-Leu107), β5 (Leu110-Ala115), β6 (Gln118-Ser127), β7 (Ala149-His152),
β8 (Arg166-Val170), β9 (Glu188-Ile192), β10 (Arg259-Phe260), β11 (Gly264-Glu266),
β12 (Arg272-Asn274), β13 (Ala277-Trp280), β14 (Val305-Glu306), β15 (Thr308-
Val310) and β16 (Leu316-Leu318). The β-strands form four distinct sheets A (β1- β6,
β16), B (β7- β9, β13), C (β10- β12) and D (β14- β15). At the N-terminus of PFMO
there is a high β-strand presence whereas the C-terminus is highly helical. The
secondary elements for PFMO are summarised in Figure 8.14.

β1 α1

M P P I L D V I V I G G G Q A A L T T A Y F L R R T S L S   *29*

β2   α2   α3

Y L L L D E Q P G P G G A W L H A W D S L R L F S P A A W S S I A G W P M P S   *68*

α4   β3   β4

P T E P G N P T R N D V I D Y L R R Y E D R Y Q F P I Q R P V R V D T V T R L   *107*

β5   β6   α5

D D L W R V Q A G D Q Q W L A R A V I S A T G T W S K P F I P P Y E G R E L F   *146*

β7   α6   α7   β8   α8

Q G A Q I H S A H Y R T P A P F A G K R V M V V G G G N S G A Q V L A E L S S   *185*

β9   α9

V S E T L W I T Q E P P A F L P D E V D G R V L F E R A T A R W K A Q Q E G R   *224*

α10   β10

S I D E P A G G F G D I V M V P P V R E A R E R G V L V A E R P F A R F T E T   *263*

β11   β12   β13   α11

G V E W A D G R R E N L D A V I W C S G F R P A L D H L R E L G V V E A D G K   *302*

β14   β15   β16   α12   α13

V Q V E D T R V V K Q P N L W L V G Y G D W T G M A S A T L I G V T R T A R S   *341*

T A D Q V V Q A L T A T P S R R P   *358*

Figure 8.14 Diagram illustrating secondary structure of PFMO.
The α-helices are shown as red cylinders and the β-strands are shown as blue arrows.

Active site of PFMO

The PFMO monomer consists of two domains, a FAD binding domain and what is believed to be the substrate binding domain. The monomer has a molecule of FAD present in the active site which is bound by eight hydrogen bonds. The side chain N-H of Gln14 interacts with the oxygen (O2', 2.81 Å) closest to the tricyclic ring present in the riboflavin moiety. In addition, Gln14 also hydrogen bonds to the phosphate oxygen (O1P, 3.21 Å). Hydrogen bond interactions between the main chain nitrogen of residue Ala15 and the phosphate oxygen (O1P, 2.92 Å) also act to secure the FAD in place. Residues Gln14 and Ala15 are part of the Rossman motif in PFMO, present in α1, and are responsible for identifying the ADP moiety in FAD. Other hydrogen bond interactions binding the FAD in place are the side chain N-H of Gln36 to the ribose hydroxyl (O3B, 2.95 Å), the main chain nitrogen of Ala42 to the phosphate carbonyl

(O1A, 2.85 Å), the backbone N-H of residue Phe53 to the flavin carboxyl (O4, 2.84 Å), the carbonyl oxygen of Val101 to the N-H on the adenine moiety (N6A, 2.98 Å) and the backbone N-H of residue Leu332 to the carbonyl on the tricyclic ring (O2, 2.89 Å).

Overall structure of CFMO

Crystals of CFMO grew in the *C*2 space group, with two molecules 'A' and 'B', representing one dimer in the asymmetric unit Figure 8.15.
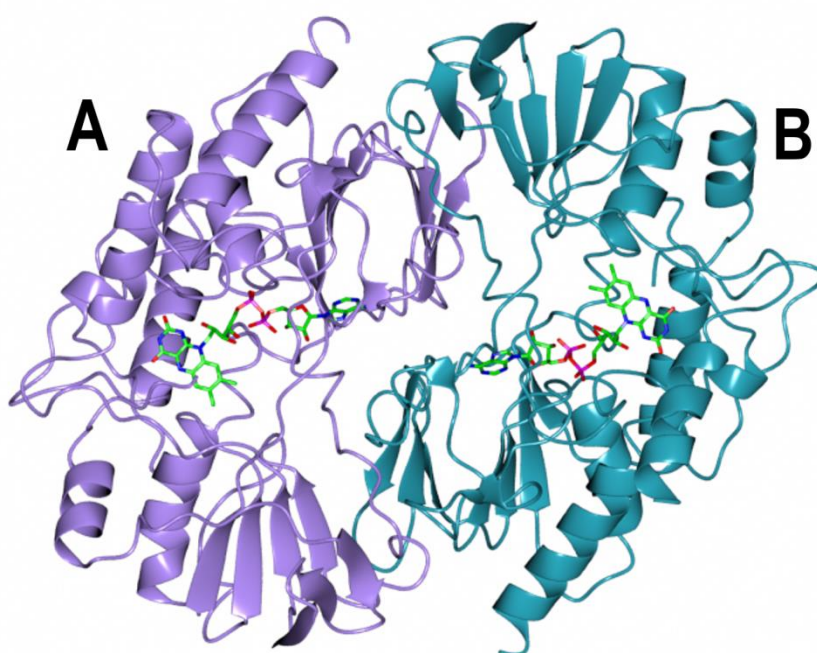
Figure 8.15 Quaternary structure of CFMO showing the two subunits A and B.
The protein backbone is shown in ribbon format with monomer A in lilac and monomer B in blue. The FAD molecules in each subunit are shown in cylinder format with the carbon atoms in green.

There was electron density for the majority of amino acids in each monomer from Ser13 to Ala361 with a stretch of missing density corresponding to eight amino acids from positions Gln236 to Asp243 (QPVGGLGD) that could not be modelled. It is believed that the missing residues are part of a flexible loop that sits over the active site containing the bound FAD (Figure 8.16). The CFMO dimer was made up of two monomers, sharing an interfacial area of approximately 1207 Å^2. Analysis of CFMO

using PISA[106] found that the interactions that stabilise the dimer included six hydrogen bonds between the backbone nitrogen of Trp108(A) and the side chain oxygen of Glu124(B) (2.79 Å), the side chain N-H of Gln320(A) interacts with the backbone carbonyl of Ile145(B) (3.23 Å), the backbone N-H of Ile145(A) and the backbone carbonyl of Gly303(B) (3.60 Å), the side chain N-H of Arg293(A) and the main chain carbonyl of Ile305(B) (2.48 Å), the main chain carbonyl of residue Gly303(A) interacts with the backbone N-H of Ile145(B) (3.67 Å) and between the backbone carbonyl of Ile145(A) and the side chain N-H of residue Gln320(B) (3.13 Å).
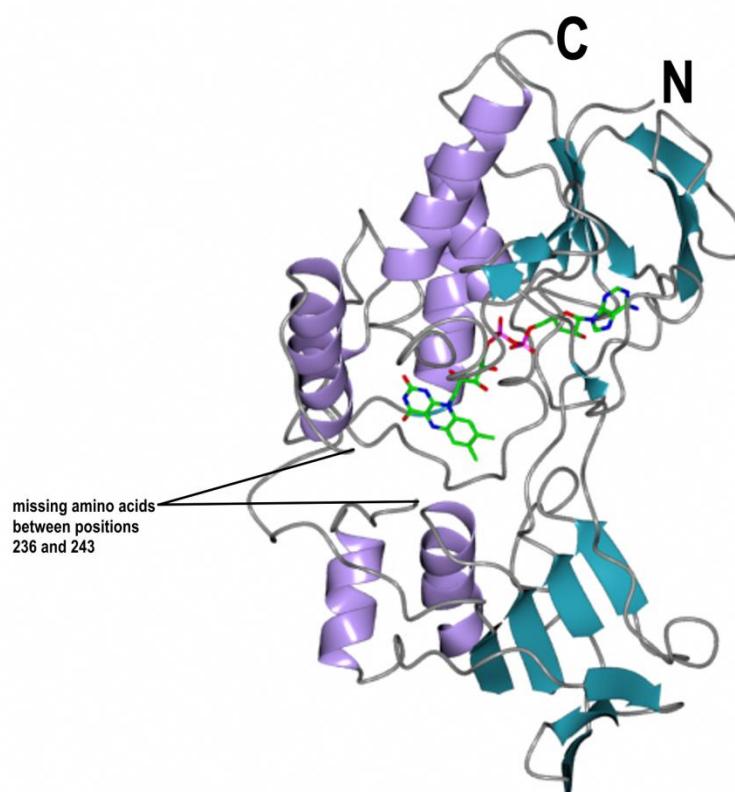
Figure 8.16 Structure of CFMO monomer A with complexed FAD.
The protein backbone is shown in ribbon format with β-strands in blue and α-helices in lilac. The FAD molecule is shown in cylinder format with carbon atoms in green. The N and C termini are also illustrated.

Each CFMO monomer consisted of twelve alpha helices: α1 (residues Gln23-Leu32), α2 (Gly50-His54), α3 (Ser65-Trp67), α4 (Arg85-Tyr99), α5 (Ser161-His163), α6 (Ala158-Phe170), α7 (Gly181-Val191), α8 (Arg215-Glu230), α9 (Pro248-Glu255), α10

(Ala397-Leu302), α11 (Glu331-Thr333), α12 (Val343-Leu358) and seventeen beta strands: β1 (residues Ser13-Ile19), β2 (Phe39-Leu42), β3 (Val103-Gln104), β4 (Ala111-Asn114), β5 (Leu119-Val122), β6 (Gln127-Ser135), β7 (Ala157-His160), β8 (Lys174-Val178), β9 (Asp196-Val200), β10 (Ala261-Val262), β11 (Arg267-Thr269), β12 (Gly272-Val274), β13 (Gly280-Ala282), β14 (Ala285-Trp288), β15 (Leu313-Ala314), β16 (Arg317-Leu318) and β17 (Leu324-Leu326). The β-strands form four distinct sheets A (β1- β6, β17), B (β7- β10, β14), C (β11- β13) and D (β15- β16). At the N-terminus of PFMO there is a high β-strand presence whereas the C-terminus is highly helical. The secondary elements for CFMO are summarised in Figure 8.17.
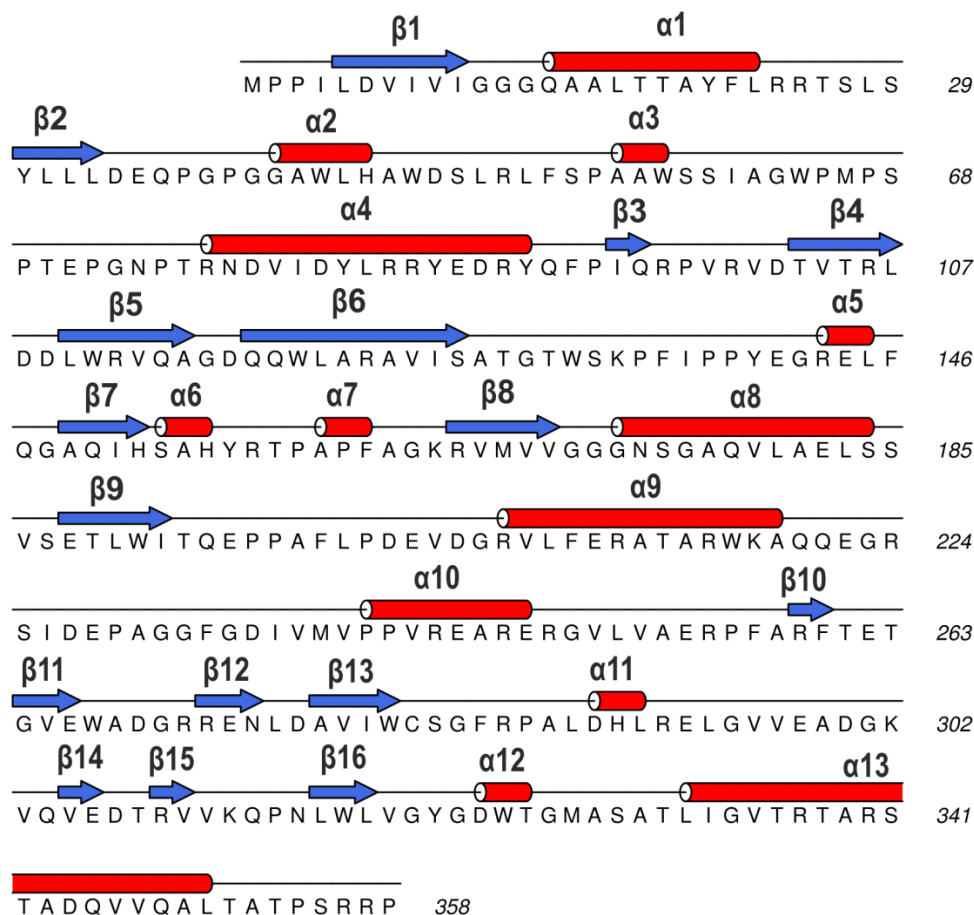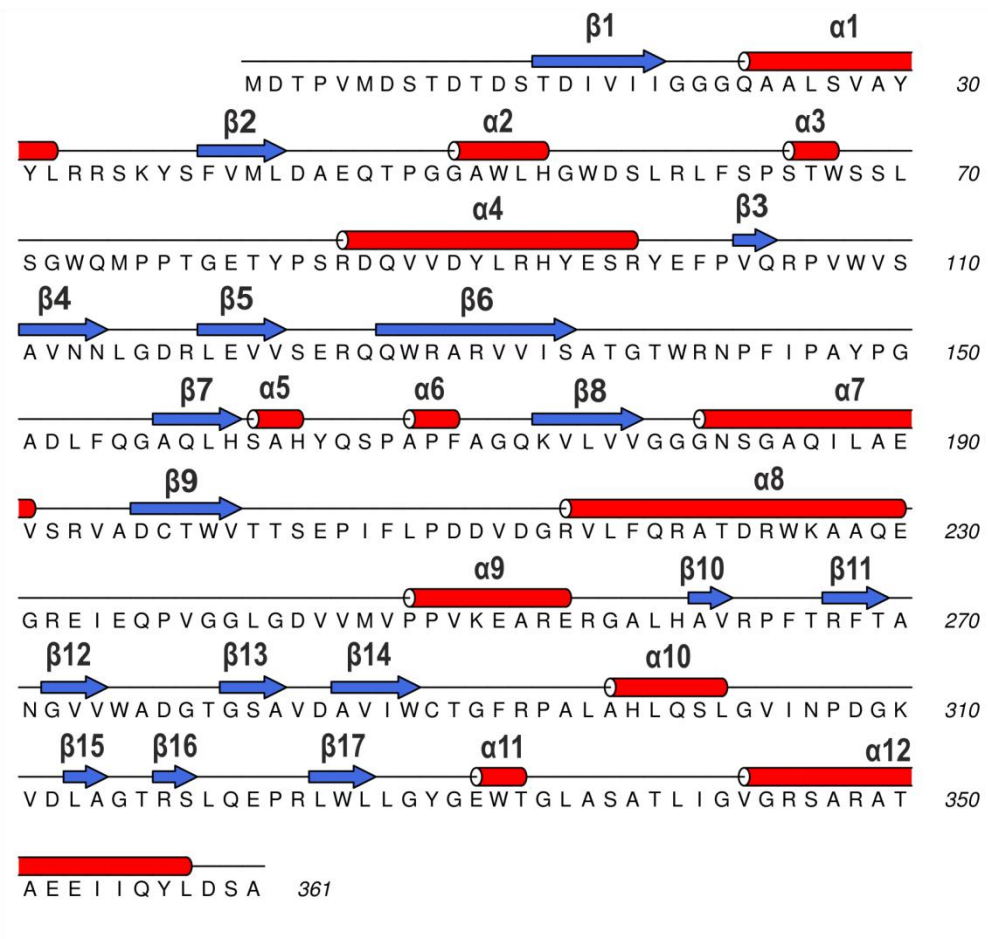

Figure 8.17 Diagram illustrating secondary structure of CFMO.
The α-helices are shown as red cylinders and the β-strands are shown as blue arrows.

172

Active site of CFMO

Each CFMO monomer consists of two domains, a FAD binding domain and what is believed to be the substrate binding domain. Each monomer has a molecule of FAD present in the active site which is bound by nine hydrogen bonds. The side chain N-H of Gln23 interacts with the oxygen (O2', 2.56 Å) closest to the tricyclic ring present in the riboflavin moiety. In addition, Gln23 also hydrogen bonds to the phosphate oxygen (O1P, 2.84 Å). Hydrogen bonding between the backbone nitrogen of residue Ala24 and the phosphate carbonyl (O1P, 2.84 Å) also act to secure the FAD within the catalytic site. Residues Gln23 and Ala24 are part of the Rossman motif in CFMO, present in α1, and are responsible for identifying the ADP moiety in FAD. Other hydrogen bond interactions securing the FAD in place are the backbone N-H of Ala51 to the phosphate carbonyl (O1A, 2.82 Å), the main chain N-H of residue Phe62 to the flavin carboxyl (O4, 2.75 Å), the carbonyl oxygen of Val109 to the N-H on the adenine moiety (N6A, 2.64 Å), the side chain N-H of residue Arg141 to the ribose hydroxyl (O2B, 3.09 Å) and between the backbone N-H of residue Leu340 to the carbonyl on the tricyclic ring (O2, 2.75 Å) and hydroxyl oxygen (O2', 3.72 Å).

8.5 Discussion

The structure of PFMO was found to be similar to that of SMFMO (4A9W[114], 60 % sequence identity, rmsd 0.9 Å over 327 Cα atoms), illustrated in figure 8.18.
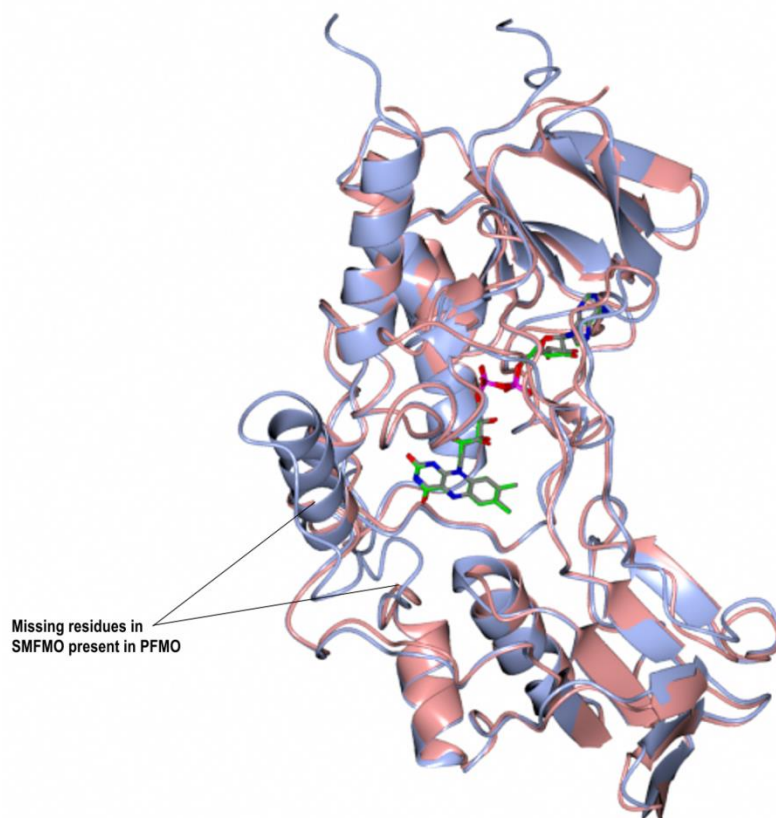
Figure 8.18 Superimposition of the PFMO monomer and SMFMO monomer A.
The protein backbone is shown in ribbon format. PFMO is shown in lilac with FAD
carbon atoms in green. SMFMO is show in pink with FAD carbon atoms in gold. The
superimposition was achieved using the ccp4mg programme.

The superimposition of PFMO and SMFMO indicates that the active sites of both flavo-
proteins sit in the same position indicated by the aligned FAD ligands. The majority of
α-helices, β-strands and loops are conserved within both structures, which could be
attributed to the conserved Rossman motifs and FMO motif. Interestingly, the amino
acids corresponding to the missing loop in SMFMO from Ala212 to Gly233
(ATERWKAQQEGREPDLPPGGFG) are present in PFMO. In PFMO the
corresponding residues Ala213 to Gly234 that are part of the α-helices, α9, and the loop
to α10. In SMFMO, the majority of α10 and the entire loop to α11 are missing. It is
thought that this loop is flexible and therefore can move to sit over the bound flavin in
the catalytic site.

The variant of SMFMO, PFMO, possessed a glutamate residue (Glu195) at the cofactor
determinant positions in SMFMO and it was hypothesised that the Glu195 would repel
the phosphate in NADPH and thus PFMO would be NADH specific. Interestingly this

was not the case; PFMO had the ability to use either NADH or NADPH as nicotinamide cofactor to reduce the flavin with equal capability, with comparable catalytic efficiencies. Owing to higher turnover rates and lower K_M PFMO had 2 fold increase in catalytic efficiency for NADH and a 3 fold increase for NADPH compared to the k_{cat}/K_M observed for SMFMO with either cofactor. As the structure of PFMO complexed with NAD(P)H has not yet been determined, the structure PFMO was superimposed with the structure of the nicotinamide binding site in SMFMO (Figure 8.19).

Figure 8.19 Superimposition of the nicotinamide binding site in SMFMO and PFMO structures.
The SMFMO structure is shown in worm format in pink with the sulfate ion in yellow. The PFMO structure is shown in worm format in lilac.

Superimposition of the nicotinamide cofactor binding site in SMFMO and PFMO revealed that the residues thought to be responsible for cofactor promiscuity in SMFMO, Gln193 and His194, sat in the same position as those in PFMO, Gln194 and Glu195. However, both Gln194 and Glu195 residues seem to be pointing away from the site in which the 2-hydroxyl phosphate would sit in NADPH (in SMFMOs case: the position of the sulfate ion) indicating that the glutamate residue would not actually

interact with the phosphate and therefore could not repel it. Thus, allowing PFMO to utilise either NADH or NADPH as cofactor. Superimposition of PFMO with the NADPH dependent mFMO structure (2XLT[103]) complexed with FAD and NADPH (Figure 8.20), found that residues possibly responsible for phosphate binding in mFMO Arg234 and Thr235 are indeed replaced by Gln194 and Glu195 in PFMO. As with the superimposition with SMFMO the Glu195 residue is pointing away from the phosphate indicating that it does not repel the phosphate in NADPH therefore allowing PFMO to utilise either cofactor. As with SMFMO the relaxation in nicotinamide cofactor in PFMO may be due to the removal of the interactions between the positively charged arginine and negatively charged oxygens on the phosphate. A variant of SMFMO with a glutamate at position 193 in SMFMO may have granted a NADH specific FMO however one could not be found.
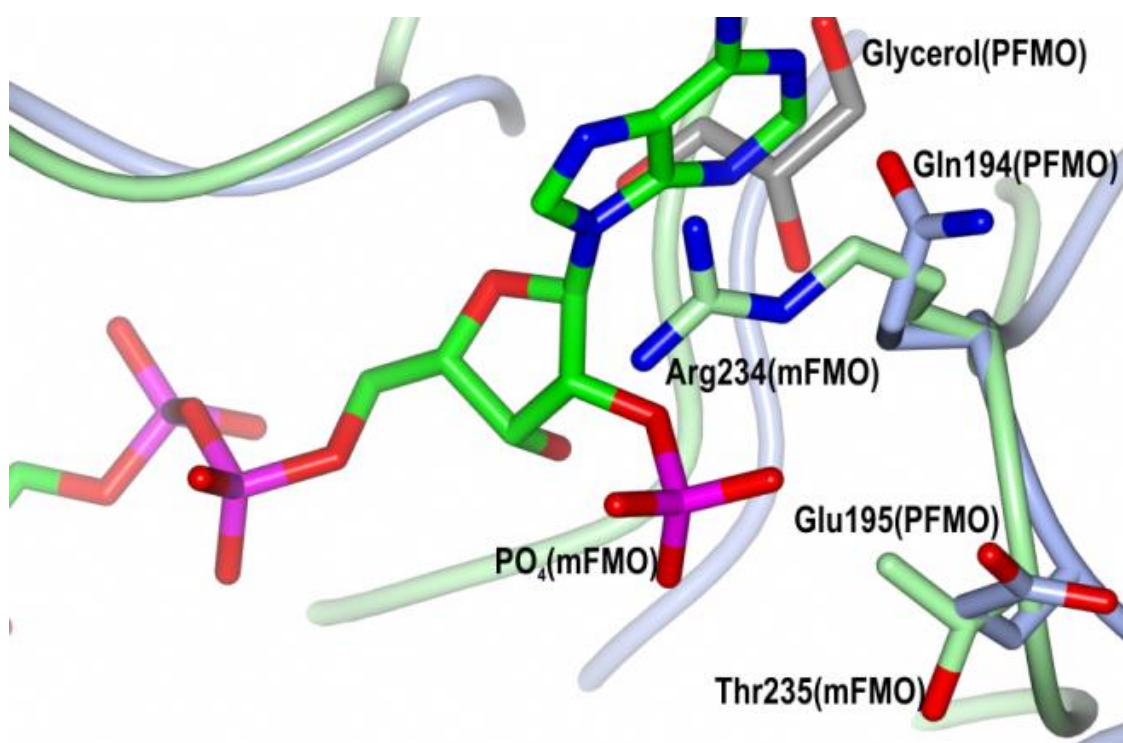
Figure 8.20 Superimposition of the NAD(P)H binding site of mFMO and PFMO. NADP ribose 2'-phosphate recognition site in mFMO (backbone shown in worm format in green and NADPH in cylinder format with carbon atoms in green), superimposed with structurally homologous regions of PFMO (shown in worm format in lilac).

176

Similar to PFMO, the monomeric tertiary structure of CFMO was found to be similar to that of SMFMO (4A9W, 59 % sequence identity, rmsd 0.9 Å over 325 Cα atoms), shown in Figure 8.21.
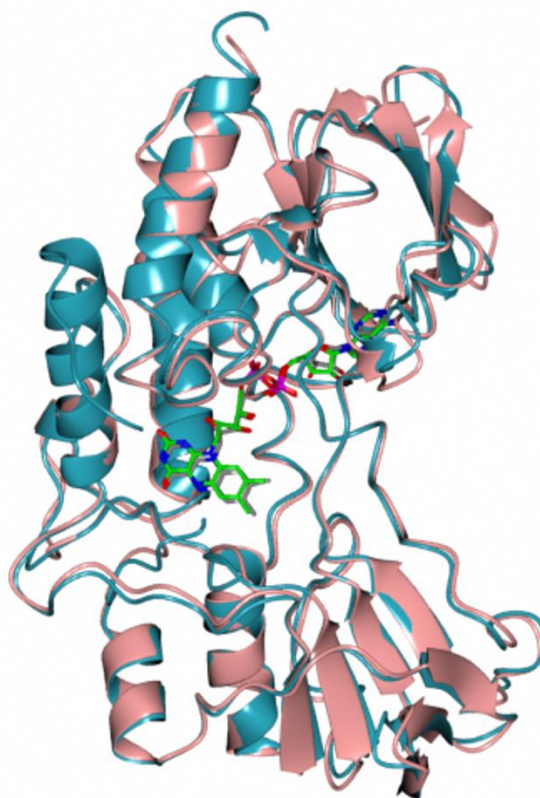
Figure 8.21 Superimposition of CFMO monomer A and SMFMO monomer A.
The protein backbone is shown in ribbon format. CFMO is shown in blue with FAD carbon atoms in green. SMFMO is show in pink with FAD carbon atoms in gold. The superimposition was achieved using the ccp4mg programme.

The superimposition of CFMO and SMFMO indicates that the active sites of both flavoproteins sit in the same position indicated by the aligned FAD ligands. Within both structures the extent of α-helices, β-strands and loops are conserved, which could be attributed to the conserved Rossman motifs and FMO motif. The amino acids corresponding to the missing loop in SMFMO from Ala212 to Gly233 (ATERWKAQQEGREPDLPPGGFG) are partly present in CFMO. In CFMO the corresponding residues Ala221 to Gly242 are part of the α-helices, α8, and the loop to α9. However eight amino acids from positions Gln236 to Asp243 (QPVGGLGD) are not modelled, leaving an incomplete loop to α9. It is thought that this loop is flexible

which sits over the bound flavin in the active site. However as with SMFMO and CFMO its flexibility results in amino acids that cannot be modelled.

CFMO was found to be another variant of SMFMO that possessed a threonine residue (Thr202) and a serine residue (Ser203) at the cofactor determinant positions in SMFMO (Gln193 and His194). This variantled to the hypothesis that the smaller hydrophilic residues Thr202 and Ser203 present in CFMO would allow the enzyme to have increased NADH specificity. As with PFMO this was not the case; CFMO had the ability to utilise either NADH or NADPH as nicotinamide cofactor with equal capability, with comparable catalytic efficiencies and binding constants. Owing to lower turnover rates and significantly lower K_M CFMO had 4 fold increase in catalytic efficiency for NADH and a 5 fold increase for NADPH compared to the catalytic efficiency observed for SMFMO with either cofactor. Interestingly, CFMO was able to use either NADH or NADPH as nicotinamide cofactor for the reduction of flavin more efficiently than SMFMO and PFMO. As the structure of CFMO complexed with NAD(P)H has not yet been determined, the structure CFMO was superimposed with the structure of the nicotinamide binding site in SMFMO (Figure 8.22). Superimposition of the nicotinamide cofactor binding site in SMFMO and CFMO revealed that the residues thought to be responsible for cofactor promiscuity in SMFMO, Gln193 and His194, sat in the same position as those in CFMO, Thr202 and Ser203. Both Thr202 and Ser203 residues are much smaller to the Gln193 and His194 residues present in SMFMO thus opening up the nicotinamide binding site. The larger cofactor site could therefore allow CFMO to utilise either NADH or its larger phosphorylated neighbour NADPH as cofactor.
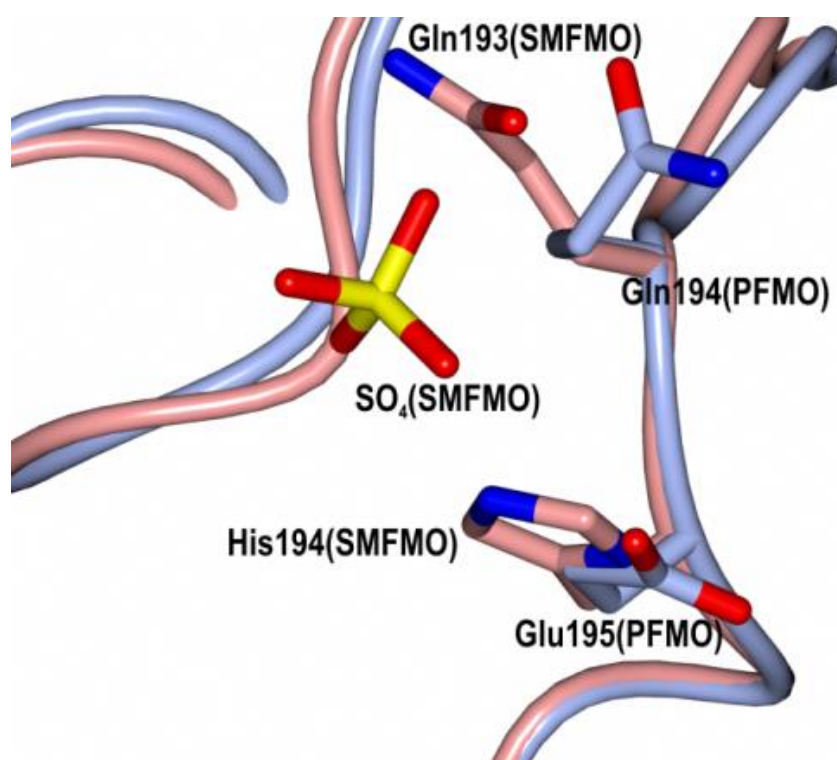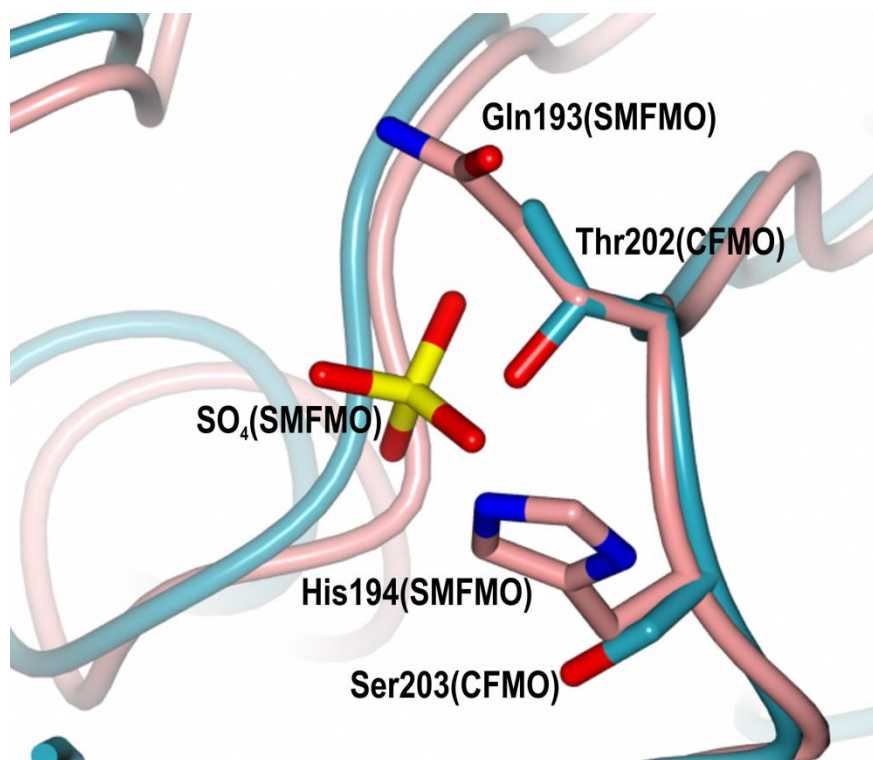
Figure 8.22 Superimposition of the nicotinamide binding site in SMFMO and CFMO structures.
The SMFMO structure is shown in worm format in pink with the sulfate ion in yellow.
The CFMO structure is shown in worm format in blue.

Superimposition of CFMO with the NADPH dependent mFMO structure (2XLT[103]) complexed with FAD and NADPH (Figure 8.23), found that residues possibly responsible for phosphate binding in mFMO Arg234 and Thr235 are also replaced by Thr202 and Ser203 in CFMO. The Thr234 residue in mFMO is partially conserved with the Ser203 in CFMO, which sits further than 4 Å away from the ribose 2' phosphate in NADPH and therefore will potentially have no interaction with the cofactor. As with the superimposition with SMFMO the Thr202 is much smaller than the Arg234 residue in mFMO which may allow CFMO to utilise either cofactor. In FPMOs such as mFMO[103], PAMO[61a] and CHMO[59] the arginine residue is essential for the recognition of the 2' phosphate in NADPH therefore the relaxation in nicotinamide cofactor in CFMO may also be due to the removal of the interactions between the positively charged arginine and negatively charged oxygens on the phosphate.
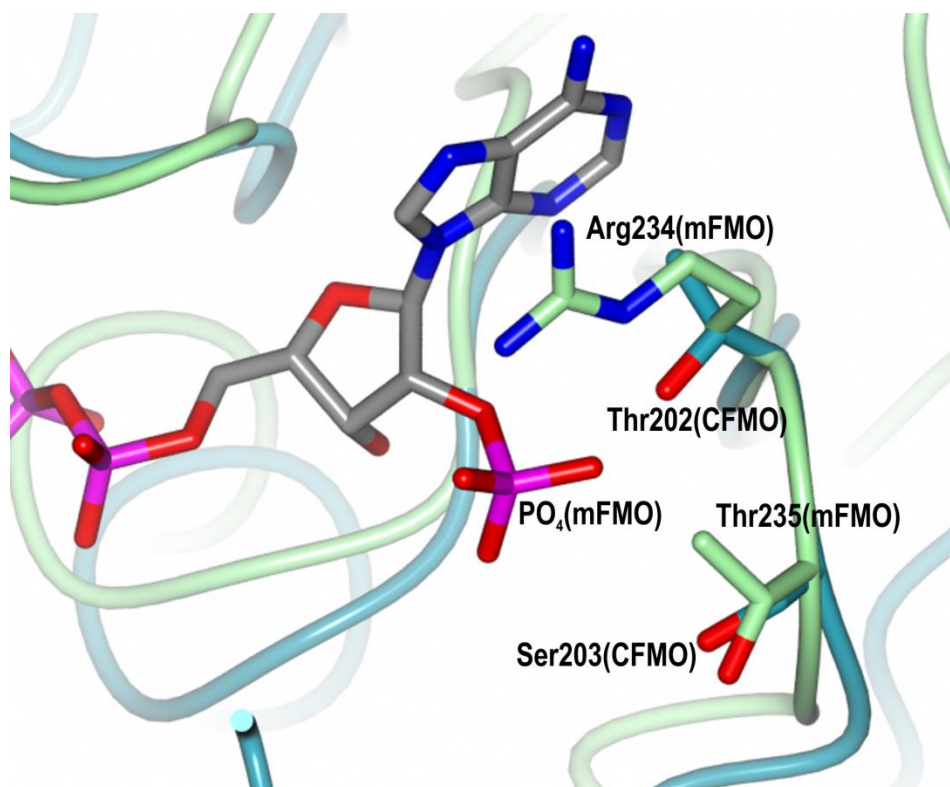
Figure 8.23 Superimposition of the NAD(P)H binding site of mFMO and CFMO. NADP ribose 2'-phosphate recognition site in mFMO (backbone shown in worm format in green and NADPH in cylinder format with carbon atoms in grey), superimposed with structurally homologous regions of CFMO (shown in worm format in blue).

Both enzymes PFMO and CFMO had the ability to catalyse the oxidation of a range of prochiral sulfides using either NADH or NADPH as cofactor. The successful reactions are summarised in Table 8.6 and Table. 8.7.

Table 8.6 Summary of results for biotransformations for SMFMO, PFMO and CFMO when NADH is used as cofactor.

Substrate	Conversion %			ee %			Configuration		
	SMFMO	PFMO	CFMO	SMFMO	PFMO	CFMO	SMFMO	PFMO	CFMO
13	27	61	17	71	30	54	*R*	*R*	*R*
15	90	97	65	25	47	22	*R*	*R*	*R*
16	8	99	64	21	14	58	*R*	*R*	*R*
19	40	50	14	80	85	66	*R*	*R*	*R*

Table 8.7 Summary of results for biotransformations for SMFMO, PFMO and CFMO when NADPH is used as cofactor.

Substrate	Conversion %			ee %			Configuration		
	SMFMO	PFMO	CFMO	SMFMO	PFMO	CFMO	SMFMO	PFMO	CFMO
13	2	28	6	57	24	43	*R*	*R*	*S*
15	33	78	32	44	32	32	*R*	*R*	*R*
16	1	73	38	34	4	64	*R*	*R*	*R*
19	9	22	47	82	57	77	*R*	*R*	*R*

The reactions involving PFMO allowed for greater conversion of substrate to product using either cofactor, when compared to SMFMO and CFMO. However, the conversions utilising NADH were slightly higher. In most cases the *ee*s were lower for PFMO than those observed for SMFMO[114], CHMO[23a, 98] and mFMO[73], however, PFMO had a greater ee for substrate **19** when NADH was employed. For substrates **13**, **15** and **16** the enantioselectivity was higher when NADH was used as cofactor suggesting that PFMO may have a 'greater preference' for one enantiomer with NADH than NADPH and thus substrate binding is different depending what cofactor is used. The (*R*)-enantiomer was the preferred configuration for the sulfoxide products for PFMO similar to SMFMO.

CFMO on the other hand had greater conversions of substrates when NADPH was employed and lower conversions using NADH, when compared to SMFMO, except in the case of substrate **16**. However, the conversions utilising NADH were noticeably higher for CFMO. As with PFMO, the *ee*s for the majority of the substrates were lower for CFMO than those observed for SMFMO, CHMO and mFMO. Compared to SMFMO, substrate **16** allowed for much greater *ee*s when either NADH or NADPH was used as cofactor suggesting that enantiomeric excess may be dependent on substrate structure. The *ee* for each substrate were similar regardless of cofactor indicating that, unlike PFMO and SMFMO, substrate binding may not depend on that cofactor is used, similar to the behaviour of the type II FMOs from *R. jostii* RHA1.[18] However, in the case of substrate **13** the (*S*)-enantiomer is preferred when NADPH was employed suggesting that cofactor may actually play a role in substrate binding. For the majority

of products the (*R*)-enantiomer was the preferred configuration for CFMO similar to PFMO and SMFMO.

NADP[+] has been reported to stabilise the flavin peroxidate species within the FPMO family.[103] As with SMFMO, PFMO and CFMO favour NADH for catalysis rather than NADPH suggesting that such stabilisation is achieved by NADH in these enzymes. In mFMO it was found that residue Asn78 stabilises the oxygenating species and mutation of this residue was found to knock out activity.[103] However, in other cases mutations in this position had little effect on activity.[108] PFMO and CFMO have a large, hydrophobic phenylalanine in this position (Phe53 and Phe62). SMFMO was also found to possess a phenylalanine at this position (Phe52) and mutagenesis studies revealed that this residue was not an absolute requirement but did have an interesting effect on active site recognition of sulfide substrates, as the mutation to the smaller amino acid valine altered the enantioselectivity.[115]

8.6 Conclusion

PFMO and CFMO are variants of SMFMO and were found to accept either NADH or NADPH as cofactor in order to reduce the flavin. The structures of PFMO and CFMO were solved which allowed for a basis for cofactor promiscuity for each enzyme to be suggested and avenues for future rational investigation studies to be proposed.

Chapter 9: Conclusion and future work

The over-expression of SMFMO when utilising *E. coli* BL21 (DE3) cells proved to be very successful in yielding yellow protein that purifies well by nickel affinity and size exclusion chromatography.

Kinetic studies found that SMFMO had the ability to employ either NADH or NADPH as nicotinamide cofactor to reduce the non-covalently bound flavin with equal capability. The only ketone yet to be transformed by SMFMO was the strained bicyclo[3.2.0]hept-2-en-6-one substrate. Both NADH and NADPH could be used as cofactors however, NADH allowed for a much greater conversion but enantioselectivity was poor for both. SMFMO had much more success with the monooxygenation of prochiral sulfides to the corresponding sulfoxides. Both cofactors could be employed although, greater conversion were obtained with NADH as cofactor. In the majority of cases, the sulfoxides products had low-moderate enantiomeric excess. The low substrate scope with respect to ketones but tolerance for sulfides indicates SMFMO behaves more like an FMO than a BVMO, which was expected as SMFMO possesses the 'FMO motif'. Such results are interesting as the majority of BVMOs and FMOs utilize NADPH as cofactor and for SMFMO, NADH seems to be the cofactor of choice which is attractive for biocatalytic applications as NADH is cheaper than NADPH.

Structural studies allowed for the structure of SMFMO to be determined and analysis and comparison of the active site and cofactor recognition site with known NADPH-dependent FMOs identified determinants for cofactor promiscuity and enantioselectivity in SMFMO. The mutational studies identified SMFMO residues Gln193 and in particular His194 as determinants for cofactor promiscuity and suggests avenues for the engineering of other FPMOs for NADH specificity. The residue Phe52 was found to have an interesting effect on the active sites recognition of sulfides in SMFMO, such that upon the mutation to a smaller amino acid side-chain the enantioselectivity of the sulfoxide product is altered. Therefore, the investigation into the nature and role of this residue in SMFMO and structurally homologous residues in other FPMOs, such as Asn78, is still ongoing.

The SMFMO variants, PFMO and CFMO were also found to accept either NADH or NADPH as nicotinamide cofactor to reduce the flavin for catalysis.

In addition to the BVMO mekA[112] and the groups of FMOs from *R. jostii* RHA1,[18] SMFMO is interesting amongst the FPMOs as it one of the only enzymes to demonstrate acceptance of NADH as nicotinamide cofactor. The characterisation of SMFMO, PFMO and CFMO has identified single-component class B FPMOs capable of NADH-dependent monooxygenation reactions and determined the structural basis for possible residues responsible for relaxed cofactor specificity in these enzymes. This will help highlight other related enzymes in the database and possible platforms for developing NADH-specific FPMOs.

In an effort to shed more light on the cofactor acceptance of these enzymes, future work could involve the possible attempts to co-crystallise SMFMO, PFMO and CFMO with either cofactor and/or substrate. This would allow for the interactions between cofactor and binding site and the substrate and active site to be determined, thus providing solid evidence for rational mutagenesis in order to make the enzymes fully NADH-specific with high enantioselectivity. In addition, another future investigation could be the mutation of the Gln194 residue on PFMO to glutamate, as this could not be achieved for SMFMO. The cofactor determinant residues in PFMO would then be Glu194 and Glu195, which could allow for strict NADH-dependence as the two glutamates may repel the negatively charge phosphate in NADPH.

Appendix One: Investigating the NADH recycling system

Introduction

FPMOs require a nicotinamide cofactor in order to reduce the bound FAD for catalysis, however to reduce costs a cofactor recycling system is employed. The cofactor recycling system comprises of an auxiliary enzyme and auxiliary substrate that do not interfere with the reaction taking place. In the case of SMFMO, NADH was found to be employed as cofactor in addition to NADPH. The recycling system often used to regenerate NADH is formate dehydrogenase (FDH) along with sodium formate (NaF). In an effort to determine the best ratio of SMFMO to FDH and therefore the best ratio for catalysis, FDH was cloned, expressed and purified in order to be employed in biotransformations alongside FMO.

Methods & Materials

The gene coding the formate dehydrogenase from *Pseudomonas sp.* (strain 101) (*Achromobacter parvulus* T1) was codon optimised for *E. coli* and purchased from Geneart in vector pMA-T. The gene sequences for the targets are shown in Figure A1.1.

```
>FDH 1203 bp
ATGGCCAAAGTTCTGTGTGTGCTGTATGATGATCCGGTTGATGGTTATCCGAAAACCTAT
GCACGTGATGATCTGCCGAAAATTGATCATTATCCGGGTGGTCAGACCCTGCCGACCCCG
AAAGCAATTGATTTTACACCGGGTCAGCTGCTGGGTAGCGTTAGCGGTGAACTGGGTCTG
CGTAAATATCTGGAAAGCAATGGTCATACCCTGGTTGTTACCAGCGATAAAGATGGTCCG
GATAGCGTTTTTGAACGTGAACTGGTTGATGCCGATGTTGTTATTAGCCAGCCGTTTTGG
CCTGCATATCTGACACCGGAACGTATTGCAAAAGCCAAAAATCTGAAACTGGCACTGACC
GCAGGTATTGGTAGCGATCATGTTGATCTGCAGAGCGCAATTGATCGTAATGTTACCGTT
GCAGAAGTGACCTATTGTAATAGCATTAGCGTTGCCGAACATGTGGTTATGATGATTCTG
AGCCTGGTTCGTAATTATCTGCCGAGCCATGAATGGGCACGTAAAGGTGGTTGGAATATT
GCAGATTGTGTTAGCCATGCATATGATCTGGAAGCTATGCATGTTGGCACCGTTGCCGCA
GGTCGTATTGGTCTGGCAGTTCTGCGTCGTCTGGCTCCGTTTGATGTTCATCTGCATTAT
ACCGATCGTCATCGTCTGCCGGAAAGCGTTGAAAAAGAACTGAATCTGACCTGGCATGCA
ACCCGTGAAGATATGTATCCGGTTTGTGATGTTGTGACCCTGAATTGTCCGCTGCATCCG
GAAACCGAACATATGATTAACGATGAAACCCTGAAACTGTTTAAACGCGGTGCCTATATT
GTTAATACCGCACGTGGTAAACTGTGTGATCGTGATGCAGTTGCACGTGCCCTGGAAAGC
GGTCGCCTGGCAGGTTATGCCGGTGATGTGTGGTTTCCGCAGCCTGCACCGAAAGATCAT
CCGTGGCGTACCATGCCGTATAATGGTATGACACCGCATATTAGCGGCACCACCCTGACC
GCACAGGCACGTTATGCAGCAGGCACCCGTGAAATTCTGGAATGTTTTTTTGAAGGTCGT
CCGATTCGTGATGAATATCTGATTGTTCAGGGTGGCGCACTGGCAGGTACAGGTGCACAT
AGCTATAGCAAAGGTAATGCAACCGGTGGTAGCGAAGAAGCAGCAAAATTCAAAAAAGCC
GTG
```

Figure A1.1 The codon optimised gene sequence for FDH.

LIC

The FDH gene was cloned into the pET-YSBLIC-3C vector using the LIC protocol described in Section 3.1.2. Each gene was amplified using the protocol in Section 3.1.1. The primers used are shown in Table A1.1.

Table A1.1 Primers used in PCR for FDH gene amplification

Gene	Forward primer	Reverse primer
FDH	CCAGGGACCAGCAATGGCCA AAGTTCTGTGTGTGCTGTATG ATGATC	GAGGAGAAGGCGCGTTACACGG CTTTTTTGAATTTTGCTGCTTCTT CGC

The insert was treated with T4 polymerase as detailed in Section 3.1.4 and the vector and insert were annealed together as described in Section 3.1.5.

Expression and Purification

The pET-YSBLIC-3C vector containing the FDH gene was expressed in *E. coli* BL21 (DE3) cells (Section 3.2.5) and the protein was purified using Ni^{2+} affinity chromatography followed by size exclusion chromatography described in Section 3.3.

Biotransformations

Varying NADH concentrations

To a screw cap vial (5 mL) containing buffer A (1 mL) substrate **15** (5 mM), NaF (0.5 M), SMFMO (200 μL at 3 mg mL^{-1}), FDH (278 μL at 2.5 mg mL^{-1}) and varying concentrations of NADH (0.1, 0.2 and 0.5 mM) were added and placed in an incubator (r.t.) for 24 h.

For each reaction time points at t= 1, 2 , 4, 6 and 24 h were taken and the product was extracted using EtOAc (Section 3.4.4) and analysed by GC (Section 3.4.5).

Varying SMFMO:FDH ratio

To a screw cap vial (5 mL) containing buffer A (1 mL) substrate **15** (5 mM), NaF (0.5 M), NADH (0.5 mM) and varying amounts of SMFMO and FDH (ratios (SMFMO:FDH) 1:4, 1:2, 1:0.5, 1:0.25, 1:0.1) were added and placed in an incubator (r.t.) for 24 h.

For each reaction time points at t= 1, 2 , 4, 6 and 24 h were taken and the product was extracted using EtOAc (Section 3.4.4) and analysed by GC (Section 3.4.5).

Results

Cloning, expression and purification

The FDH gene was cloned into the pET-YSBLIC-3C and expressed in *E. coli* BL21 cells, with a total volume of 2 L, successfully. The resulting 44 kDa protein was purified by Ni^{2+} affinity chromatography followed by size exclusion chromatography (Section 3.3) yielding pure FDH in high concentrations (25 mg).

Biotransformations

Varying NADH concentrations

SMFMO had the ability to employ NADH as cofactor to reduce the bound flavin. In order to investigate the amount of NADH needed to allow for greater substrate conversion, varying concentrations of the cofactor were employed in the biotransformation of substrate **15**. The results can are shown in Figure A1.2
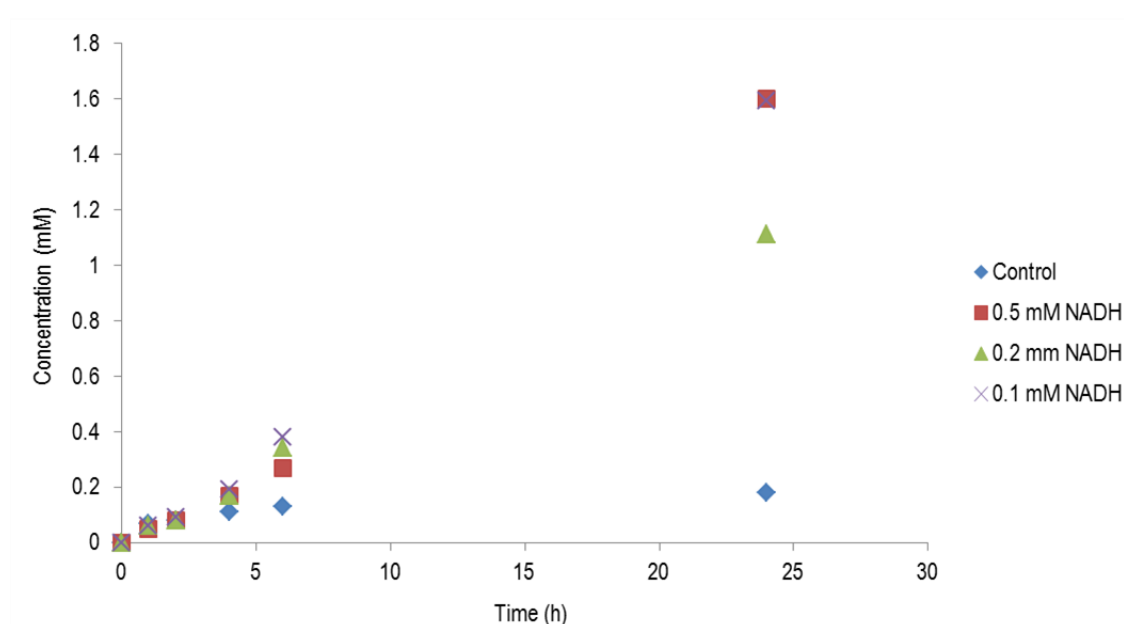
Figure A1.2 Graph illustrating the concentration of sulfoxide product from the oxygenation of substrate **15** when varying amounts of NADH cofactor is utilised at different time points.

Between 1 and 6 h the concentration of product is similar for each NADH concentration. After 24 h the concentration of the sulfoxide product is similar for both the 0.5 mM and 0.1 mM NADH concentrations. In addition, for the 0.2 mM NADH reaction the concentration of product is still above 1 mM. The control biotransformation utilised 0.5 mM of NADH in the absence of the FDH recycling system and the concentration of sulfoxide product plateaued after 4 h. The results indicate that the recycling system is effective and as little as 0.1 mM of cofactor can be utilised to reduce the flavin.

Varying SMFMO:FDH ratios

In an effort to investigate the most effective SMFMO to recycling system ratio, and thus regenerate the NADH cofactor efficiently for catalysis a number of biotransformations were carried out utilising various ratios of enzyme to FDH. The results are shown in Figure A1.3.
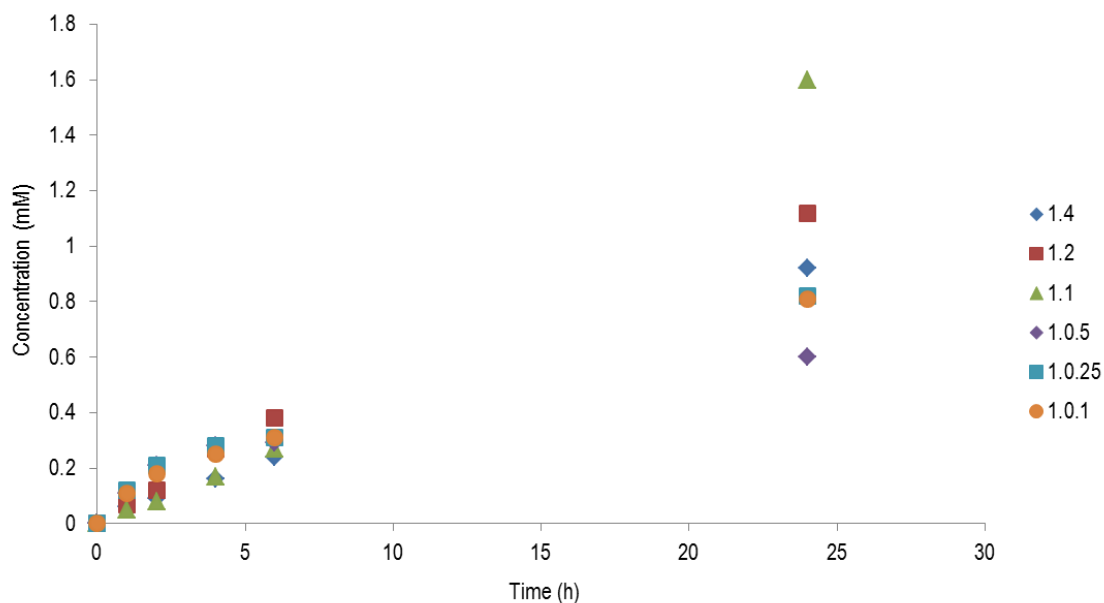
Figure A1.3 Graph illustrating the concentration of sulfoxide product from the oxygenation of substrate **15** when varying SMFMO to FDH ratios are utilised at different time points.

Between 1 and 6 h the concentration of the sulfoxide product is similar for all SMFMO:FDH ratio biotransformation reactions indicating that the conversion of sulfide 15 to its corresponding sulfoxide is not dependent on the SMFMO:FDH ratio. However, after 24 h the concentrations for the sulfoxide products was greatest for the 1:1 ratio of SMFMO and FDH.

Appendix Two: Fusion of SMFMO and cofactor recycling enzyme FDH

Introduction

As FDH and NaF were found to be an appropriate recycling system for NADH in biotransformations involving SMFMO the attempted fusion of SMFMO and FDH was attempted. This would allow for the expression of both genes within the same vector, thus eliminating the separate expression and protein purification for both SMFMO and FDH.

Methods & Materials

Fusion-cloning

The fusion cloning of SMFMO and FDH was achieved by Jared Cartwright (Technology Facility, University of York).

Expression using chaperones

The *E. coli* BL21 (DE3) cells expressing the Gro7 plasmid containing two chaperone genes, groES and groEL, was obtained from Claudia Spandolf with the Grogan group.

Starter cultures containing the chaperone genes were set up as described in Section 3.2.2 with the addition of chloramphenicol (30 μg mL^{-1}). Fresh LB media was inoculated with starter culture (1:100) and incubated at 37°C for 1 h 45 mins. After incubation the culture was centrifuged (5 min, 4000 rpm) and the pellet was resuspended in ice cold CaCl$_2$ (50 mM, half the culture volume) and left on ice for 30 min. The mixture was centrifuged (10 min, 1500 rpm) and the pellet was resuspended in ice cold CaCl$_2$ (50 mM, 1/50th original culture volume). The cells (200 μL) were then added to a pre-cooled Eppendorf tube. To the cells the vector containing the FDH-SMFMO and SMFMO-FDH fusions (2 μL) was added and left on ice for 30 mins. The mixture was heat shocked for 90 s and returned to ice for 2 min. SOC media (800 μL) was then added to the mixture and incubated at 37°C for 1 h. After incubation 100 μL of the mixture was plated onto a LB agar plate containing kanamycin and chloramphenicol (30 μg mL^{-1}).

The remaining mixture was centrifuged (5 min, 4000 rpm) and the pellet was resuspended in 150 µL of supernatant and plated onto a second LB agar plate. The plates were incubated at 37°C overnight.

Expression tests were carried out as described in Section 3.2.3 with the addition of chloramphenicol (30 µg mL^{-1}) and arabinose (125 µL).

Large scale expression was carried out in accordance with Section 3.2.5 with the addition of chloramphenicol (30 µg mL^{-1}) and arabinose (1:40).

Purification of the FDH-SMFMO and SMFMO-FDH fusion proteins was carried out as described in Section 3.3.

Results

The SMFMO-FDH fusion was cloned into the pET-YSBLIC-3C vector. The SMFMO and FDH were fusion-cloned in both orientations (FDH-SMFMO and SMFMO-FDH) and a short thirteen amino acid linker region was included (SerLeuSerThrProProThrProSerThrProProThr) to help stabilise the final constructs (Figure A2.1).

FDH-SMFMO (2 + 3)

```
MAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGGQTLPTPKAIDFTPGQLLGSVSGELGL
RKYLESNGHTLVVTSDKDGPDSVFERELVDADVVISQPFWPAYLTPERIAKAKNLKLALT
AGIGSDHVDLQSAIDRNVTVAEVTYCNSISVAEHVVMMILSLVRNYLPSHEWARKGGWNI
ADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNLTWHA
TREDMYPVCDVVTLNCPLHPETEHMINDETLKLFKRGAYIVNTARGKLCDRDAVARALES
GRLAGYAGDVWFPQPAPKDHPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFFEGR
PIRDEYLIVQGGALAGTGAHSYSKGNATGGSEEAAKFKKAVSLSTPPTPSTPPTMDSVDV
VVIGGGQSGLSAGYFLRRSGLSYVILDAEASPGGAWQHAWHSLHLFSPAGWSSIPGWPMP
ASQGPYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERLRVVARDGRQWLARAVIS
ATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAIIGGGNSGAQILAEVSTVA
ETTWITQHEPAFLADDVDGRVLFERATERWKAQQEGREPDLPPGGFGDIVMVPPVLDARA
RGVLAAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSHLKGLDLVTPQGQVEVDG
SGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTAYCADHQDR
```

SMFMO-FDH (1+4)

```
MDSVDVVVIGGGQSGLSAGYFLRRSGLSYVILDAEASPGGAWQHAWHSLHLFSPAGWSSI
PGWPMPASQGPYPARAEVLAYLAQYEQKYALPVLRPIRVQRVSHFGERLRVVARDGRQWL
ARAVISATGTWGEAYTPEYQGLESFAGIQLHSAHYSTPAPFAGMRVAIIGGGNSGAQILA
EVSTVAETTWITQHEPAFLADDVDGRVLFERATERWKAQQEGREPDLPPGGFGDIVMVPP
VLDARARGVLAAVPPPARFSPTGMQWADGTERAFDAVIWCTGFRPALSHLKGLDLVTPQG
QVEVDGSGLRALAVPSVWLLGYGDWNGMASATLIGVTRYAREAVRQVTAYCADHQDRSLS
TPPTPSTPPTMAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGGQTLPTPKAIDFTPGQL
LGSVSGELGLRKYLESNGHTLVVTSDKDGPDSVFERELVDADVVISQPFWPAYLTPERIA
KAKNLKLALTAGIGSDHVDLQSAIDRNVTVAEVTYCNSISVAEHVVMMILSLVRNYLPSH
EWARKGGWNIADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESV
EKELNLTWHATREDMYPVCDVVTLNCPLHPETEHMINDETLKLFKRGAYIVNTARGKLCD
RDAVARALESGRLAGYAGDVWFPQPAPKDHPWRTMPYNGMTPHISGTTLTAQARYAAGTR
EILECFFEGRPIRDEYLIVQGGALAGTGAHSYSKGNATGGSEEAAKFKKAV
```

Figure A2.1 Protein sequences of FDH-SMFMO and SMFMO-FDH fusion constructs. SMFMO is shown in pink, FDH is shown in blue and the amino acid linker is shown in black.

Expression of both FDH-SMFMO and SMFMO-FDH genes was attempted in *E. coli* Bl21 (DE3) cells which expressed a protein approximately 82 kDa (FDH plus SMFMO) however the expression was insoluble. *E. coli* Rossetta (DE3) 2 pLys cells were also employed to express the fusion-cloned genes however, expression was unsuccessful as no band at 82 kDa was present on the SDS-PAGE gel.

The unsuccessful expression systems led to the use of chaperone proteins groES and groEL. The chaperone proteins were used to assist the expression of the FDH-SMFMO and SMFMO-FDH fusion proteins. Expression tests indicated that both fusion proteins were expressed as bands at 82 kDa were present. The expression was scaled to 1 L LB media and the protein was purified. At the purification stage the SMFMO-FDH fusion

protein proved unsuccessful as no peak was present on the chromatogram after purification by Ni^{2+} affinity chromatography. The FDH-SMFMO fusion protein purification proved more successful as after Ni^{2+} affinity and size exclusion chromatography pure pale yellow protein was obtained at a concentration of 2.3 mg.

The expression of the FDH-SMFMO fusion protein was scaled to 8 L LB media and the protein was purified (Figure A2.2).
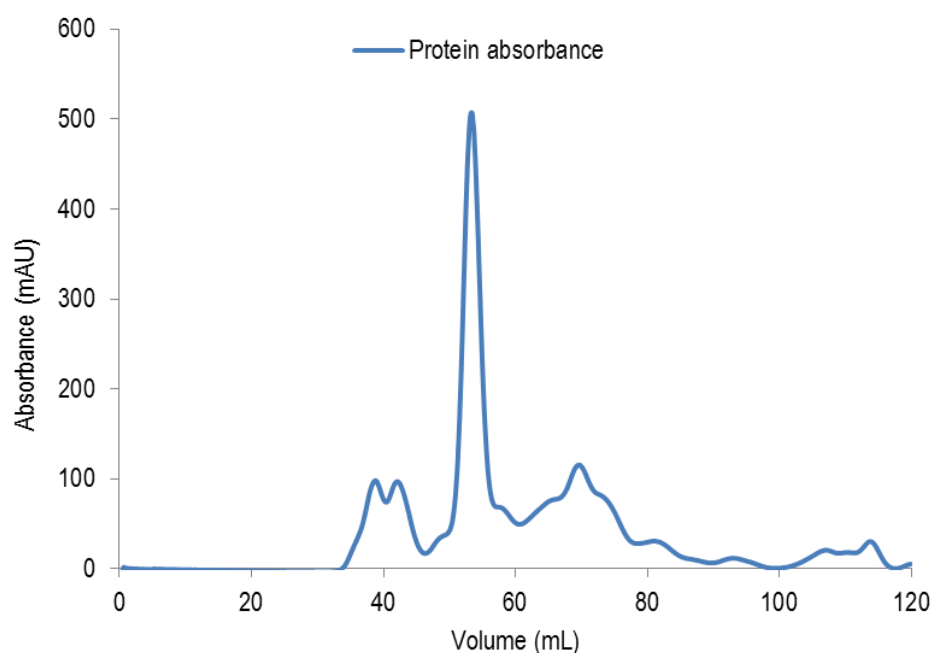
Figure A2.2 Chromatogram for FDH-SMFMO purification by size exclusion chromatography.
Protein absorbance at 280 nm is indicated by the blue trace.

The protein eluted at approximately 55 mL. Using the calibration curve provided by GE Healthcare, the position of the elution peak gave the molecular weight at ~109 kDa, indicating that the 82 kDa fusion protein is still linked to part of the chaperone protein. In addition, other peaks present on the chromatograph indicate aggregates are present. The fractions corresponding to the peak were pale yellow in colour, indicating that FAD was still present within the fusion protein. The fractions were analysed by SDS-PAGE (Section 3.2.4). The gel can be seen in Figure A2.3.
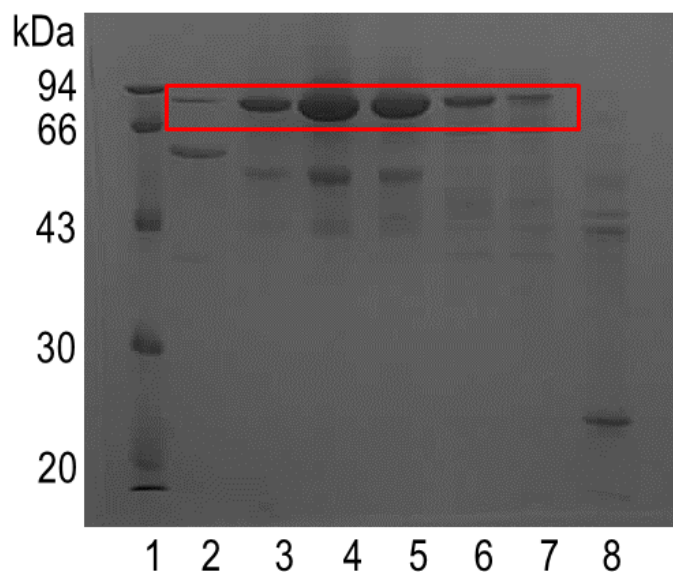
Figure A2.3 SDS-PAGE gel of FDH-SMFMO fusion purification by size exclusion chromatography.
Lane 1: Bio-Rad low weight molecular marker, lane 2-8: collected fractions after FPLC run. The protein can be seen at ~82 kDa.

The FDH-SMFMO protein can be seen as a dark band on the gel at ~82 kDa. The gel shows other impurities at lower molecular weights suggesting that the protein is impure. The fractions were combined to give ~4.3 mg of FDH-SMFMO fusion protein (Section 3.3.3).

In an attempt to investigate if the fusion protein was able to catalyse the oxygenation of sulfides in addition to cofactor recycling biotransformation reactions were set up with substrate **15** as described in Section 3.4.4. The SMFMO and FDH were replaced by the FDH-SMFMO fusion protein at the same concentration. Unfortunately, after 1 h an egg white like precipitate occurred indicating that the protein was unstable and had become denatured. GC analysis of the time samples also suggested that the enzyme was unstable as the substrate conversion after 24 h was only 14 %, whereas for wild type SMFMO the conversion was 90 %.

Future work in this area could include, repeating the biotransformation with cell extract to investigate if any oxidation of the sulfide occurs. In addition to the fusion proteins, the co-expression of SMFMO and FDH could be attempted.

List of abbreviations

^{13}C	Carbon NMR
^{1}H	Proton NMR
A	Measured absorbance
Å	Angstrom, $1\text{x}10^{-10}$ m
APS	Ammonium persulfate
B factor	Displacement parameter
btp	Bis-tris propane
BV	Baeyer-Villiger
BVMO	Baeyer-Villiger monooxygenase
C	Concentration
CC$_{1/2}$	Pearson correlation coefficient
CFMO	*Cellvibrio* sp. flavin-containing monooxygenase
CHMO	Cyclohexanone monooxygenase
Cα	Alpha carbon
dATP	Deoxyadenosine triphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphate monomeric units
DTT	Dithiothreitol

dTTP	Thymidine triphosphate
E	Enzyme
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
EP	Enzyme-product complex
ES	Enzyme-substrate complex
et al.	*Et alia* (and others)
EtOAc	Ethyl acetate
EtOH	Ethanol
FAD	Flavin adenine dinucleotide
F_C	Calculated structure factor
FDH	Formate dehydrogenase
FMN	Flavin mononucleotide
FMO	Flavin-containing monooxygenase
F_O	Observed structure factor
FPLC	Fast protein liquid chromatography
FPMO	Flavoprotein monooxygenase
g	Gram
G.O.I	Gene of interest
GC	Gas chromatography

GC-MS	Gas chromatography-mass spectrometry
h	Hour
HRV 3C	Human rhinovirus 3C
I	Transmitted intensity
I_0	Intensity of incident light
<I/σ(I)>	Measure of intensity above an average signal to noise ratio
IPTG	Isopropyl β-1-thiogalactopyranoside
kb	Kilo base
k_{cat}	Turnover number
K_{cat}/K_M	Specificity constant
kDa	Kilo Dalton
K_M	Michaelis-Menten constant (binding constant)
kPa	Kilo Pascal
l	Path length
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
Li_2SO_4	Lithium sulfate
LIC	Ligation independent cloning
m	meter
M	Molar
m-	Meta

m/z	Mass/charge
mA	Milli amps
mFMO	*Methylophaga* Flavin-containing MonoOxygenase
mg	Milli gram
MgSO$_4$	Magnesium sulfate
min	Minutes
mL	Milli litre
mM	Milli molar
MW	Molecular weight
n.d.	Not determined
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosophate
NaF	Sodium formate
NaOH	Sodium hydroxide
ng	Nano gram
nm	Nano meter
NMO	N-hydroxylating Monooxygenase
nmol	Nano moles
NMR	Nuclear magnetic resonance
o-	Ortho
^{o}C	Degrees Celsius

OD$_{600nm}$	Optical density at 600 nm
P	Product
p-	Para
PAMO	Phenylacetone monooxygenase
PCR	Polymerase chain reaction
PFMO	*Pseudomonas stutzeri* flavin-containing monooxygenase
PhOCH$_2$	Phenoxymethyl-
PML	Plymouth Marine Laboratories
*p*mol	Pico moles
r.t	Room temperature
R$_{free}$	Measure of agreement between observed and calculated data from a randomly chosen test set
R$_{merge}$	Measure of discrepancy between observed reflections
rmsd	Root mean square deviation
RNA	Ribonucleic acid
R$_{p.i.m}$	Measure of data quality after averaging of multiple measurements
rpm	Revolutions per minute
s	Seconds
S	Substrate
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate

SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide electrophoresis
SEC	Size exclusion chromatography
SMFMO	*Stenotrophomonas maltophilia* flavin-containing monooxygenase
SSM	Site saturation mutagenesis
t	Temperature
TEMED	Tetramethylethylenediamine
TrxRs	Thioredoxin reductases
UV-Vis	Ultraviolet-visible
V	Volts
V$_{max}$	Maximum velocity (maximum absorption)
w/v	Weight to volume
YSBL	York Structural Biology Lab
α	Alpha
β	Beta
γ	Gamma
ε	Extinction coefficient
µg	Micro gram
µL	Micro litre
µM	Micro moloar
µmol	Micro moles

σ   Sigma

The following substrates **1**-**20** were employed for biotransformations.

1	**2**	**3**	**4**	**5**	**6**

7	**8**	**9**	**10**	**11**	**12**

13	**14**	**15: *p*-Me**
		16: H
		17: *o*-Cl
		18: *m*-Cl
		19: *p*-Cl
		20: *p*-NO$_2$
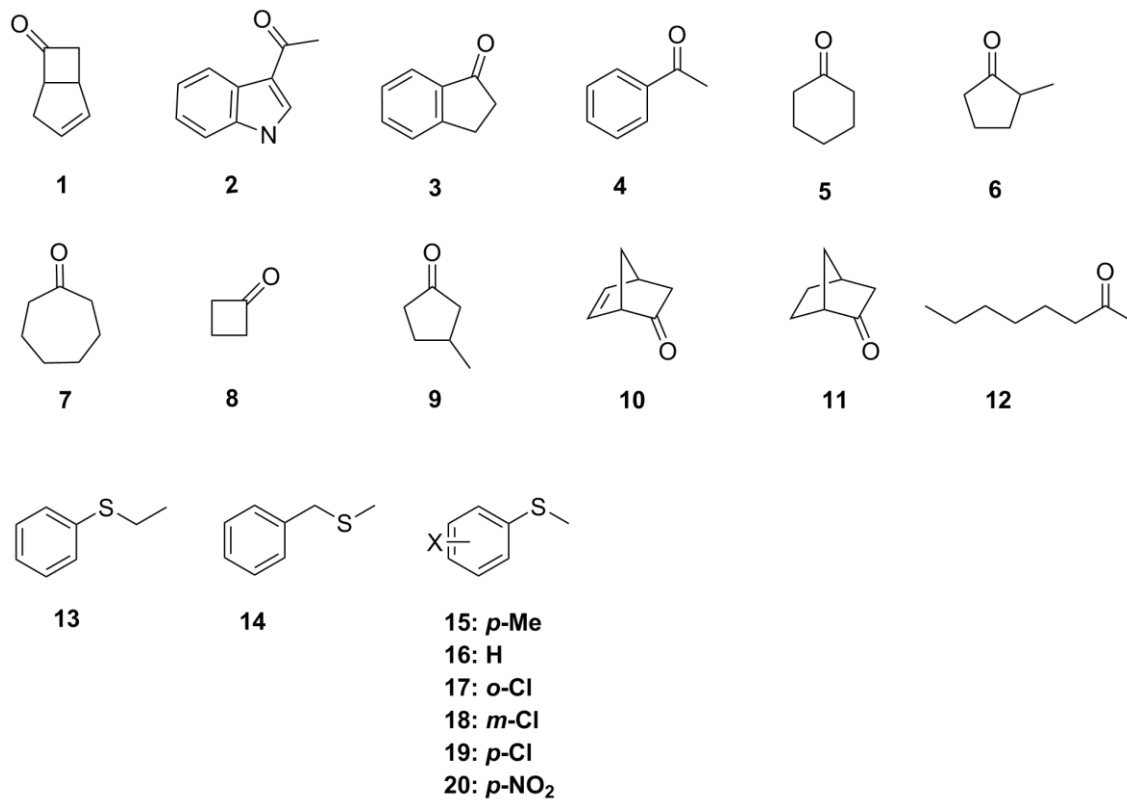
Substrate **1**   Bicyclo[3.2.0]hept-2-en-6-one

Substrate **2**   3-acetylindole

Substrate **3**   1-indanone

Substrate **4**   Acetophenone

Substrate **5**   Cyclohexanone

Substrate **6**   2-methylcyclopentanone

Substrate **7**   Cycloheptanone

Substrate **8** Cyclobutanone

Substrate **9** 3-methylcyclopentanone

Substrate **10** 5-norbornen-2-one

Substrate **11** Norcamphor

Substrate **12** Octan-2-one

Substrate **13** Ethyl-phenyl sulfide

Substrate **14** Benzyl-methyl sulfide

Substrate **15** Methyl-*p*-tolyl sulfide

Substrate **16** Thioanisole

Substrate **17** *ortho*-chloro thioanisole

Substrate **18** *meta*-chloro thioanisole

Substrate **19** *para*-chloro thioanisole

Substrate **20** *para*-nitro thioanisole

References

1.   (a) Bugg, T. D. H., Introduction to Enzyme and Coenzyme Chemistry. *Wiley, Chichester, U.K.* **2012**; (b) Reetz, M. T., Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **2013**, *135*, 12480-12496.

2.   (a) Gal, J., The Discovery of Biological Enantioselectivity: Louis Pasteur and the Fermentation of Tartaric Acid, 1857—A Review and Analysis 150 yr Later. *Chirality* **2008**, *20*, 5-19; (b) Pasteur, L., Mémoire sur la fermentation de l'acide tartrique. *C. R. Acad. Sci.* **1858**, *46*, 615–618.

3.   Fischer, E., Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985-2993.

4.   Buchner, E., Alkoholische Gährung ohne Hefezellen. *Ber. Dtsch. Chem. Ges.* **1897**, *30*, 117-124.

5.   Anastas, P. T. W., J. C. , Green Chemistry: Theory and Practice. *Oxford University Press: New York* **1998**, 30.

6.   Grogan, G., Practical Biotransformations: A Beginner's Guide. *Wiley, Chichester, U.K.* **2009**.

7.   Hogg, J. A., Steroids, the Steroid Community, and Upjohn in Perspective: A Profile of Innovation. *Steroids* **1992**, *57*, 593-616.

8.   Guengerich, F. P.; Munro, A. W., Unusual Cytochrome P450 Enzymes and Reactions. *J. Biol. Chem.* **2013**, *288*, 17065-17073.

9.   Woodward, R. B.; Sondheimer, F.; Taub, D.; Heusler, K.; McLamore, W. M., The Total Synthesis of Steroids1. *J. Am. Chem. Soc.* **1952**, *74*, 4223-4251.

10.   (a) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Keßeler, M.; Stürmer, R.; Zelinski, T., Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 788-824; (b) Nestl, B. M.; Hammer, S. C.; Nebel, B. A.; Hauer, B., New Generation of Biocatalysts for Organic Synthesis. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3070-3095.

11.   van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W., Flavoprotein Monooxygenases, A Diverse Class of Oxidative Biocatalysts. *J. Biotechnol.* **2006**, *124*, 670-89.

12.   (a) Doering, W. v. E.; Dorfman, E., Mechanism of the Peracid Ketone-Ester Conversion. Analysis of Organic Compounds for Oxygen-18. *J. Am. Chem. Soc.* **1953**, *75*, 5595-5598; (b) Doering, W. v. E.; Speers, L., The Peracetic Acid Cleavage of Unsymmetrical Ketones. *J. Am. Chem. Soc.* **1950**, *72*, 5515-5518.

13.   Hall, M.; Stueckler, C.; Kroutil, W.; Macheroux, P.; Faber, K., Asymmetric Bioreduction of Activated Alkenes Using Cloned 12-Oxophytodienoate Reductase

Isoenzymes OPR-1 and OPR-3 from *Lycopersicon esculentum* (Tomato): A Striking Change of Stereoselectivity. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 3934-3937.

14.   Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; Grosdidier, A.; Hernandez, C.; Ioannidis, V.; Kuznetsov, D.; Liechti, R.; Moretti, S.; Mostaguir, K.; Redaschi, N.; Rossier, G.; Xenarios, I.; Stockinger, H., ExPASy: SIB Bioinformatics Resource Portal. *Nucleic Acids Res.* **2012**, *40*, 597-603.

15.   Coon, M. J., Cytochrome P450: Nature's Most Versatile Biological Catalyst. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 1-25.

16.   Sciara, G.; Kendrew, S. G.; Miele, A. E.; Marsh, N. G.; Federici, L.; Malatesta, F.; Schimperna, G.; Savino, C.; Vallone, B., The Structure of ActVA-Orf6, A Novel Type of Monooxygenase Involved in Actinorhodin Biosynthesis. *EMBO J.* **2003**, *22*, 205-215.

17.   Adams, M. A.; Jia, Z., Structural and Biochemical Evidence for an Enzymatic Quinone Redox Cycle in *Escherichia coli*: Identification of a Novel Quinol Monooxygenase. *J. Biol. Chem.* **2005**, *280*, 8358-8363.

18.   Riebel, A.; de Gonzalo, G.; Fraaije, M. W., Expanding the Biocatalytic Toolbox of Flavoprotein Monooxygenases from *Rhodococcus jostii* RHA1. *J. Mol. Catal. B: Enzym.* **2013**, *88*, 20-25.

19.   Cashman, J. R.; Zhang, J., Human Flavin-Containing Monooxygenases. *Annu. Rev. Pharmacol. Toxicol.* **2006**, *46*, 65-100.

20.   (a) Beam, M. P.; Bosserman, M. A.; Noinaj, N.; Wehenkel, M.; Rohr, J., Crystal Structure of Baeyer−Villiger Monooxygenase MtmOIV, the Key Enzyme of the Mithramycin Biosynthetic Pathway. *Biochemistry* **2009**, *48*, 4476-4487; (b) Seo, M.-J.; Zhu, D.; Endo, S.; Ikeda, H.; Cane, D. E., Genome Mining in *Streptomyces*. Elucidation of the Role of Baeyer−Villiger Monooxygenases and Non-Heme Iron-Dependent Dehydrogenase/Oxygenases in the Final Steps of the Biosynthesis of Pentalenolactone and Neopentalenolactone. *Biochemistry* **2011**, *50*, 1739-1754.

21.   Colonna, S.; Gaggero, N.; Carrea, G.; Ottolina, G.; Pasta, P.; Zambianchi, F., First Asymmetric Epoxidation Catalysed by Cyclohexanone Monooxygenase. *Tetrahedron Lett.* **2002**, *43*, 1797-1799.

22.   (a) de Gonzalo, G.; Mihovilovic, M. D.; Fraaije, M. W., Recent Developments in the Application of Baeyer–Villiger Monooxygenases as Biocatalysts. *ChemBioChem* **2010**, *11*, 2208-2231; (b) Mihovilovic, M. D.; Müller, B.; Stanetty, P., Monooxygenase-Mediated Baeyer-Villiger Oxidations. *Eur. J. Org. Chem.* **2002**, 3711-3730; (c) Torres Pazmiño, D. E.; Dudek, H. M.; Fraaije, M. W., Baeyer–Villiger Monooxygenases: Recent Advances and Future Challenges. *Curr. Opin. Chem. Biol.* **2010**, *14*, 138-144.

23.   (a) Colonna, S.; Gaggero, N.; Pasta, P.; Ottolina, G., Enantioselective Oxidation of Sulfides to Sulfoxides Catalysed by Bacterial Cyclohexanone Monooxygenases. *Chem. Commun.* **1996**, 2303-2307; (b) Light, D. R.; Waxman, D. J.; Walsh, C., Studies

on the Chirality of Sulfoxidation Catalyzed by Bacterial Flavoenzyme Cyclohexanone Monooxygenase and Hog Liver FAD-Containing Monooxygenase. *Biochemistry* **1982**, *21*, 2490-2498; (c) Pasta, P.; Carrea, G.; Holland, H. L.; Dallavalle, S., Synthesis of Chiral Benzyl Alkyl Sulfoxides by Cyclohexanone Monooxygenase from *Acinetobacter* NCIB 9871. *Tetrahedron: Asymmetry* **1995**, *6*, 933-936.

24.   Ogilby, P. R., Singlet Oxygen: There is Indeed Something New Under The Sun. *Chem. Soc. Rev.* **2010**, *39*, 3181-3209.

25.   Massey, V., Activation of Molecular Oxygen by Flavins And Flavoproteins. *J. Biol. Chem.* **1994**, *269*, 22459-22462.

26.   Ghisla, S.; Massey, V., Mechanisms of Flavoprotein-Catalyzed Reactions. *Eur. J. Biochem.* **1989**, *181*, 1-17.

27.   Entsch, B.; van Berkel, W. J., Structure and Mechanism of *para*-Hydroxybenzoate Hydroxylase. *FASEB J.* **1995**, *9*, 476-83.

28.   Sutton, W. B., Mechanism of Action and Crystallization af Lactic Oxidative Decarboxylase from *Mycobacterium Phlei*. *J. Biol. Chem.* **1957**, *226*, 395-405.

29.   De Jong, E.; Van Berkel, W. J. H.; Van Der Zwan, R. P.; De Bont, J. A. M., Purification and Characterization of Vanillyl-Alcohol Oxidase from *Penicillium simplicissimum*. *Eur. J. Biochem.* **1992**, *208*, 651-657.

30.   http://www.sigmaaldrich.com/united-kingdom.html.

31.   Massey, V., The Chemical and Biological Versatility of Riboflavin. *Biochem. Soc. Trans.* **2000**, *28*, 283–296.

32.   Moonen, M. J. H.; Fraaije, M. W.; Rietjens, I. M. C. M.; Laane, C.; van Berkel, W. J. H., Flavoenzyme-Catalyzed Oxygenations and Oxidations of Phenolic Compounds. *Adv. Synth. Catal.* **2002**, *344*, 1023-1035.

33.   White-Stevens, R. H.; Kamin, H., Studies of a Flavoprotein, Salicylate Hydroxylase: I. Preparation, Properties, and the Uncoupling of Oxygen Reduction from Hydroxylation. *J. Biol. Chem.* **1972**, *247*, 2358-2370.

34.   Entsch, B.; Cole, L. J.; Ballou, D. P., Protein Dynamics and Electrostatics in the Function of *p*-Hydroxybenzoate Hydroxylase. *Arch. Biochem. Biophys.* **2005**, *433*, 297-311.

35.   (a) Gibson, M.; Nur-e-alam, M.; Lipata, F.; Oliveira, M. A.; Rohr, J., Characterization of Kinetics and Products of the Baeyer−Villiger Oxygenase MtmOIV, The Key Enzyme of the Biosynthetic Pathway toward the Natural Product Anticancer Drug Mithramycin from *Streptomyces argillaceus*. *J. Am. Chem. Soc.* **2005**, *127*, 17594-17595; (b) Prado, L.; Fernández, E.; Weiβbach, U.; Blanco, G.; Quirós, L. M.; Braña, A. F.; Méndez, C.; Rohr, J.; Salas, J. A., Oxidative Cleavage of Premithramycin B is one of the Last Steps in the Biosynthesis of the Antitumor Drug Mithramycin. *Chem. Biol.* **1999**, *6*, 19-30.

36.   Fraaije, M. W.; Kamerbeek, N. M.; van Berkel, W. J.; Janssen, D. B., Identification of a Baeyer-Villiger Monooxygenase Sequence Motif. *FEBS Lett.* **2002**, *518*, 43-47.

37.   (a) Baldwin, T. O.; Christopher, J. A.; Raushel, F. M.; Sinclair, J. F.; Ziegler, M. M.; Fisher, A. J.; Rayment, I., Structure of Bacterial Luciferase. *Curr. Opin. Struct. Biol.* **1995**, *5*, 798-809; (b) Fisher, A. J.; Raushel, F. M.; Baldwin, T. O.; Rayment, I., Three-Dimensional Structure of Bacterial Luciferase from *Vibrio harveyi* at 2.4 Å Resolution. *Biochemistry* **1995**, *34*, 6581-6586.

38.   Taylor, D. G.; Trudgill, P. W., Camphor Revisited: Studies of 2,5-Diketocamphane 1,2-Monooxygenase from *Pseudomonas putida* ATCC 17453. *J. Bacteriol.* **1986**, *165*, 489-497.

39.   (a) Duffner, F. M.; Kirchner, U.; Bauer, M. P.; Müller, R., Phenol/Cresol Degradation by the Thermophilic *Bacillus thermoglucosidasius* A7: Cloning and Sequence Analysis of Five Genes Involved in the Pathway. *Gene* **2000**, *256*, 215-221; (b) Kirchner, U.; Westphal, A. H.; Müller, R.; van Berkel, W. J. H., Phenol Hydroxylase from *Bacillus thermoglucosidasius* A7, a Two-protein Component Monooxygenase with a Dual Role for FAD. *J. Biol. Chem.* **2003**, *278*, 47545-47553; (c) van den Heuvel, R. H. H.; Fraaije, M. W.; van Berkel, W. J. H., Direction of the Reactivity of Vanillyl-Alcohol Oxidase with 4-Alkylphenols. *FEBS Lett.* **2000**, *481*, 109-112.

40.   Kadiyala, V.; Spain, J. C., A Two-Component Monooxygenase Catalyzes Both the Hydroxylation of *p*-Nitrophenol and the Oxidative Release of Nitrite from 4-Nitrocatechol in *Bacillus sphaericus* JS905. *Appl. Environ. Microbiol.* **1998**, *64*, 2479-2484.

41.   (a) Galán, B.; Díaz, E.; Prieto, M. A.; García, J. L., Functional Analysis of the Small Component of the 4-Hydroxyphenylacetate 3-Monooxygenase of *Escherichia coli* W: a Prototype of a New Flavin:NAD(P)H Reductase Subfamily. *J. Bacteriol.* **2000**, *182*, 627-636; (b) Prieto, M. A.; Garcia, J. L., Molecular Characterization of 4-Hydroxyphenylacetate 3-Hydroxylase of *Escherichia coli*. A Two-Protein Component Enzyme. *J. Biol. Chem.* **1994**, *269*, 22823-22829.

42.   (a) Chaiyen, P.; Suadee, C.; Wilairat, P., A Novel Two-Protein Component Flavoprotein Hydroxylase. *Eur. J. Biochem.* **2001**, *268*, 5550-5561; (b) Sucharitakul, J.; Chaiyen, P.; Entsch, B.; Ballou, D. P., The Reductase of *p*-Hydroxyphenylacetate 3-Hydroxylase from *Acinetobacter baumannii* Requires *p*-Hydroxyphenylacetate for Effective Catalysis. *Biochemistry* **2005**, *44*, 10434-10442; (c) Thotsaporn, K.; Sucharitakul, J.; Wongratana, J.; Suadee, C.; Chaiyen, P., Cloning and Expression of *p*-Hydroxyphenylacetate 3-Hydroxylase from *Acinetobacter baumannii*: Evidence of the Divergence of Enzymes in the Class of Two-Protein Component Aromatic Hydroxylases. *Biochim. Biophys. Acta - Gene Structure and Expression* **2004**, *1680*, 60-66.

43.   Panke, S.; Witholt, B.; Schmid, A.; Wubbolts, M. G., Towards a Biocatalyst for (*S*)-Styrene Oxide Production: Characterization of the Styrene Degradation Pathway of *Pseudomonas* sp. Strain VLB120. *Appl. Environ. Microbiol.* **1998**, *64*, 2032-2043.

44.   Panke, S.; Wubbolts, M. G.; Schmid, A.; Witholt, B., Production of Enantiopure Styrene Oxide by Recombinant *Escherichia coli* Synthesizing a Two-Component Styrene Monooxygenase. *Biotechnol. Bioeng.* **2000**, *69*, 91-100.

45.   Schmid, A.; Hofstetter, K.; Feiten, H.-J.; Hollmann, F.; Witholt, B., Integrated Biocatalytic Synthesis on Gram Scale: The Highly Enantioselective Preparation of Chiral Oxiranes with Styrene Monooxygenase. *Adv. Synth. Catal.* **2001**, *343*, 732-737.

46.   Panke, S.; Held, M.; Wubbolts, M. G.; Witholt, B.; Schmid, A., Pilot-Scale Production of (*S*)-Styrene Oxide from Styrene by Recombinant *Escherichia coli* Synthesizing Styrene Monooxygenase. *Biotechnol. Bioeng.* **2002**, *80*, 33-41.

47.   Ukaegbu, U. E.; Kantz, A.; Beaton, M.; Gassner, G. T.; Rosenzweig, A. C., Structure and Ligand Binding Properties of the Epoxidase Component of Styrene Monooxygenase. *Biochemistry* **2010**, *49*, 1678-1688.

48.   (a) Keller, S.; Wage, T.; Hohaus, K.; Hölzer, M.; Eichhorn, E.; van Pée, K.-H., Purification and Partial Characterization of Tryptophan 7-Halogenase (PrnA) from *Pseudomonas fluorescens*. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 2300-2302; (b) Yeh, E.; Garneau, S.; Walsh, C. T., Robust In Vitro activity of RebF and RebH, a Two-Component Reductase/Halogenase, Generating 7-Chlorotryptophan during Rebeccamycin Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3960-3965.

49.   Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K.-H.; Naismith, J. H., Tryptophan 7-Halogenase (PrnA) Structure Suggests a Mechanism for Regioselective Chlorination. *Science* **2005**, *309*, 2216-2219.

50.   (a) Baeyer, A.; Villiger, V., Einwirkung des Caro'schen Reagens auf Ketone. *Ber. Dtsch. Chem. Ges.* **1899**, *32*, 3625-3633; (b) ten Brink, G. J.; Arends, I. W. C. E.; Sheldon, R. A., The Baeyer−Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev.* **2004**, *104*, 4105-4124.

51.   (a) Renz, M.; Meunier, B., 100 Years of Baeyer−Villiger Oxidations. *Eur. J. Org. Chem.* **1999**, 737-750; (b) Strukul, G., Transition Metal Catalysis in the Baeyer−Villiger Oxidation of Ketones. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 1198-1209.

52.   Criegee, R., Die Umlagerung der Dekalin-peroxydester als Folge von kationischem Sauerstoff. *Justus Liebigs Ann. Chem.* **1948**, *560*, 127-135.

53.   Sheng, D.; Ballou, D. P.; Massey, V., Mechanistic Studies of Cyclohexanone Monooxygenase: Chemical Properties of Intermediates Involved in Catalysis. *Biochemistry* **2001**, *40*, 11156-11167.

54.   Ryerson, C. C.; Ballou, D. P.; Walsh, C., Mechanistic Studies on Cyclohexanone Oxygenase. *Biochemistry* **1982**, *21*, 2644-2655.

55.   Kamerbeek, N. M.; Janssen, D. B.; van Berkel, W. J. H.; Fraaije, M. W., Baeyer−Villiger Monooxygenases, an Emerging Family of Flavin-Dependent Biocatalysts. *Adv. Synth. Catal.* **2003**, *345*, 667-678.

56.   Donoghue, N. A.; Norris, D. B.; Trudgill, P. W., The Purification and Properties of Cyclohexanone Oxygenase from *Nocardia globerula* CL1 and *Acinetobacter* NCIB 9871. *Eur. J. Biochem.* **1976**, *63*, 175-192.

57.   Mihovilovic, M. D.; Rudroff, F.; Grötzl, B.; Kapitan, P.; Snajdrova, R.; Rydz, J.; Mach, R., Family Clustering of Baeyer–Villiger Monooxygenases Based on Protein Sequence and Stereopreference. *Angew. Chem. Int. Ed. Engl.* **2005,** *44*, 3609-3613.

58.   (a) Reetz, M. T.; Brunner, B.; Schneider, T.; Schulz, F.; Clouthier, C. M.; Kayser, M. M., Directed Evolution as a Method To Create Enantioselective Cyclohexanone Monooxygenases for Catalysis in Baeyer–Villiger Reactions. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 4075-4078; (b) Reetz, M. T.; Daligault, F.; Brunner, B.; Hinrichs, H.; Deege, A., Directed Evolution of Cyclohexanone Monooxygenases: Enantioselective Biocatalysts for the Oxidation of Prochiral Thioethers. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 4078-4081.

59.   Mirza, I. A.; Yachnin, B. J.; Wang, S.; Grosse, S.; Bergeron, H.; Imura, A.; Iwaki, H.; Hasegawa, Y.; Lau, P. C.; Berghuis, A. M., Crystal Structures of Cyclohexanone Monooxygenase Reveal Complex Domain Movements and a Sliding Cofactor. *J. Am. Chem. Soc.* **2009**, *131*, 8848-54.

60.   Higson, F. K.; Focht, D. D., Bacterial Degradation of Ring-Chlorinated Acetophenones. *Appl. Environ. Microbiol.* **1990**, *56*, 3678-3685.

61.   (a) Malito, E.; Alfieri, A.; Fraaije, M. W.; Mattevi, A., Crystal Structure of a Baeyer-Villiger Monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13157-13162; (b) Fraaije, M.; Wu, J.; Heuts, D. H. M.; Hellemond, E.; Spelberg, J. L.; Janssen, D., Discovery of a Thermostable Baeyer–Villiger Monooxygenase by Genome Mining. *App. Microbiol. Biotechnol.* **2005**, *66*, 393-400.

62.   de Gonzalo, G.; Ottolina, G.; Zambianchi, F.; Fraaije, M. W.; Carrea, G., Biocatalytic Properties of Baeyer–Villiger Mmonooxygenases in Aqueous–Organic Media. *J. Mol. Catal. B: Enzym.* **2006**, *39*, 91-97.

63.   Gonzalo, G. d.; Pazmiño, D. E. T.; Ottolina, G.; Fraaije, M. W.; Carrea, G., Oxidations Catalyzed by Phenylacetone Monooxygenase from *Thermobifida fusca*. *Tetrahedron: Asymmetry* **2005**, *16*, 3077-3083.

64.   Krueger, S. K.; Williams, D. E., Mammalian Flavin-Containing Monooxygenases: Structure/Function, Genetic Polymorphisms and Role in Drug Metabolism. *Pharmacol. Ther.* **2005**, *106*, 357-387.

65.   Zhou, J.; Shephard, E. A., Mutation, Polymorphism and Perspectives for the Future of Human Flavin-Containing Monooxygenase 3. *Mutat. Res.- Rev. Mutat.* **2006**, *612*, 165-171.

66.   Cashman, J. R., Some Distinctions Between Flavin-Containing and Cytochrome P450 Monooxygenases. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 599-604.

67.   Alfieri, A.; Malito, E.; Orru, R.; Fraaije, M. W.; Mattevi, A., Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105*, 6572-6577.

68.   Suh, J.-K.; Poulsen, L. L.; Ziegler, D. M.; Robertus, J. D., Molecular Cloning and Kinetic Characterization of a Flavin-Containing Monooxygenase from *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.* **1996**, *336*, 268-274.

69.   Suh, J.-K.; Poulsen, L. L.; Ziegler, D. M.; Robertus, J. D., Yeast Flavin-Containing Monooxygenase Generates Oxidizing Equivalents that Control Protein Folding in the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2687-2691.

70.   Eswaramoorthy, S.; Bonanno, J. B.; Burley, S. K.; Swaminathan, S., Mechanism of Action of a Flavin-Containing Monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9832-9837.

71.   Zhao, Y.; Christensen, S. K.; Fankhauser, C.; Cashman, J. R.; Cohen, J. D.; Weigel, D.; Chory, J., A Role for Flavin Monooxygenase-Like Enzymes in Auxin Biosynthesis. *Science* **2001**, *291*, 306-309.

72.   Choi, H. S.; Kim, J. K.; Cho, E. H.; Kim, Y. C.; Kim, J. I.; Kim, S. W., A Novel Flavin-Containing Monooxygenase from *Methylophaga* sp. Strain SK1 and its Indigo Synthesis in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **2003**, *306*, 930-936.

73.   Rioz-Martinez, A.; Kopacz, M.; de Gonzalo, G.; Torres Pazmino, D. E.; Gotor, V.; Fraaije, M. W., Exploring the Biocatalytic Scope of a Bacterial Flavin-Containing Monooxygenase. *Org. Biomol. Chem.* **2011**, *9*, 1337-1341.

74.   Cho, H. J.; Cho, H. Y.; Kim, K. J.; Kim, M. H.; Kim, S. W.; Kang, B. S., Structural and Functional Analysis of Bacterial Flavin-Containing Monooxygenase Reveals its Ping-Pong-Type Reaction Mechanism. *J. Struct. Biol.* **2011**, *175*, 39-48.

75.   Beaty, N. B.; Ballou, D. P., The Reductive Half-Reaction of Liver Microsomal FAD-Containing Monooxygenase. *J. Biol. Chem.* **1981**, *256*, 4611-4618.

76.   Cashman, J. R., Structural and Catalytic Properties of the Mammalian Flavin-Containing Monooxygenase. *Chem. Res. Toxicol.* **1995**, *8*, 165-181.

77.   Beaty, N. B.; Ballou, D. P., Transient Kinetic Study of Liver Microsomal FAD-Containing Monooxygenase. *J. Biol. Chem.* **1980**, *255*, 3817-3819.

78.   Beaty, N. B.; Ballou, D. P., The Oxidative Half-Reaction of Liver Microsomal FAD-Containing Monooxygenase. *J. Biol. Chem.* **1981**, *256*, 4619-4625.

79.   (a) Ge, L.; Seah, S. Y. K., Heterologous Expression, Purification, and Characterization of an l-Ornithine N5-Hydroxylase Involved in Pyoverdine Siderophore Biosynthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **2006**, *188*, 7205-7210; (b) Meneely, K. M.; Barr, E. W.; Bollinger, J. M.; Lamb, A. L., Kinetic Mechanism of Ornithine Hydroxylase (PvdA) from *Pseudomonas aeruginosa*: Substrate Triggering of

O$_2$ Addition but Not Flavin Reduction. *Biochemistry* **2009**, *48*, 4371-4376; (c) Meneely, K. M.; Lamb, A. L., Biochemical Characterization of a Flavin Adenine Dinucleotide-Dependent Monooxygenase, Ornithine Hydroxylase from *Pseudomonas aeruginosa*, Suggests a Novel Reaction Mechanism. *Biochemistry* **2007**, *46*, 11930-11937.

80.   (a) Chocklett, S. W.; Sobrado, P., Aspergillus fumigatus SidA Is a Highly Specific Ornithine Hydroxylase with Bound Flavin Cofactor. *Biochemistry* **2010**, *49*, 6777-6783; (b) Mayfield, J. A.; Frederick, R. E.; Streit, B. R.; Wencewicz, T. A.; Ballou, D. P.; DuBois, J. L., Comprehensive Spectroscopic, Steady State, and Transient Kinetic Studies of a Representative Siderophore-associated Flavin Monooxygenase. *J. Biol. Chem.* **2010**, *285*, 30375-30388.

81.   (a) Thariath, A.; Socha, D.; Valvano, M. A.; Viswanatha, T., Construction and Biochemical Characterization of Recombinant Cytoplasmic Forms of the IucD Protein (Lysine:N6-hydroxylase) Encoded by the pColV-K30 Aerobactin Gene Cluster. *J. Bacteriol.* **1993**, *175*, 589-596; (b) Thariath, A. M.; Fatum, K. L.; Valvano, M. A.; Viswanatha, T., Physico-chemical Characterization of a Recombinant Cytoplasmic Form of Lysine: N6-Hydroxylase. *Biochim. Biophys. Acta - Protein Structure and Molecular Enzymology* **1993**, *1203*, 27-35.

82.   Stehr, M.; Diekmann, H.; Smau, L.; Seth, O.; Ghisla, S.; Singh, M.; Macheroux, P., A Hydrophobic Sequence Motif Common to N-Hydroxylating Enzymes. *Trends Biochem. Sci.* **1998**, *23*, 56-57.

83.   Olucha, J.; Meneely, K. M.; Chilton, A. S.; Lamb, A. L., Two Structures of an N-Hydroxylating Flavoprotein Monooxygenase: Ornithine Hydroxylase from *Pseudomonas aeruginosa*. *J. Biol. Chem.* **2011**, *286*, 31789-31798.

84.   Robinson, R.; Badieyan, S.; Sobrado, P., C4a-Hydroperoxyflavin Formation in N-Hydroxylating Flavin Monooxygenases is Mediated by the 2′-OH of the Nicotinamide Ribose of NADP+. *Biochemistry* **2013**, *52*, 9089-9091.

85.   Stehr, M.; Smau, L.; Singh, M.; Seth, O.; Macheroux, P.; Ghisla, S.; Diekmann, H., Studies with Lysine N6-Hydroxylase. Effect of a Mutation in the Assumed FAD Binding Site on Coenzyme Affinities and on Lysine Hydroxylating Activity. *Biol. Chem.* **1999**, *380*, 47-54.

86.   Ambrosi, C.; Leoni, L.; Putignani, L.; Orsi, N.; Visca, P., Pseudobactin Biogenesis in the Plant Growth-Promoting Rhizobacterium *Pseudomonas* Strain B10: Identification and Functional Analysis of the l-Ornithine N5 -Oxygenase (psbA) Gene. *J. Bacteriol.* **2000**, *182*, 6233-6238.

87.   (a) Dudek, H.; Torres Pazmiño, D.; Rodríguez, C.; Gonzalo, G.; Gotor, V.; Fraaije, M., Investigating the Coenzyme Specificity of Phenylacetone Monooxygenase from *Thermobifida fusca*. *App. Microbiol. Biotechnol.* **2010**, *88*, 1135-1143; (b) Kamerbeek, N. M.; Fraaije, M. W.; Janssen, D. B., Identifying Determinants of NADPH Specificity in Baeyer–Villiger Monooxygenases. *Eur. J. Biochem.* **2004**, *271*, 2107-2116.

88. Hollmann, F.; Lin, P.-C.; Witholt, B.; Schmid, A., Stereospecific Biocatalytic Epoxidation: The First Example of Direct Regeneration of a FAD-Dependent Monooxygenase for Catalysis. *J. Am. Chem. Soc.* **2003**, *125*, 8209-8217.

89. Hilker, I.; Wohlgemuth, R.; Alphand, V.; Furstoss, R., Microbial transformations 59: First Kilogram Scale Asymmetric Microbial Baeyer-Villiger Oxidation with Optimized Productivity using a Resin-Based In Situ SFPR Strategy. *Biotechnol. Bioeng.* **2005**, *92*, 702-710.

90. Cotton, H.; Elebring, T.; Larsson, M.; Li, L.; Sörensen, H.; von Unge, S., Asymmetric Synthesis of Esomeprazole. *Tetrahedron: Asymmetry* **2000**, *11*, 3819-3825.

91. Bong, Y. K.; Clay, M. D.; Collier, S. J.; Mijts, B.; Vogel, M.; Zhang, X.; Zhu, J.; Nazor, J.; Smith, D.; Song, S., Synthesis of Prazole Compounds. *Patent* **2013**, *Patent number: US 2013/0017580 A1*.

92. Dover, L. G.; Alahari, A.; Gratraud, P.; Gomes, J. M.; Bhowruth, V.; Reynolds, R. C.; Besra, G. S.; Kremer, L., EthA, a Common Activator of Thiocarbamide-Containing Drugs Acting on Different Mycobacterial Targets. *Antimicrob. Agents Chemother.* **2007**, *51*, 1055-1063.

93. Fraaije, M. W.; Kamerbeek, N. M.; Heidekamp, A. J.; Fortin, R.; Janssen, D. B., The Prodrug Activator EtaA from *Mycobacterium tuberculosis* is a Baeyer-Villiger Monooxygenase. *J. Biol. Chem.* **2004**, *279*, 3354-3360.

94. Willand, N.; Dirie, B.; Carette, X.; Bifani, P.; Singhal, A.; Desroses, M.; Leroux, F.; Willery, E.; Mathys, V.; Deprez-Poulain, R.; Delcroix, G.; Frenois, F.; Aumercier, M.; Locht, C.; Villeret, V.; Deprez, B.; Baulard, A. R., Synthetic EthR Inhibitors Boost Antituberculous Activity of Ethionamide. *Nat. Med.* **2009**, *15*, 537-544.

95. Allen, M. J.; Tait, K.; Muhling, M.; Weynberg, K.; Bradley, C.; Trivedi, U.; Gharbi, K.; Nissimov, J.; Mavromatis, K.; Jensen, C. N.; Grogan, G.; Ali, S. T., Genome sequence of *Stenotrophomonas maltophilia* PML168, which Displays Baeyer-Villiger Monooxygenase Activity. *J. Bacteriol.* **2012**, *194*, 4753-4754.

96. Rao, S. T.; Rossmann, M. G., Comparison of Super-Secondary Structures in Proteins. *J. Mol. Biol.* **1973**, *76*, 241-256.

97. Fogg, M. J.; Wilkinson, A. J., Higher-Throughput Approaches to Crystallization and Crystal Structure Determination. *Biochem. Soc. Trans.* **2008**, *36*, 771-775.

98. Carrea, G.; Redigolo, B.; Riva, S.; Colonna, S.; Gaggero, N.; Battistel, E.; Bianchi, D., Effects of Substrate Structure on the Enantioselectivity and Stereochemical Course of Sulfoxidation Catalyzed by Cyclohexanone Monooxygenase. *Tetrahedron: Asymmetry* **1992**, *3*, 1063-1068.

99. Reetz, M. T.; Kahakeaw, D.; Lohmer, R., Addressing the Numbers Problem in Directed Evolution. *ChemBioChem* **2008**, *9*, 1797-1804.

100.   Golchoubian, H.; Hosseinpoor, F., Effective Oxidation of Sulfides to Sulfoxides with Hydrogen Peroxide under Transition-Metal-Free Conditions. *Molecules* **2007**, *12*, 304-311.

101.   Brzozowski, A. M.; Walton, J., Clear Strategy Screens for Macromolecular Crystallization. *J. Appl. Crystallogr.* **2001**, *34*, 97-101.

102.   Gustafsson, T. N.; Sandalova, T.; Lu, J.; Holmgren, A.; Schneider, G., High-Resolution Structures of Oxidized and Reduced Thioredoxin Reductase from *Helicobacter pylori*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 833-843.

103.   Orru, R.; Pazmino, D. E.; Fraaije, M. W.; Mattevi, A., Joint Functions of Protein Residues and NADP(H) in Oxygen Activation by Flavin-Containing Monooxygenase. *J. Biol. Chem.* **2010**, *285*, 35021-35028.

104.   Yachnin, B. J.; Sprules, T.; McEvoy, M. B.; Lau, P. C. K.; Berghuis, A. M., The Substrate-Bound Crystal Structure of a Baeyer–Villiger Monooxygenase Exhibits a Criegee-like Conformation. *J. Am. Chem. Soc.* **2012**, *134*, 7788-7795.

105.   Yu, Q.; Schaub, P.; Ghisla, S.; Al-Babili, S.; Krieger-Liszkay, A.; Beyer, P., The Lycopene Cyclase CrtY from *Pantoea ananatis* (Formerly Erwinia uredovora) Catalyzes an FAD red-dependent Non-redox Reaction. *J. Biol. Chem.* **2010**, *285*, 12109-12120.

106.   Krissinel, E.; Henrick, K., Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774-797.

107.   Holm, L.; Sander, C., Mapping the Protein Universe. *Science* **1996**, *273*, 595-602.

108.   McDonald, C. A.; Fagan, R. L.; Collard, F.; Monnier, V. M.; Palfey, B. A., Oxygen Reactivity in Flavoenzymes: Context Matters. *J. Am. Chem. Soc.* **2011**, *133*, 16809-16811.

109.   Andreadeli, A.; Platis, D.; Tishkov, V.; Popov, V.; Labrou, N. E., Structure-Guided Alteration of Coenzyme Specificity of Formate Dehydrogenase by Saturation Mutagenesis to Enable Efficient Utilization of NADP+. *FEBS J.* **2008**, *275*, 3859-3869.

110.   Hatrongjit, R.; Packdibamrung, K., A Novel NADP+-Dependent Formate Dehydrogenase from *Burkholderia stabilis* 15516: Screening, Purification and Characterization. *Enzyme Microb. Technol.* **2010**, *46*, 557-561.

111.   Pazmiño, D. E. T.; Snajdrova, R.; Rial, D. V.; Mihovilovic, M. D.; Fraaije, M. W., Altering the Substrate Specificity and Enantioselectivity of Phenylacetone Monooxygenase by Structure-Inspired Enzyme Redesign. *Adv. Synth. Catal.* **2007**, *349*, 1361-1368.

112.   Völker, A.; Kirschner, A.; Bornscheuer, U.; Altenbuchner, J., Functional Expression, Purification, and Characterization of the Recombinant Baeyer-Villiger

Monooxygenase MekA from *Pseudomonas veronii* MEK700. *App. Microbiol. Biotechnol.* **2008**, *77*, 1251-1260.

113.   Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **2007**, *23*, 2947-2948.

114.   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.

115.   Jensen, C. N.; Ali, S. T.; Allen, M. J.; Grogan, G., Mutations of an NAD(P)H-Dependent Flavoprotein Monooxygenase that Influence Cofactor Promiscuity and Enantioselectivity. *FEBS Open Bio.* **2013**, *3*, 473-478.