Old Yellow Enzymes from Extremophiles: finding and characterizing potential biocatalysts

Suzanne Litthauer

Submitted in fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences
Department of Microbial, Biochemical and Food Biotechnology
University of the Free State

July 2012

Supervisor: Dr. D.J. Opperman
Co-supervisor: Prof. E. van Heerden

DECLARATION

It is herewith declared that this dissertation submitted for the degree Magister Scientiae (Biochemistry) at the University of the Free State is the independent work of the undersigned and has not previously been submitted by her at another university or faculty. Copyright of this dissertation is hereby ceded in favour of the University of the Free State.

Suzanne Litthauer

Department of Microbial, Biochemical and Food Biotechnology

Faculty of Natural and Agricultural Sciences

University of the Free State

South Africa

ACKNOWLEDGEMENTS

CHAPTER 1

LITERATURE REVIEW

CHAPTER 2

SEQUENCE-BASED ANALYSIS OF OYE HOMOLOGUES FROM BACTERIA AND ARCHAEA

CHAPTER 3

CLONING AND HETEROLOGOUS EXPRESSION OF NEW OYE GENES

CHAPTER 4

PURIFICATION AND CHARACTERISATION OF OYES FROM *R. METALLIDURANS* CH3 AND *S. SOLFATARICUS* P2

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH

LIST OF TABLES

LIST OF FIGURES

Fox & Karplus, 1994; Kitzing *et al.*, 2005).

modulator residue. Group A1$_\gamma$ (approximately 40% of OYE homologues) consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. Group B1$_\gamma$ (approximately 15% of the OYE homologues) consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Glu or Ser as flavin modulator. Groups B2$_\gamma$ (approximately 30% of the OYE homologues) and B3$_\gamma$ (approximately 15% of the OYE homologues) both have the His-pair in the catalytic site, but combined with a Cys and Thr residue as flavin modulator, respectively.

Group B4$_\alpha$ (approximately 5% of OYE homologues) consists of OYEs with His-pair combined with a Val in the position of flavin modulator. Group B5$_F$ (approximately 13% of the OYE homologues) have the His-pair in the catalytic site, but combined with a Arg, His or Val in the position of flavin modulator residue.

showing peaks for cyclohexanone and the product, 2-cyclohexenone.

NON-SI ABBREVIATIONS

A	Absorbance
AD	Aldehyde dehydrogenase
BCA	Bicinchoninic acid
BED	β-estradiol
Bicine	N,N-bis(2-hydroxyethyl)-glycine
BLAST	Basic Logical Alignment Search Tool
CL	Citral lyse
bp	Base pairs
BSA	Bovine serum albumin
cDNA	Complimentary DNA
(c-THN)-TPN	alpha-$O^{2'}$-6B-cyclo-1,4,5,6-tetrahydro-nicotinamide adenine dinucelotide phosphate
Da	Daltons
DNA	Deoxyribonucleic acid
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
dNTPs	Deoxyribonucleoside triphosphates
EBP	Estrogen binding protein
EDTA	Ethylenediaminetetraacetate
ER	Enoate reductase
FAD	Flavin adenine dinucleotide
FDH	Formate dehydrogenase
FMN	Riboflavin 5'-monophosphate
G6PDH	glucose-6-phosphate dehydrogenase
GDH	Glucose dehydrogenase
gDNA	Genomic DNA
IMAC	immobilized metal-affinity chromatography

IPTG	Isopropyl β-D-thiogalactoside
K_{cat}	Catalytic constant
K_m	Michaelis constant
LB	Luria-Bertani broth
LR	Levodione reductase
MES	2-(N-morpholino)ethanesulfonic acid
ML	Maximum likelihood
MOPS	3-(N-morpholino)propanesulfonic acid
MR	Morphinone reductase from *Pseudomonas putida*
M_r	Molecular weight
MUSCLE	Multiple Sequence Comparison by Log-Expectation
MUSTER	Multi-Source Threader
MWCO	Molecular weight cut off
NADH	Nicotinamide adenine dinucleotide (reduced)
NADPH	Nicotinamide adneine dinucleotide phosphate (reduced)
NCBI	National Center for Biotechnology Information
NEMR	*N*-ethylmaleimide reductase
NNI	Nearest-Neighbour-Interchange
OD	Optical density
OPR	12-oxophytodienoate reductase from *Solanum lycopersicum*
ORF	Open reading frame
OYE	Old Yellow Enzyme
OYE1	Old Yellow Enzyme homologue form *Saccharomyces pastorianus*
OYE2	Old Yellow Enzyme homologue from *Saccharomyces cerevisiae*
*p*ADH	*prim*-alcohol dehydrogenase
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank

PETN reductase	Pentaerythritol tetranitrate reductase from *Enterobacter cloacae*
PHB	*p*-hydroxybenzaldehyde
*Cm*OYE	OYE homologue from *Cupriavidus metallidurans* CH34
*s*ADH	*sec*-alcohol dehydrogenase
SDS	Sodium dodecyl sulphate
*Ss*OYE	OYE homologue from *Sulfolobus solfataricus* P2
TAE	Tris, Acetic acid, EDTA
TIM-barrel	$(\alpha/\beta)_8$-barrel fold
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
U	Units
V_e	Elution volume
V_o	Void volume
V_{max}	Maximum initial velocity
WAG model	Whelan and Goldman model
XenA	Xenobiotic reductase A from *Pseudomonas putida*
x *g*	Gravitational force
X-Gal	5-bromo-4-chloro-3-indolyl β-D-galactoside
YqjM	Old Yellow Enzyme homologue from *Bacillus subtilis*

CHAPTER 1

LITERATURE REVIEW

1.1. INTRODUCTION

The old yellow enzyme (OYE) family is a diverse group of flavoproteins that has been shown to catalyse the asymmetric reduction of activated C=C bonds of a wide variety of α/β-unsaturated carbonyl compounds, many of which are of importance in the synthesis of fine chemicals and pharmaceuticals (Oberdorfer *et al.*, 2011; Toogood *et al.*, 2010). First described as a yellow enzyme involved in the oxidation of NADPH by molecular oxygen, old yellow enzyme was the first flavin-containing enzyme characterised and played an important role by serving as a model enzyme in studies aimed at understanding the role of flavin cofactors in proteins (Karplus *et al.*, 1995). More than 70 years later, the substrate range of OYE has broadened to now also include nitroalkenes, carboxylic acids, nitrate esters, nitroglycerin, nitroaromatic explosives and cyclic triazines, while the physiological function of the enzyme remains mostly unknown (Toogood *et al.*, 2010; Williams & Bruce, 2002).

The first OYE described (OYE1), was isolated from *Saccharomyces pastorianus* (brewers' bottom yeast) in 1932 by Warburg and Christian and called 'das gelbe Ferment' (the yellow enzyme) (Warburg & Christian, 1932). Two years later, the discovery of a second yellow enzyme (OYE2) from *Saccharomyces cerevisiae* led to the renaming of the first as the "old yellow enzyme" (OYE) – a name that is still used to describe this family of flavoenzymes (Haas, 1938). In the early 1990s, the OYEs purified and characterised from *S. pastorianus* and *S. cerevisiae* were found to be heterogenous in terms of chromatographic behaviour and structure, with at least two variants of amino acid sequence observed for each OYE (Karplus *et al.*, 1995). Probing of cDNA libraries from *S. pastorianus* and *S. cerevisiae* revealed that both yeast genomes contained two OYE genes (Karplus *et al.*, 1995; Stott *et al.*, 1993). The first successful expression of homogeneous recombinant OYE in *E. coli* - OYE1 from *S. pastorianus* and OYE2 from *S. cerevisiae* – allowed comparisons to the previously purified heterogeneous proteins. Subsequent studies revealed that these OYEs exist as both homo- and heterodimers that arise from two OYE isozymes in each species, with each subunit containing FMN as cofactor. It was even shown that an 'unnatural'

OYE heterodimer can be created by mixing the OYE1 and OYE2 monomers from the two different yeast species (Stott *et al.,* 1993).

In the following years, a large number of OYE family members has been identified in bacteria, yeast, plants and parasitic eukaryotes, mostly due to the advent of automated protein sequence alignments and the availability of large genome data sets from whole genome sequencing (Toogood *et al.*, 2010; Williams & Bruce, 2002). While originally classified according to the enzyme's ability to bind phenolic compounds and the subsequent development of intense long wavelength absorbance bands in absorption spectra (Karplus *et al.*, 1995), enzymes are now regarded as OYE homologues if significant amino acid sequence and/or structural homology to known OYE members are apparent (Toogood *et al.*, 2010; Williams & Bruce, 2002). While OYEs have been shown to function in a variety of oligomeric states – from monomers to dodecamers – the overall monomeric structure is conserved throughout the whole OYE family. This review will introduce the overall structural characteristics of the OYE family, while also demonstrating their biocatalytic potential with respect to the vast substrate range and diverse catalytic abilities of these flavoproteins.

1.2. STRUCTURAL ASPECTS

Members of the OYE family are known to exist as monomers, dimers, tetramers and even in multiple oligomeric states, such as octamers and dodecamers, as observed among some thermostable OYEs (Adalbjörnsson *et al.*, 2010; Kitzing *et al.*, 2005; Toogood *et al.*, 2010). The most significant structural differences between homologues involve the position and amino acid composition of surface loops, between homologues in different sub-classes, as well as in the same subclass (Oberdorfer *et al.*, 2011; Toogood *et al.*, 2010). Generally, loops occurring at the carboxy-terminal end of the barrel form the active site and some similarity in these loops is observed among members of the OYE family (Kitzing *et al.*, 2005). The conserved monomeric structure of all analysed OYEs is a characteristic $(\alpha/\beta)_8$-barrel (TIM-barrel) fold, with additional secondary structural elements that tend to vary among homologues (Oberdorfer *et al.*, 2011). Sequence and structural differences that do occur are distinct and appear to divide the OYE homologues into two subclasses – "thermophilic-like" OYEs and "classical" OYEs (such as OYE1). In comparison to the "classical" counterparts, "thermophilic-like" OYEs (so called due to the dominance of thermophilic OYEs in the subclass) display compact monomers, variation in interaction angles at the subunit-subunit interface, shortened surface loops and changes in subunit-subunit interface, all of which may play a role in the thermostability of these OYE homologues (Oberdorfer *et al.*, 2011; Toogood *et al.*, 2010).

1.2.1. The two OYE subclasses: overview of structure and active-site architecture

1.2.1.1.   Monomeric structure

OYE1 from *S. pastorianus* serves as a satisfactory model to demonstrate the monomeric structure of the classical OYE family (figure 1.1). OYE1 is described as a 399 residue protein folded into a single, eight-stranded α/β-barrel domain, along with four additional β-strands and five additional α-helices (Fox & Karplus, 1994). Two β-strands, βA and βB, form a β-hairpin before the first β-strand of the barrel (β1) which caps the N-terminal end of the protein, closing the inside of the barrel off to the surrounding solvent. A compact 36-residue subdomain occurs at the C-terminal end between β3 and α3, consisting of helices αA and αB, as well as strands βC and βD. Situated between strand β8 and helix α8, αD binds the flavin phosphoryl groups at its amino terminus (Fox & Karplus, 1994).

Homologues of the "thermophilic-like" subclass show similarity to YqjM from the soil bacterium *Bacillus subtilis* (Kitzing *et al.*, 2005). The monomeric structure of YqjM (figure 1.2) is described as a compact single TIM barrel domain with short, 3-4 residue loops at the amino-terminal end of the barrel and longer loops (5 to 30 residues long) at the carboxy-terminal end that form the active site. However, YqjM differs structurally from the "classical" OYE1 in numerous ways. A 3_{10}-helix and two short antiparallel β-sheets (before β1) cap the barrel at the amino-terminal end, while additional secondary structural elements (4 helices and two strands) occur in the C-terminal loops. The active site region of YqjM is distinctly different from that of the "classical" OYEs, particularly with respect to the conformation and length of loops L3, L5 and L6 (loops are labelled according to the helix they precede). In the "classical" OYEs 12-oxophytodienoate reductase 1 (OPR1) from *Solanum lycopersicum* (tomato), pentaerythritol tetranitrate (PETN) reductase (from *Enterobacter cloacae*) and morphinone reductase (MR; from *Pseudomonas putida* M10), the equivalent of loop L3 forms part of the active site's hydrophobic tunnel, being extended and folding into an additional two-stranded β-sheet which covers half of the active site (Barna *et al.*, 2001; Barna *et al.*, 2002; Breithaupt *et al.*, 2001; Kitzing *et al.*, 2005). This variation is expected to play a role in substrate discrimination. At the active site of YqjM, the loop is folded to form a short 3_{10}-helix packed on the wall of the barrel, resulting in an active site that is widely accessible to potential substrates (Kitzing *et al.*, 2005).
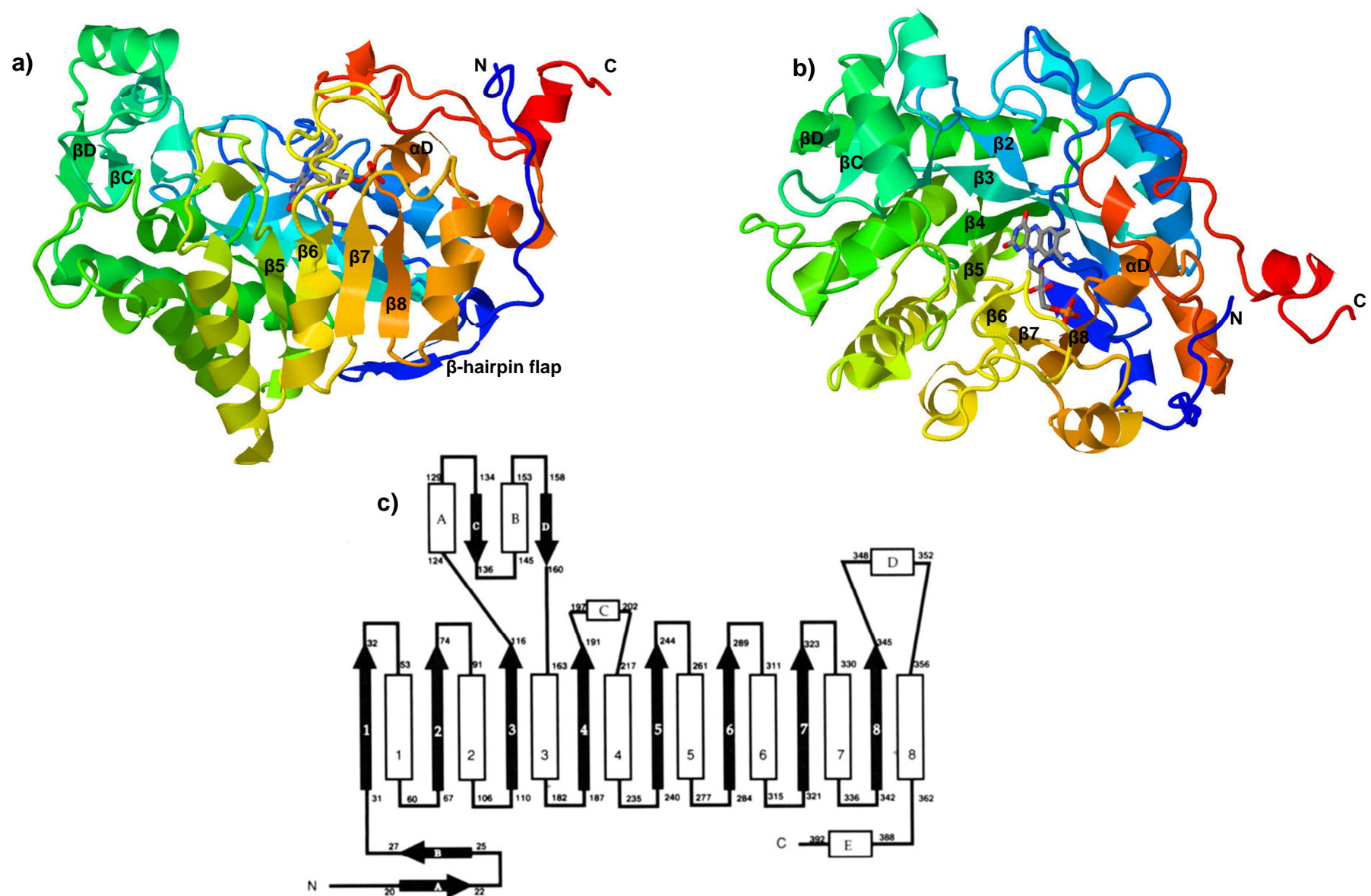
Figure 1.1: Overall monomeric structure of OYE1. Diagrams of a single subunit, with barrel views **a)** from the side and **b)** perpendicular to the barrel. Diagrams are coloured with gradient from blue (N-terminal) to red (C-terminal). FMN is shown as a stick model. Most barrel strands, as well as the β-hairpin flap, strands βC and βD, helix αD and termini, are labeled. Diagram generated from PDB files in JMol. **c)** Topography of OYE1. Strands (arrows) and helices (boxes) of the barrel are numbered 1-8 in order as they appear. Extra-barrel elements are identified with letters. Numbers indicate the residue ranges involved in the secondary structural elements (Adapted from Fox & Karplus, 1994)

4
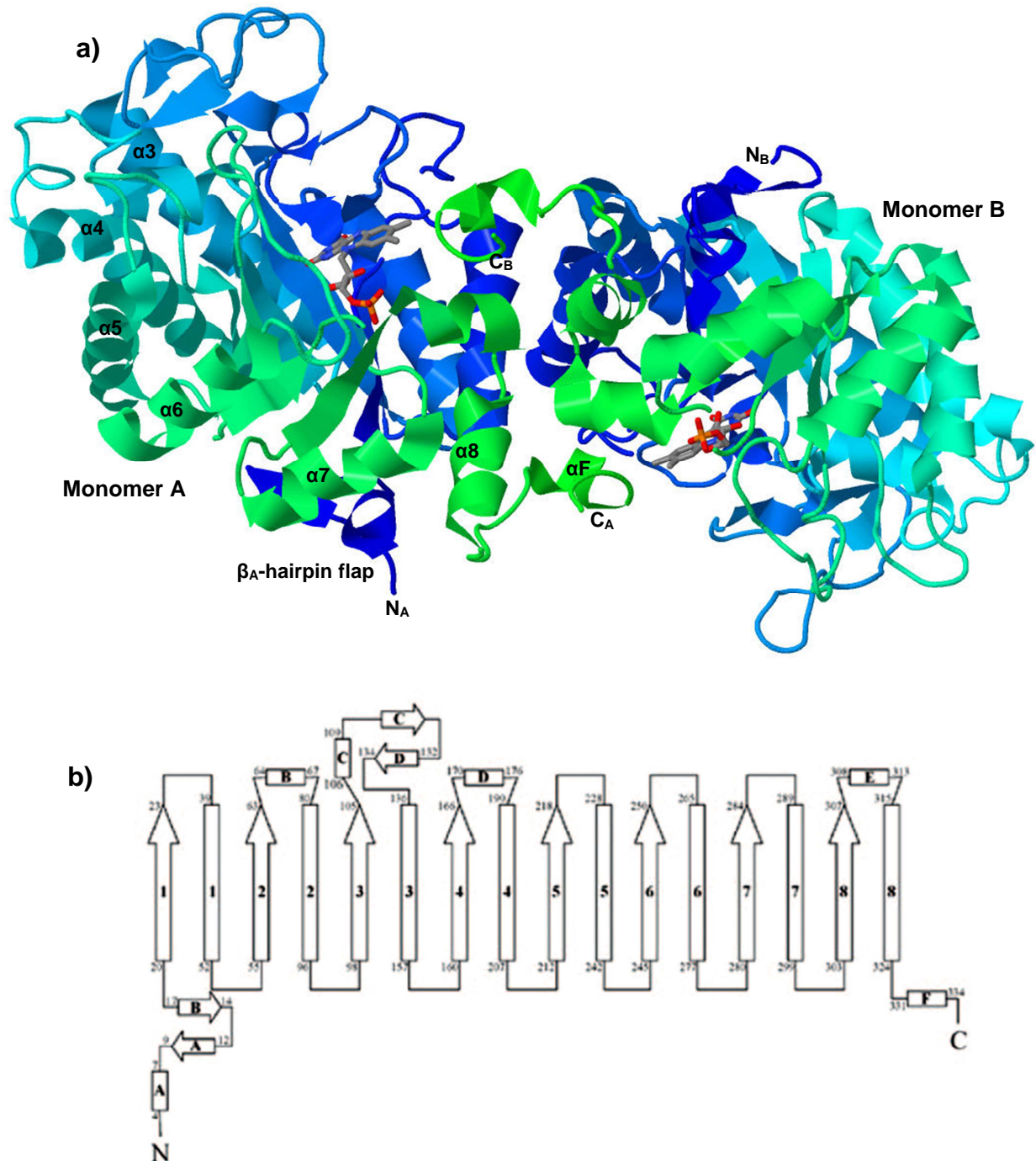
Figure 1.2: Overall dimeric structure of YqjM. **a)** Diagram of the functional unit (dimer), coloured with gradient from blue (N-terminal) to green (C-terminal). FMN is shown as a stick model. Barrels are viewed from the side, with monomers A and B labeled. β-hairpin flap and termini for both subunits are labeled, as well as most barrel helices. Diagram generated from PDB files in JMol. **b)** Topography of YqjM. Strands (arrows) and helices (boxes) of the barrel are numbered 1-8 in order as they appear. Extra-barrel elements are identified with letters. Numbers indicate the residue ranges involved in the secondary structural elements (Adapted from Kitzing *et al.*, 2005).

1.2.1.2. Dimer interface

The dimer interface of OYE1 (figure 1.3) represents about 5% of the monomer's total surface area and involves helices 4, 5 and 6 in both monomers, each interacting with its symmetry mate (Fox & Karplus, 1994). The symmetry-related amino termini of the two monomers' helix 4 point directly towards each other and involve a potentially electrostatically unfavourable interaction which is compensated for by hydrogen bonding involving the hydroxyls of Ser216 and the carboxylates of Glu218. These interactions, along with five direct hydrogen bonds between the monomers (involving a Thr, two Gly, two Ser and one Lys residue) and six hydrogen bonds employing one or two bridging water molecules, constitute the dimer interface. The dimer interface formed by helices 4, 5 and 6 are similar to the dimer interface observed in morphinone reductase (involving helices 2 and 8, as well as the amino-terminal β-strands), but is in stark contrast to the dimer interfaces observed in the "thermophilic-like" OYE YqjM (Kitzing et al., 2005).
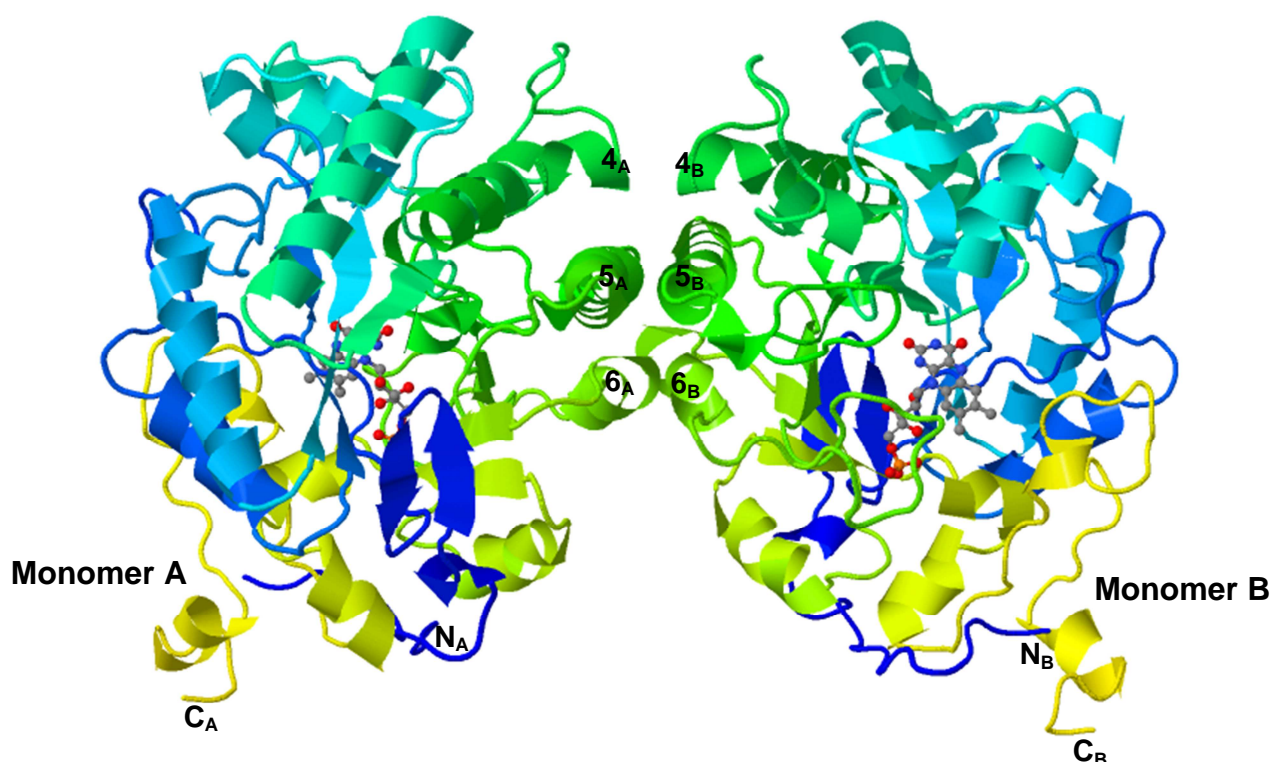
Figure 1.3: Diagram of OYE1 dimer, coloured with gradient from blue (N-terminal) to yellow (C-terminal). FMN shown as a ball-and-stick model and monomers are labeled A and B. For both monomers, helices 4, 5 and 6 involved in the dimer interaction are labeled accordingly. Termini for both monomers are also labeled. Diagram generated from available PDB files in JMol (Adapted from Fox & Karplus, 1994).

For "thermophilic-like" YqjM, interaction between corresponding monomers (figure 1.2) occurs firstly at helix 1 of both monomers (where the subunits are in close proximity), around the 2-fold dimer axis which passes between helix 1 of the two monomers (Kitzing *et al.*, 2005). Monomers are orientated such that the axes of the two β-barrels are approximately parallel to each other, while the monomers open up in opposite directions. In addition, helix αE and the entire carboxy-terminus (which includes helix αF) interact with their symmetry mates in the corresponding monomers. Interactions between helix α1 and αF of the two monomers occur such that the sites are arranged perpendicularly and are stabilized by the formation of a large hydrophobic cluster formed by Met, Pro, Phe, Ile residue pairs and one Ala residue. A hydrogen bonding network of Ser29, His44, Arg48, Leu311, Arg312, Gln333, Tyr334, Arg336, Gly337 and Trp338 encloses the hydrophobic cluster (Kitzing *et al.*, 2005). The section of the functional dimer-dimer interface constituted by C-terminal residues Arg312, Gln333, Tyr334 and Arg336 are highly conserved among members of the "thermophilic-like" OYE sub-class (Toogood *et al.*, 2010).

Evidence suggests that dimers are the active unit of YqjM, with the protein occurring in a tetrameric form (figure 1.4) (Kitzing *et al.*, 2005). Firstly, 5% of the accessible monomer surface acts as contact area between monomers A and C, while monomers A and D do not interact at all. In addition, conjoining monomers appear to share an active site: the hydrophobic cluster directs the carboxy-terminal end of one subunit towards the active site of the neighbouring monomer, with Arg336 from the carboxy-terminal protruding as an Arg finger into the active site of the adjacent monomer, forming part of the substrate binding pocket. This active site architecture contrasts with the "classical" OYEs (Fox & Karplus, 1994; Kitzing *et al.*, 2005). The carboxy-terminus of morphinone reductase forms similar structures than that observed in YqjM, but these fold back into their own subunit and do not contribute to the active site of the neighbouring monomer (Barna *et al.*, 2002). Tetrameric YqjM (resembling a four-petalled clover leaf with a central hole) is assembled such that the active site of the monomers open up in different directions to the solvent, while being connected via the central hole (Kitzing *et al.*, 2005). Dimers AB and CD connect mainly though hydrophobic interactions (involving Val260, Phe261, Pro262, Tyr264, Val266, Met285 and Met291 and their symmetry mates) and, to a lesser extent, polar interactions (established by Gly263, Glu270, Asn298 and their symmetry mates) that occur in helix-loop-helix motifs of helices 6 and 7 (Kitzing *et al.*, 2005).
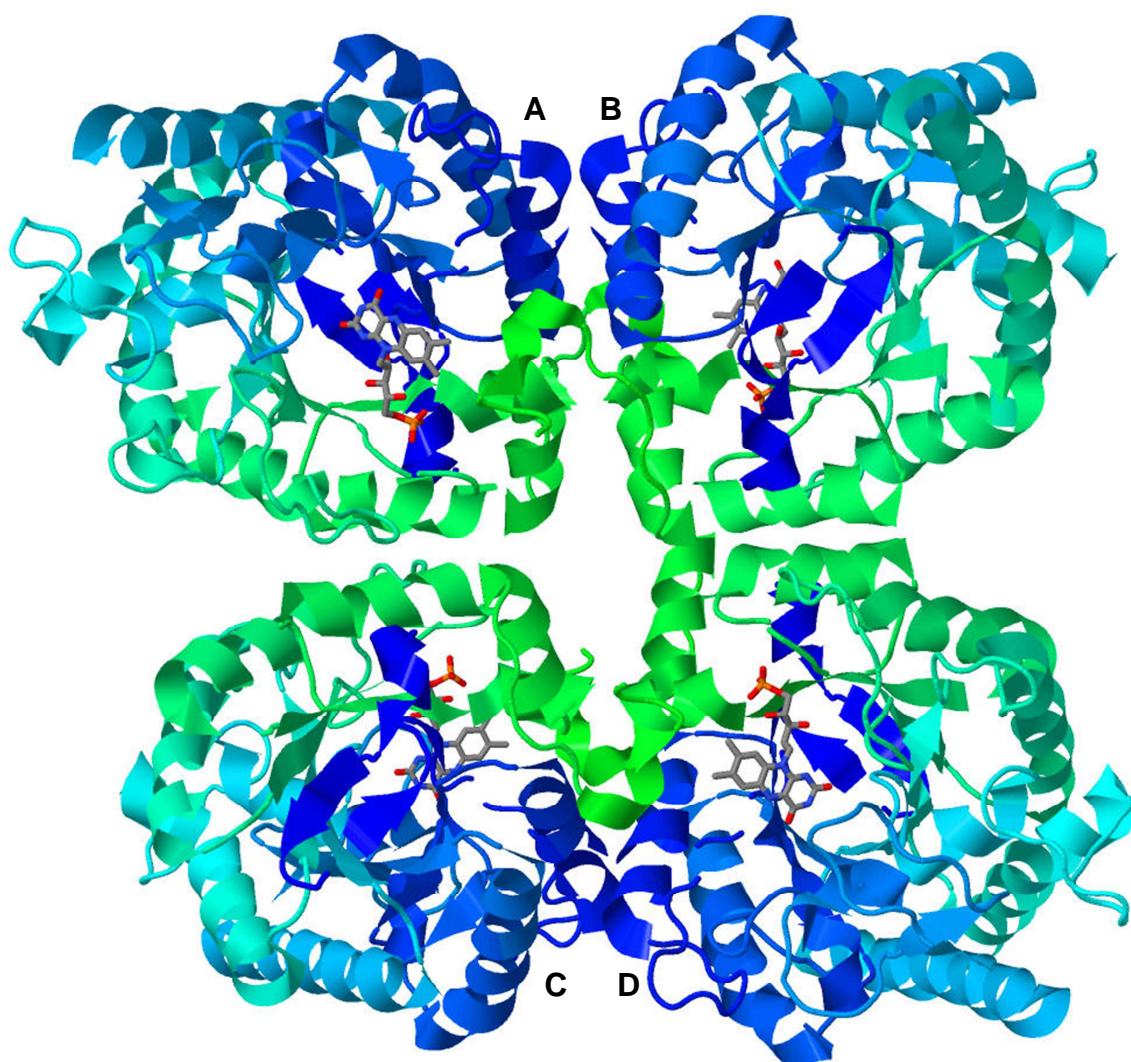
Figure 1.4:   Diagram of tetrameric YqjM (formed by interaction of functional subunits AB and CD), coloured with gradient from blue (N-terminal) to green (C-terminal). FMN shown as a stick model and monomers are labeled A-D. Diagrams generated from PDB files in JMol (Adapted from Fox & Karplus, 1994).

In OYE1, FMN is situated near the centre of the α/β-barrel, bound at the carboxy-terminal end and partly buried by carboxy-terminal loops that cover the end of the barrel and isolate the barrel's interior from solvent (Fox & Karplus, 1994). While the *re*-face of the flavin is completely buried by strand β1, the surrounding solvent gains access to the *si*-face at the bottom of the active site pocket. In the flavin binding environment, FMN is held in place by an extensive hydrogen bonding network (figure 1.5), mostly through direct interactions with the flavin's ribityl group and pyrimidine ring. Only four protein side-chains are involved in direct hydrogen bonding to the flavin, all from strands β1, β3, β5 and β8: Thr37 binds to the isoalloxazine ring system through one hydrogen bond to O4 and Gln114 through two (to O2 and N3), while Arg348 and Arg243 bind the ribityl chain through two and three hydrogen bonds respectively. These residues that bind the flavin through side-chain interactions are highly conserved in both OYE sub-classes (Toogood *et al.*, 2010). The remaining hydrogen bonds involve the main chain carbonyl and amide groups of residues present on the ends of strands β1, β2, β7 and β8, along with two water molecules (Fox & Karplus, 1994). Phe35 binds the O2' atom through interaction with its main-chain carbonyl oxygen and is highly conserved among members of both OYE sub-classes, while the remaining residues that bind the flavin through interactions with main-chain atoms are poorly conserved (Toogood *et al.*, 2010). Binding of the phosphoryl group of the flavin is achieved through two tripeptide groups – Gly324, Asp325, Phe326 from turn after β7 and Gly347, Arg348 and Phe349 from turn after β8 (Fox & Karplus, 1994). Residues in these tripeptides adopt conformations that allow the main chain amides to form hydrogen bonds with the phosphoryl group of the flavin, either directly or through a water molecule. The more hydrophobic end of the flavin is less tightly bound, with the aromatic side chains of Phe296, Phe374 and Tyr375 loosely surrounding the dimethylbenzene edge of the flavin, allowing solvent access to the flavin (Fox & Karplus, 1994).

YqjM binds the flavin cofactor (figure 1.6) in a similar manner and orientation as its "classical" OYE counterparts – the pyrimidine ring is anchored by interactions with Gln102 (with N3 and O2) and Arg215, as well as hydrogen bonding of N1 to His167 and N3 to His164 (Kitzing *et al.*, 2005). Side chains of Arg215, Ser23, Ser249 and Gn265, as well as the main chain of Pro24, bind the ribityl chain. Ser23, which binds the O2' atom of FMN by interaction with its side-chain, is highly conserved among members of the "thermophilic-like" OYE subclass (Toogood *et al.*, 2010). The FMN phosphate group is embedded in an electropositive groove formed by loops 7 and 8, with the extensive network of polar interactions (between Arg, Gly, Met, Phe and Glu residues with the phosphate group) allowing the binding of FMN as opposed to FAD (Kitzing *et al.*, 2005). However, instead of the Phe residue observed in "classical" OYEs, the Arg residue protruding from the neighbouring monomer forms part of the lateral part of the flavin binding pocket. As opposed to a

Thr residue in "classical" OYEs, O4 of the isoalloxazine ring is bound to Cys26 through hydrogen bonding – a residue that can adopt numerous conformations to interact with either O4 or N5 through hydrogen bonding, while the side chains of Cys26 can interact with Tyr28. In the oxidised state, the FMN cofactor of YqjM is butterfly-bent at the N5-N10 hinge, which is not intensified upon reduction. This is in contrast to "classical" OYEs OPR, MR and OYE1, where the isoalloxazine ring of oxidised FMN is planar and undergoes butterfly bending upon reduction (Fox & Karplus, 1994; Kitzing *et al.*, 2005).
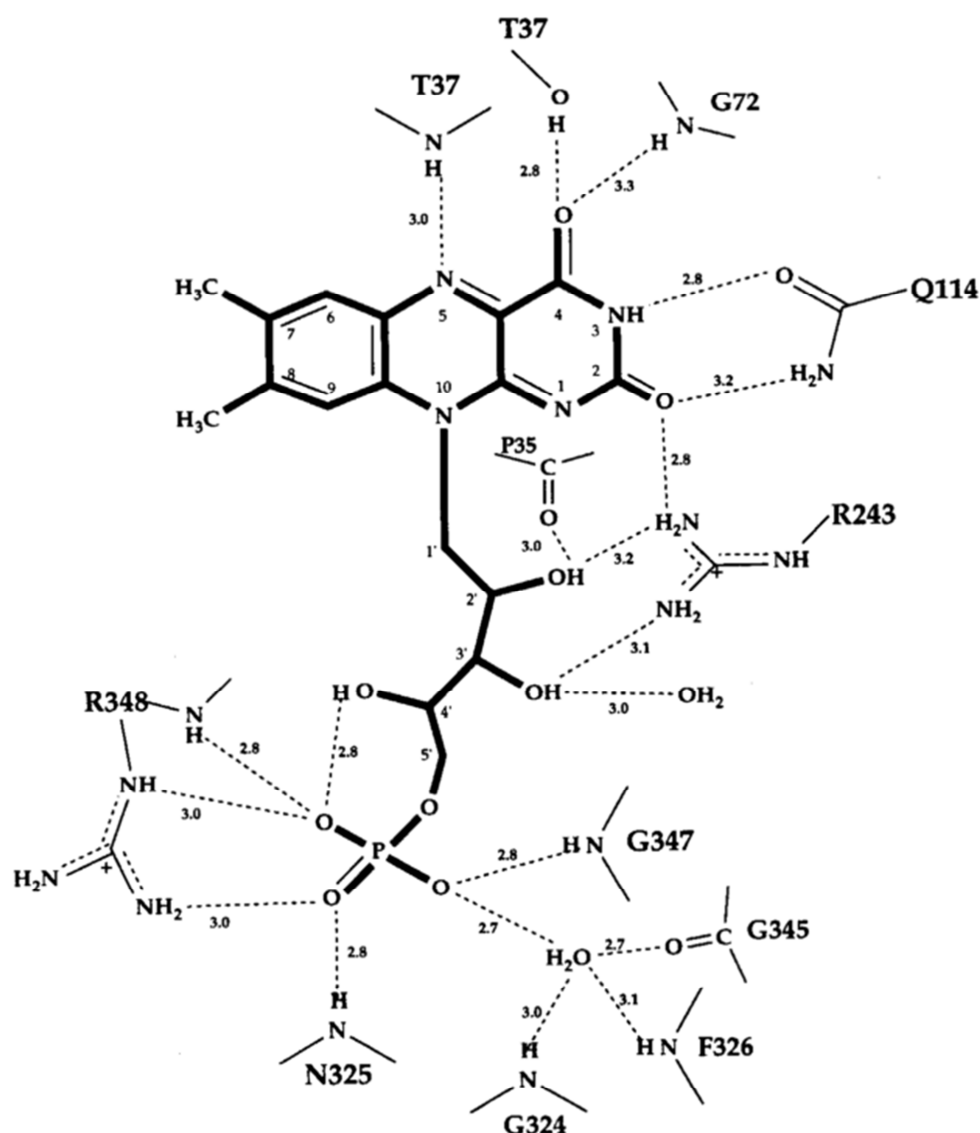
Figure 1.5: Flavin environment of OYE1 showing FMN and the interactions with surrounding amino acid residues (labeled). Hydrogen bonds indicated with dotted lines (Fox & Karplus, 1994).
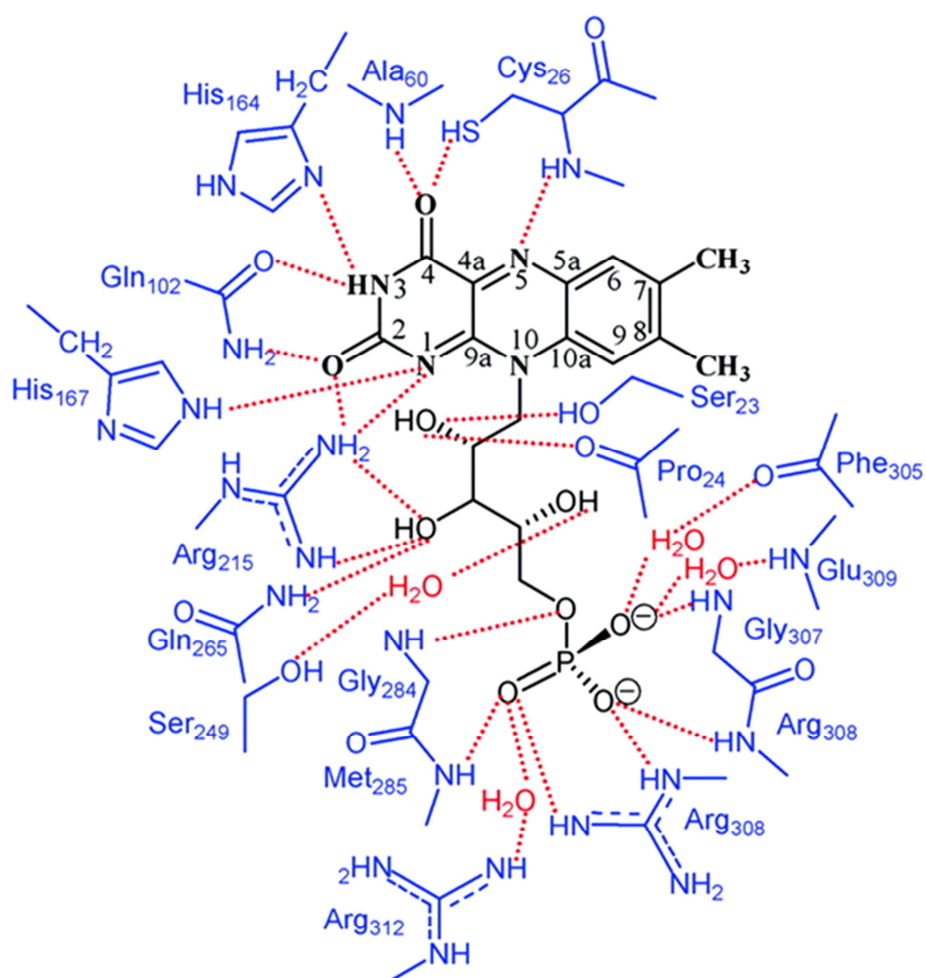
Figure 1.6: Flavin environment of YqjM showing FMN and the interactions with surrounding amino acid residues (labeled). Hydrogen bonds indicated with dotted line (Kitzing et al., 2005).

The importance of Gln114 in OYE1 was highlighted when mutation of Gln114 to Asn – a residue with side chain one methylene group shorter - caused little change to the global enzyme structure, but decreased the enzyme's ability to bind ligands (Brown *et al.*, 2002). The Q114N mutation moves the amide group of residue 114 out of H-bonding distance of O2 and N3 of the isoalloxazine ring. This results in the FMN prosthetic group to be repositioned to form new interactions to replace the lost H-bonds: compared to the wild type, the isoalloxazine is shifted and tilted at a 12˚ angle, allowing atom O2 to hydrogen bond to the main-chain nitrogen of Asn194, a guanidino nitrogen on the side-chain of Arg243 and the His191 side-chain. The disruption of the substrate binding site is demonstrated by the absence of a chloride ion which occurs in crystallised wild type OYE1. This chloride ion occurs in the substrate binding pocket near His191 and Asn194 and is displaced upon ligand binding (Fox & Karplus, 1994). Due to change in orientation of the FMN, His191 and Asn194 are involved in hydrogen-bonding of the flavin and thus unavailable for ligand binding (Brown *et al.*, 2002).   Analysis of the structure of the Q114N mutant complexed to *p*-hydroxybenzaldehyde revealed that binding of the ligand resulted in FMN moving back to a similar position as in wild type OYE, but with the N3 atom of the flavin buried without hydrogen-bonding partners. However, the resulting limited movement of the flavin combined with the unavailability of residues His191 and Asn194 was shown to cause a lower binding affinity for typical OYE substrates, while also being detrimental to catalysis with NADPH and cyclohexenone (Brown *et al.*, 2002).

The role of Thr37 was also investigated by mutagenesis, with construction of a T37A-OYE1 mutant and subsequent ligand binding studies (Xu *et al.* 1999). While mutation of the Thr37 to an Ala residue had a minimal effect on the structure or substrate-binding ability of the enzyme, the absence of the hydroxyl functional group appeared to affect the redox potential of the enzyme. The hydrogen bond formed between O4 of the FMN and the side-chain hydroxyl group of Thr37 in wild type OYE causes electrons to be withdrawn from the flavin ring. This results not only in enhanced electrophilicity of the oxidised enzyme, but also serves to stabilise the negatively charged flavin. Reduction of the oxidised T37A mutant by NADPH was impaired while rates of the oxidation-half reaction were increased. This indicated that the lower redox potential of the T37A mutant rendered the reduced substrate (NADPH) less favourable, while the rate of transfer of electrons to oxidant substrates is increased (Brown *et al.*, 2002).

Similarly, the importance of the Cys residue (corresponding to Cys26 in YqjM) involved in binding of the flavin in the "thermophilic-like" OYE subclass, was demonstrated through mutagenesis studies with xenobiotic reductase A (XenA), a "thermophilic-like" OYE from *Pseudomonas putida* 86 (Spiegelhauer *et al.*, 2010). In XenA, the Cys25 residue corresponds to Cys26 in YqjM and appears to modulate the electronic structure of the flavin's isoalloxazine ring by restraining its planarity. The

role of Cys25 in XenA was studied by replacement of the residue with a serine and an alanine residue respectively. Exchange of Cys25 to alanine had very little effect on the reduction potential, reactivity and structure of XenA. However, exchange of Cys25 to serine resulted in an increase in the reduction potential, an increase in the rate constant of the reductive half-reaction and a decrease in the rate constant of the oxidative half-reaction. It was proposed that the role of Cys25 in XenA lies in modulating the reduction potential of the enzyme (Spiegelhauer *et al.*, 2010). A more in depth discussion of the flavin modulator residue in OYE homologues can be found in chapter 2.

1.2.1.4.   Active site residues – substrate binding and catalysis

Binding of substrate and nicotinamide cofactor by OYE1 (figure 1.8a) was investigated by structural analysis of OYE complexed with *p*-hydroxybenzaldehyde, β–estradiol and the NADP analogue (c-THN)TPN (figure 1.7) (Fox & Karplus, 1994). Upon binding of a phenolic compound, the phenol ring is stacked above and parallel to the *si*-face of the isoalloxizine ring, orientating the phenolate oxygen above the C2 atom of the flavin. The phenolic compound is orientated such that hydrogen bonding occurs between the phenolate oxygen and hydrogen bond-donating pair His191 and Asn194, as well as between the aldehyde carbonyl and the hydroxyl of Tyr375. Positioning of the phenolate oxygen facilitates formation of a charge-transfer complex where the phenolate ion acts as the charge-transfer donor and the isoalloxazine ring of FMN as acceptor. Binding of β-estradiol and the NADP analogue (c-THN)TPN indicated that residues His191 and Asn194 were involved in a similar manner in binding these two compounds. The amide oxygen of the nicotinamide ring of the NADP analogue is positioned such that C-4 of the nicotinamide ring is close to N-5 of the FMN, facilitating hydride transfer to the flavin. The involvement of His191 and Asn194 was confirmed by construction of mutants H191N and N194H and double mutant H191N/N194H (Brown *et al.*, 1998).
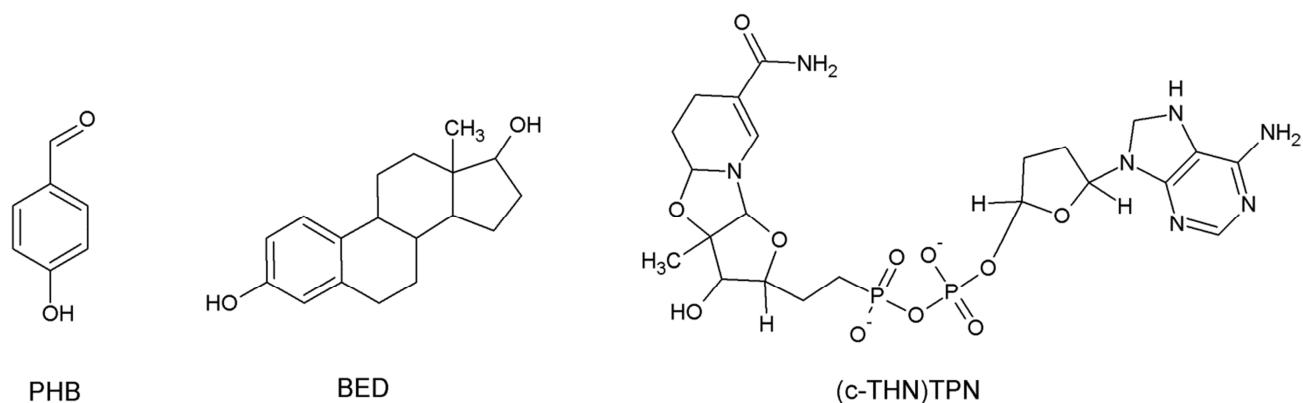
Figure 1.7: Chemical structures of the OYE ligands used during crystallographic analysis: *p*-hydroxybenzaldehyde (PHB), β-estradiol (BED) and alpha-$O^{2'}$-6B-cyclo-1,4,5,6-tetrahydro-nicotinamide adenine dinucelotide phosphate [(c-THN)TPN]. (Adapted from Fox & Karplus,1994).

Upon replacement of His191 with Asn, binding affinity for phenolic compounds and the reduction of cyclohexenone (using NADPH as reductant) were greatly reduced, indicating the importance of His191 in the orientation of cyclohexenone to facilitate hydride transfer from N5 of the flavin at the Cβ (Brown *et al.*, 1998). Mutation of Asn194 to His affected the ability of the enzyme to bind FMN. This was a surprising result, as OYE homologues from the "thermophilic-like" subclass, including YqjM, are known to have two histidine residues in these positions. Impaired binding of phenolic compounds and reduction of cyclohexenone was also observed for the double mutant, as the His and Asn residues, although present, are orientated differently in the active site. Binding of and hydride transfer from the nicotinamide cofactor was not greatly affected in the mutants, possibly due to other factors such as π-π interactions and binding to additional amino acids. While the NADPH-dependant reduction of FMN to $FMNH_2$ was not impaired, the above mutations were detrimental to the reduction of the enone by $FMNH_2$ (Brown *et al.*, 1998).

Another residue of importance in the catalytic site of OYE is Tyr196, being well-positioned to act as an acid by donating a proton to the α-carbon of enones during reductions (Kohli & Massey, 1998). The role of Tyr196 in catalysis in OYE1 was studied by mutation of Tyr196 to phenylalanine. As changes to the structure of the active site and the distribution of electrons in the FMN are minimal in the Y196F mutant, ligand binding and the reductive half-reaction were left unaffected. However, the oxidative half-reaction with α/β-unsaturated carbonyl compounds was slowed drastically, as reduction requires both hydride transfer from the flavin to C_β and donation of a proton from Tyr196 to C_α. Since no proton donor is oriented toward the C_α, the Y196F mutant is equipped to only perform hydride transfer (Kohli & Massey, 1998).

The substrate binding pocket of YqjM (figure 1.8b) differs rather significantly from that of the "classical" OYEs (Kitzing *et al.*, 2005). Structural analysis of enzyme complexed to phenolic aromatic compounds, revealed that substrate is bound in such a way as to align the compound parallel above the *si*-face of the FMN, allowing hydride transfer from N5 of the flavin. Proper orientation of the substrate *p*-hydroxybenzaldehyde is achieved through hydrogen bonding of the anionic hydroxyl group of the ligand (via the phenolic oxygen) to residues His164 and His167, while the aldehyde group is bound by the hydroxyl group of Tyr28. This results in the aldehyde oxygen being rotated 180° and pointing in the opposite direction as observed in substrate binding by "classical" OYEs. These residues responsible for the binding of the proximal functional group are similar in both YqjM and the "classical" OYE homologues OYE1, MR, OPR1 and PETN reductase, with motifs consisting either of His pairs or His and Asn residues. In fact, the active site of OYEs in both sub-classes consist mainly of aromatic residues, with the pair of hydrogen bond-donating residues (responsible for substrate, cofactor and inhibitor binding) being conserved (Toogood *et al.*, 2010). While the equivalent of YqjM Tyr28 is also present (and serves the same function) in the "classical" OYE, the Tyr375 from OYE1 occurs on a COOH-terminal fragment, whereas Tyr28 from YqjM occurs on the amino-terminal fragment and is properly positioned by Cys26 (Kitzing *et al.*, 2005). When binding *p*-nitrophenol, proper positioning of the substrate is achieved through additional hydrogen bonding of the nitro groups to Arg336* that protrudes into the binding pocket from the neighbouring monomer. Tyr169 of YqjM corresponds to the acid catalyst residue Tyr196 in OYE1 (Kitzing *et al.*, 2005).
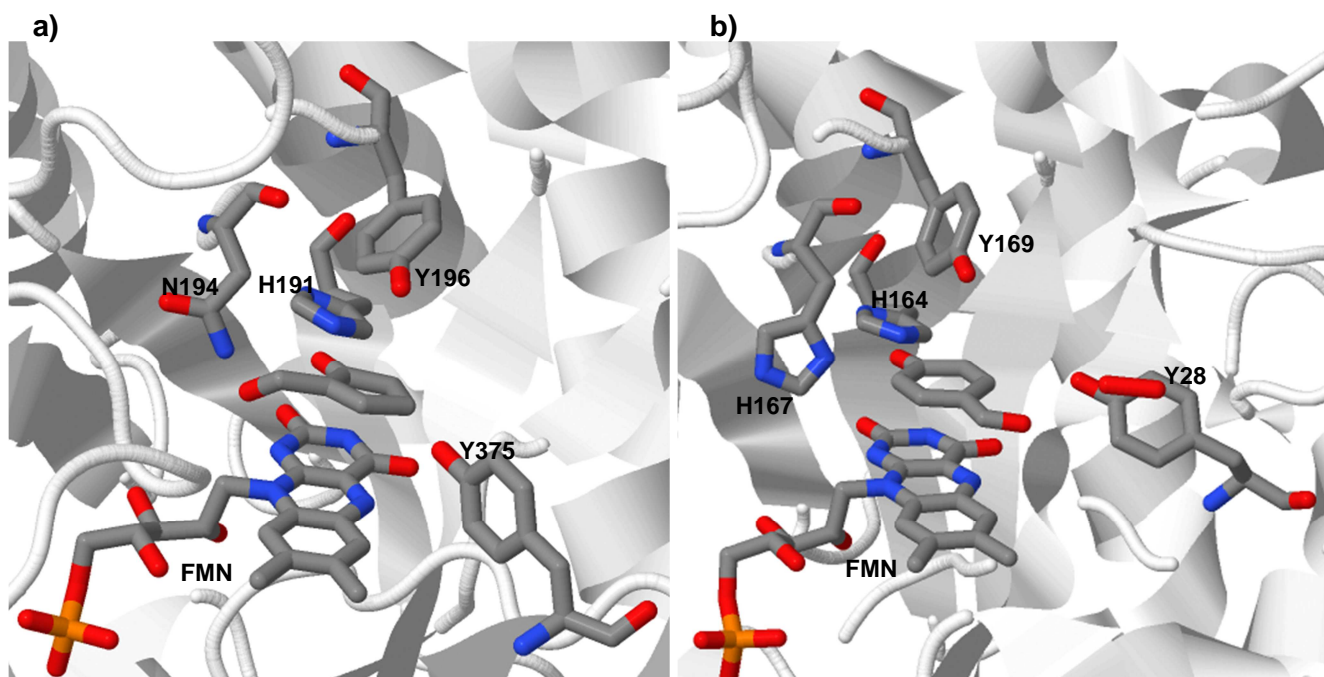
Figure 1.8: Active-site environment of **a)** OYE1 complexed with 2-hydroxymethyl-cyclopent-2-enone and **b)** YqjM complexed with *p*-hydroxybenzaldehyde showing residues important in substrate binding and catalysis (Adapted from Fox & Karplus, 1994; Kitzing *et al.*, 2005).

1.3. CATALYTIC MECHANISM

Members of the OYE family catalyse the reduction of activated alkenes by means of a two-stage mechanism (figure 1.9): a reductive half-reaction whereby the nicotinamide cofactor – NADH or NADPH – is oxidised by hydride transfer to the FMN cofactor of the OYE, followed by an oxidative half-reaction involving hydride transfer from the reduced FMN to the activated C=C of the alkene substrate (Breithaupt *et al.*, 2001; Kohli & Massey, 1998). Kinetic analysis of the NADPH-dependent reduction of quinones and numerous other α/β-unsaturated carbonyl compounds revealed the reduction of the FMN cofactor by NADPH to be the rate-limiting step, while kinetic studies involving the systematic variation of NADPH and oxidant concentrations resulted in parallel Lineweaver-Burk plots (Karplus *et al.*, 1995). As a result, the OYE enzyme family are described as performing the NAD(P)H oxidation and subsequent substrate reduction by means of a bi-bi ping pong mechanism with both reductant and oxidant binding in the same active site, such that the first product, NADP$^+$, leaves the binding pocket before binding of the alkene substrate (Karplus *et al.*, 1995).

R$_1$ = electron-withdrawing group

Y = Tyr
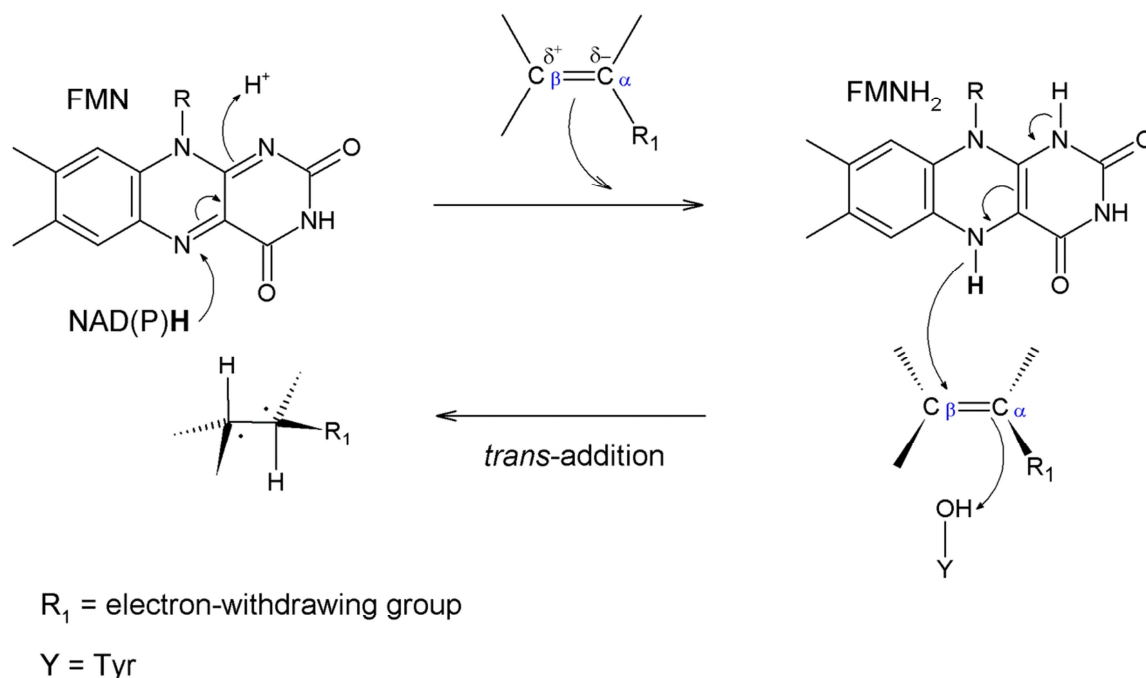
Figure 1.9:   Overall reaction catalysed by members of the OYE family (Adapted from Hall *et al.*, 2006).

1.3.1. The reductive half-reaction

While OYEs can utilise both NADH and NADPH, NADPH is considered to be the physiological substrate of the reductive half-reaction, as NADH appears to reduce the enzyme-bound FMN less efficiently (Brigé *et al.*, 2006; Williams *et al.*, 2004). Analysis of NADPH binding by OYE1 (figure 1.10) using stopped-flow spectrophotometry revealed that the oxidised enzyme and NADPH undergo rapid equilibrium binding to form a Michaelis complex which is dependent on the concentration of NADPH (Brown *et al.*, 2002; Karplus *et al.*, 1995). The appearance of long wavelength absorbance bands and a perturbation of the flavin absorption spectrum after initial binding indicated the formation of a charge-transfer complex. Formation of the charge transfer complex is followed by a biphasic reduction of the flavin, which is independent of NADPH concentration. The fast measurable development of the charge transfer complex is presumed to be due to optimal positioning of the pyrimidine ring of the NADPH over the flavin, allowing π-electron overlap and subsequent hydride transfer (Karplus *et al.*, 1995).

a)

OYE FMN$_{ox}$

NADPH

Reductive
half-reaction

OYE FMN$_{red}$

NADP$^+$

b)

$$EFMN + NADPH \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \underset{\substack{\text{Michaelis} \\ \text{complex}}}{EFMN\text{-}NADPH} \underset{k4}{\overset{k_3}{\rightleftharpoons}} \underset{\substack{\text{Charge-transfer} \\ \text{complex}}}{EFMN^\cdot NADPH^\cdot} \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EFMNH_2{}^\cdot NADP^+ \underset{k_8}{\overset{k_7}{\rightleftharpoons}} EFMNH_2 + NADP^+$$
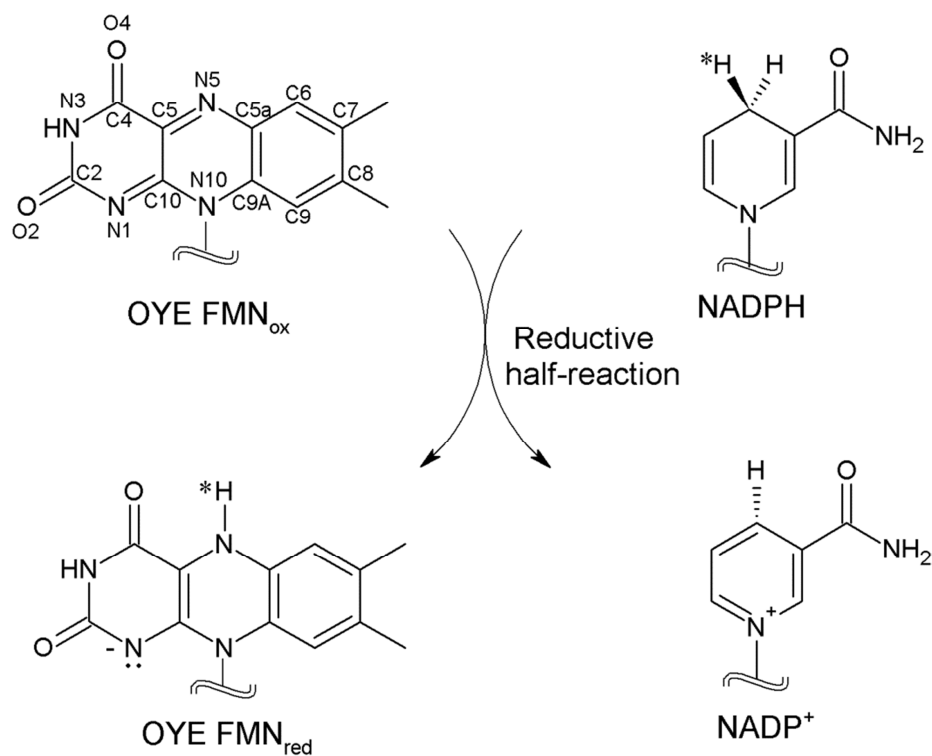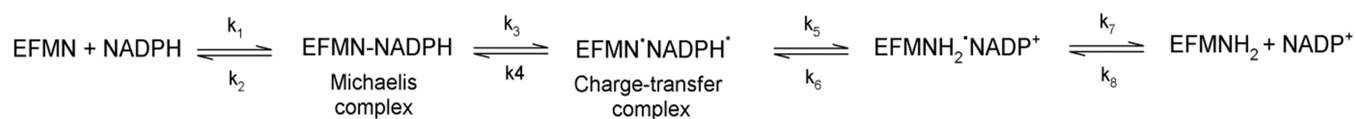
Figure 1.10: **a)** Graphical representation and **b)** reaction mechanism of the reductive reaction catalysed by OYE1 and employing NADPH as nicotinamide cofactor. *H refers to the H$^-$ (pro-*R*) transferred from NADPH to the enzyme-bound FMN during reduction of the enzyme. Reaction mechanism indicates kinetics of the OYE1 catalysed reductive reaction: upon binding of NADPH to the oxidise enzyme, a Michaelis complex is formed (rates k_1 and k_2) in a concentration-dependant manner. This is followed by formation of the charge-transfer complex (k_3 and k_4), after which the FMN undergoes a biphasic reduction (k_5 and k_6). Following the reduction, NADP$^+$ is released (k_7 and k_8) (Adapted from Breithaupt *et al.*, 2001; Karplus *et al.*, 1995; Kohli *et al.*, 1998).

1.3.2. The oxidative half-reaction

The physiological oxidants of the OYE family remain mostly unknown, although a diverse set of compounds have been shown to act as oxidative substrates – from molecular oxygen to methylene blue, ferricyanide and quinines (Vaz *et al.*,1995; Williams & Bruce, 2002; Xu *et al.*, 1999). Although OYE was originally described as an enzyme capable of reducing NADPH in the presence of molecular oxygen, it is unlikely that molecular oxygen is the physiological substrate for the oxidative half reaction (Barna *et al.*, 2001; Karplus *et al.*, 1995; Williams & Bruce 2002) . This is due to the slow reaction rate, along with the formation of undesirable hydrogen peroxide and superoxide. A wide variety of activated alkene substrates have been found to serve as oxidants, some of which will be discussed later in the review. Due to limitations in the development of anaerobic bioreactor systems, most biotransformations by OYE are performed under aerobic conditions (Adalbjörnsson *et al.*, 2010; Fryszkowska *et al.*, 2009; Toogood *et al.*, 2008). Although aerobic biotransformations are effective, a definite reduction in the steady-state reaction rate in the presence of oxygen has been observed, likely due to competition for reduced FMN between molecular oxygen and the activated alkene substrate (Adalbjörnsson *et al.*, 2010).

The oxidation half-reaction results in a net *trans* addition of [2H], which occurs with absolute stereospecificity (figure 1.11) (Hall *et al.*, 2006; Karplus *et al.*, 1995; Kohli *et al.*, 1998). The reaction proceeds via hydride transfer from the N5 atom of the reduced flavin to Cβ of the substrate, followed by transferral of a solvent-derived proton to the substrate's Cα. The reaction is similar to an asymmetric conjugate Michael-addition of hydride onto an activated alkene and is thought to normally occur in a concerted manner (Williams & Bruce, 2002). In order to be a suitable oxidative substrate for an OYE, a compound must contain an electron-withdrawing (or activating) substituent (Stuermer *et al.*, 2007). These substituents are typically aldehyde, ketone, nitro, carboxylic acid, ester, anhydride, lactone or imide groups. The importance of these activating groups is demonstrated by the reduction of the ketone-activated substrate 2-cyclohexenone by PETN reductase (figure 1.13) (Barna *et al.*, 2001; Breithaupt *et al.*, 2001). The carbonyl oxygen of the substrate hydrogen bonds with the previously-mentioned His-pair (His181 and His184) in the active site, resulting in polarization of the olefinic bond and consequent activation of the Cβ for nucleophlic attack by the reduced flavin. In cases, such as OYE1, where the His-pair is substituted instead for a His/Asn pair, the mechanisms remain comparable: the ND2 atom of the asparagine in OYE1 appears to be equivalent in position to the ND1 atom of the later histidine residue of PETN reductase, thus the resulting electronic effects on the substrate oxygen are likely to be similar (Padhi *et al.*, 2009). The conserved active site histidine pair (or equivalent residues) thus serves an activating function in addition to the function of positioning the Cβ of the substrate at a suitable

distance and orientation to the N5 atom of the flavin. The source of the proton donated to the Cα of the substrate during the reduction half reaction appears to differ between OYEs. As discussed in 1.2.1.4., crystallographic and mutagenesis studies in OYE1 demonstrated the role of residue Y196 to act as an active-site acid (Kohli & Massey, 1998). However, the absence of the equivalent tyrosine residue in the OYE morphinone reductase indicates that the residue is not essential for catalysis in all OYEs (Messiha *et al.*, 2005). In addition, mutagenesis studies targeting the equivalent tyrosine residues in PETN reductase and an enoate reductase from *Thermus scotoductus* SA-01 had a small effect on activity of the enzymes (Khan *et al.*, 2005; Opperman *et al*, 2010). In morphinone reductase, a cysteine residue (C191) replaces Y196 of OYE1 and mutagenesis studies have determined that C191 does not act as a crucial acid in the mechanism of the reduction of the olefinic bonds of 2-cyclohexenone and codeinone (Barna *et al.*, 2002). Further investigation of the reaction mechanism of morphinone reductase revealed that the proton necessary for the oxidative half-reaction may even be derived from water present in the solvent (Messiha *et al.*, 2005).

The kinetics of the reduction half-reaction is demonstrated by the reduction of codeinone by the reduced form of the OYE morphinone reductase (figure 1.12) (Craig *et al.,* 1998). The reduction reaction proceeds in three kinetically resolvable steps, with four identifiable intermediate species. The first step involves the dihydroflavin form of the flavin (the first species, formed as a result of the now-complete oxidative half-reaction), which is observed along with a small component of oxidized FMN. A reduced enzyme-codeinone charge-transfer complex is formed, with the rate of formation dependant on codeinone concentration, resulting in the presence of the second species – the two-electron reduced FMN-codeinone charge transfer intermediate. The second step represents the reoxidation of the flavin, which is independent of codeinone concentration, and is accountable for the formation of the third species – the oxidized FMN-hydrocodone complex. The final species observed is likely the release of hydrocodone from the oxidized enzyme. This scheme differs from the observed reduction half-reaction when OYE1 reduces 2-cyclohexenone, during which the substrate-reduced FMN charge transfer step is absent (Massey & Schopfer, 1986). Instead, it is understood that an enzyme-2-cyclohexenone complex was formed by rapid equilibrium binding, after which electron transfer occurred (Massey & Schopfer, 1986).
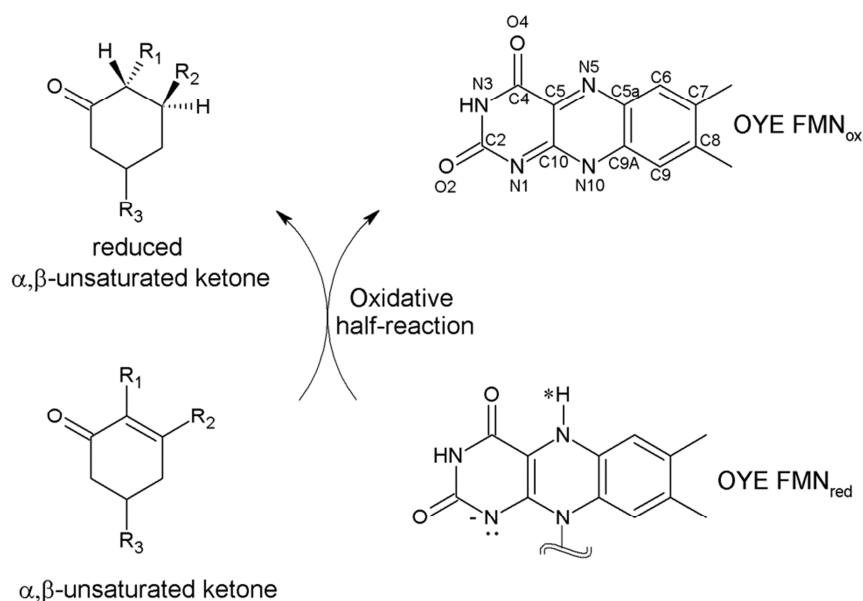
Figure 1.11: Graphical representation of the oxidative half- reaction catalysed by OYE1 and cofactor. *H refers to the H⁻ transferred from NADPH to the enzyme-bound FMN during reduction of the enzyme and which is subsequently involved in the re-oxidation of the enzyme (Adapted from Breithaupt *et al.*, 2001; Karplus *et al.*, 1995; Kohli *et al.*, 1998).
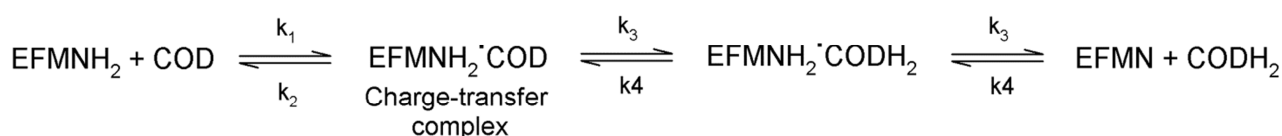
Figure 1.12: Reaction mechanism indicating kinetics oxidative half-reaction catalysed by morphinone reductase with codeinone (COD) as oxidative substrate: firstly occurs formation of a two-electron reduced FMN-codeinone charge transfer intermediate (rates k_1 and k_2). This is followed by formation of the oxidised FMN-hydrocodone complex (k_3 and k_4). Finally, hydrocodone is released from the oxidized enzyme (k_5 and k_6) (Adapted from Craig *et al.,* 1998).
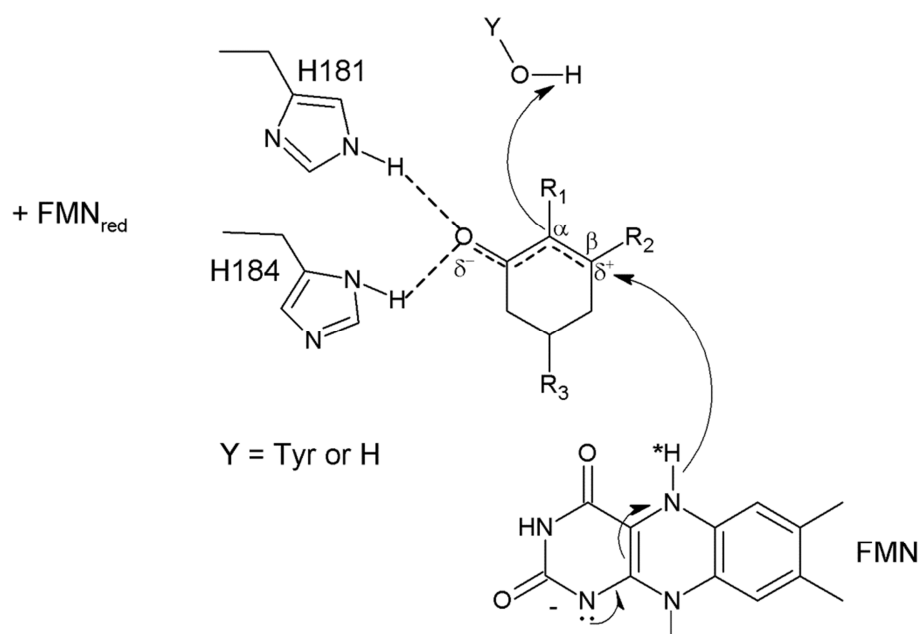
Figure 1.13: Reduction of an activated alkene by PETN reducase. *H refers to the FMN-derived H⁻ (Adapted from Fryszkovska *et al.*, 2009).

1.4. VIABILITY OF OYEs AS BIOCATALYSTS

The use of OYEs in the asymmetric reduction of C=C for the synthesis of chiral compounds is an increasingly efficient and popular biocatalytic alternative to conventional use of (transition-)metal catalysts (Oberdorfer *et al.*, 2011). However, one of the problems associated with these biocatalytic alternatives is the requirement of external NAD(P)H cofactor for enzymatic reduction (Stuermer *et al.*, 2007).

To avoid the requirement of external cofactor recycling (as well as the tedious process of protein purification), many asymmetric bioreductions of activated C=C bonds have been performed using whole cells, most often fermenting baker's yeast (not surprising, considering the isolation of the first OYE from brewer's bottom yeast) (Stuermer *et al.*, 2007). Yet even this solution poses problems. Biotransformations of α,β-unsaturated carboxylic acids using *Clostridium* spp., *Proteus mirabilis, Enterobacter glomerians* and *Acetobacterium woodii* yielded impressive stereoselectivities, but the practical application of these oxygen-sensitive organisms proved difficult due to limitations in the industrial-scale growth of anaerobes (Stueckler *et al.*, 2007). In addition, chemoselectivity of whole-cell bioreductions is often poor, mostly due to the competing C=C reduction reaction and conversion of the activating carbonyl group to the equivalent alcohol by alcohol dehydrogenases (Hall *et al.*, 2006; Müller *et al.*, 2007; Stuermer *et al.*, 2007). Since enoate reductases and alcohol dehydrogenases depend on the same nicotinamide cofactor, redox-decoupling of the two enzyme activities is difficult. As a result, the asymmetric bioreduction of conjugated enals and enones to yield the corresponding saturated aldehyde or ketones often occurs via low chemoselectivity due to the presence of side reactions (Hall *et al.*, 2006; Müller *et al.*, 2007; Stuermer *et al.*, 2007).

It has been reported that the relative rates of the desired C=C reduction versus the C=O reduction depends on the substrate and the enzymes involved (Hall *et al.*, 2006; Müller *et al.*, 2007). Generally, aldehydes are quickly reduced to the corresponding *prim*-alcohol. This can occur via formation of an achiral allylic alcohol (which is not an OYE substrate) and subsequent depletion of substrate, or via over-reduction of the saturated aldehyde formed by the OYE to yield the saturated *prim*-alcohol as final product (Stuermer *et al.*, 2007). With conjugated enones, on the other hand, the competing *sec*-alcohol dehydrogenase activities are less dominant and formation of the *sec*-alcohol is often in the same order of magnitude as the enone reduction. Whole cell biotransformations of (*E/Z*)-citral by organisms such as *Candida parapsilosis* DMS 70125 serve as an example of competing side-reactions influencing yields of the desired product (figure 1.14) (Hall *et al.*, 2006). Relative yields of (*R*)-citronellal and (*R*)-citronellol were dependent on the relative activites of the enoate reductase and the *prim*-alcohol dehydrogenase. Other competing activities

that occur include the loss of acetaldehyde to form sulcatone through citral lyase activity, as well as the generation of geranic or neric acid through the oxidation of the aldehyde moiety (Stuermer *et al.*, 2007).

Despite the above-mentioned and other challenges (such as low *ee* values due to the presence of more than one OYE-like enzyme with different enantiopreferences), numerous whole-cell biotransformations are successfully implemented (a few being described below), often as a result of strict control of reaction conditions (Toogood *et al.*, 2010). However, these problems can be overcome by the use of purified enzymes in conjunction with large quantities of NAD(P)H or, preferably, a suitable NAD(P)H cofactor recycling system to overcome the limitation of high-cost large-scale usage of NAD(P)H.
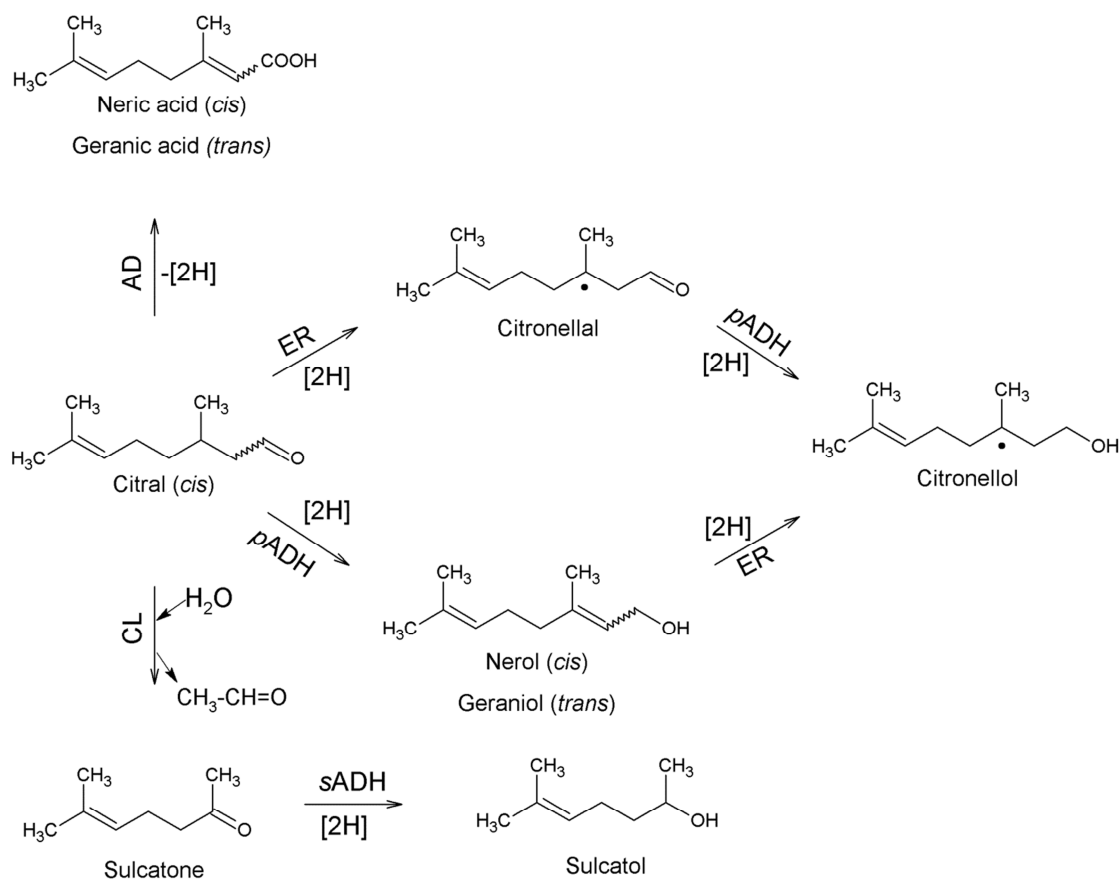
Figure 1.14: Schematic representation of *Candida parapsilosis* whole-cell transformation of (*cis/trans*)-citral, indicating competing enzymatic reactions catalysed by enoate reductase (ER), *prim*-alcohol dehydrogenase (*p*ADH), citral lyase (CL), *sec*-alcohol dehydrogenase (*s*ADH) an aldehyde dehydrogenase (AD). Introduction of new chiral centres is indicated by '*' (Adapted from Hall *et al.*, 2006).

1.4.1. Sources of cofactors for biotransformations with purified OYE

A variety of cofactor regeneration systems exist. These systems offer a cost-effective alternative, as a small amount of $NAD(P)^+$ is required, while the reductive substrate for the recycling enzyme is often inexpensive (Fryszkowska *et al.*, 2009; Hollmann *et al.*, 2010). The first example is cofactor recycling using the enzyme glucose-6-phosphate dehydrogenase (G6PDH). In this system, G6PDH catalyses the oxidation of D-glucose-6-phosphate to D-glucono-1,5-lactone-6-phosphate at the expense of $NADP^+$, thus reducing $NADP^+$ to form NADPH that is used by OYE. Since OYEs appear to have a preference of NADPH over NADH, the G6PDH system would be expected to be the preferred system. However, the presence of D-glucose-6-phosphate and D-glucono-1,5-lactone-6-phosphate in large quantities may cause a decrease in pH, thus affecting both the OYE-catalysed and the cofactor regenerating enzyme reactions. An alternative system (NAD^+/GDH) employs the enzyme glucose dehydrogenase in a NAD^+-specific cofactor recycling system, whereby glucose dehydrogenase oxidises glucose at the expense of NAD^+ to produce NADH (Mueller *et al.*, 2010).

It has been shown that the source of reducing equivalents can affect the yield and enantiopurity of products and can even affect the likelihood of side-reactions occurring (Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008b). For example, PETN reductase catalysed reduction of 2-methylcyclopentenone saw an improvement of *ee* by 6% when an $NADP^+$/G6PDH co-factor recycling system was used (Fryszkowska *et al.*, 2009). In addition, implimentation of this recycling system improved the *ee* of the reduction of ketoisophorone by PETN reductase by 36%. Similarly, reduction of ketoisophorone by OPR1 saw an improvement of *ee* by 40% upon implimentation of the $NADP^+$/G6PDH recycling system (compared to the use of NADH or NADPH alone), while implimentation of a recycling system using formate dehydrogenase (NAD^+/FDH) provided no improvement in *ee* (Hall *et al.*, 2007). In another example from this study, switching from use of $NAD(P)H$ or NAD^+/FDH to the G6PDH recycling system resulted in a 33-44% increase in yield in the reduction of ketoisophorone by OPR3. However, yields in the reduction of citral by OPR1 decreased from >99% to less than 20% when use of $NAD(P)H$ was substituted by a recycling system. This reduction of citral by OPR1 also demonstrated the effect of a cofactor-recycling system on the efficiency of side-reactions – implimentation of the NAD^+/FDH recycling system resulted in the reduction of the carbonyl group becoming the predominant reaction, yielding the allylic alcohol at >95% yield. The allylic alcohol product was not detected significantly in reactions where $NAD(P)H$ or the NAD^+/GDH or $NADP^+$/G6PDH recycling systems were used (Hall *et al.*, 2007). However, comparative studies involving OYE1 confirmed that the effects of the different cofactor sources varies among OYEs (Hall *et al.*, 2008a). While the reduction of citral to the (S)-enantiomeric product by OYE1 resulted in poor *ee* values (15-20%) in the presence of $NAD(P)H$ or

the NADP$^+$/G6PDH recycling system, implimentation of the NAD$^+$/GDH saw the production of the (*R*)-enantiomeric product at improved *ee* of 77% (Hall *et al.*, 2008a).

Recently, an alternative to the conventional cofactor-recycling systems was proposed (figure 1.15) – an approach that employs free FMN, EDTA and white light as a source of reducing equivalents (Grau *et al.*, 2009; Taglieber *et al.*, 2008). By directly regenerating the FMN of the OYE, the undesired carbonyl reduction of substrates and products by NAD(P)H-dependent ketoreductases (present in both whole-cell and (contaminated) isolated protein biotransformations) is avoided. In addition, this exploitation of the photoexcitability of flavins to directly regenerate the OYE FMN circumvents the costly use of unstable nicotanamide cofactors. The efficacy of the light-driven approach was demonstrated when photoenzymatic reductions of citral and ketoisophorone were performed using crude *E. coli* extracts containing heterologously expressed YqjM, and EDTA as the sacrificial electron donor (cosubstrate) (Grau *et al.*, 2009). The full conversion of citral to citronellal showed that undesired ketoreductase activity is successfully inhibited even in crude cell extracts. In terms of enantioselectivity, the photoenzymatic reduction of ketoisophorone resulted in production of (*R*)-levodione at 69% *ee*. The observed *ee* is 30% lower than that obtained during reduction of ketoisophorone by isolated YqjM in conjunction with the NADP$^+$/G6PDH recycling system, but 22-32% higher than in conjunction with the NAD$^+$/FDH recycling system of NAD(P)H (Grau *et al.*, 2009; Hall *et al.*, 2008b). However, the phytoenzymatic conversion of ketoisophorone by the crude enzyme extract resulted in 72% *ee* when using stoichiometric NAD(P)H (Grau *et al.*, 2009).

Numerous variables affected the productivity of the photoenzymatic conversions (Grau *et al.*, 2009). Deaeration of reaction mixtures greatly increased conversion compared to the trace amounts of products detected after reactions under aerobic conditions. Low conversion under aerobic conditions was attributed to the formation of hydrogen peroxide by highly-reactive reduced flavin in the presence of molecular oxygen. In addition, use of an intensive light source (250 W) compared to a conventional 40 W lamp resulted in 50% increase in productivity of the system. Use of alternative electron donors (such as NaHCO$_2$, methionine and nicotine) resulted in lower yields compared to reactions performed with EDTA, as well as significantly poorer *ee* values. While different concentrations of sacrificial electron donor did not influence the conversions significantly, concentration of the FMN influenced both yield and enantioselectivity. Initial reaction rates correlated to FMN concentration (indicating that the formation of FMNH$_2$ is the rate-limiting step). Surprisingly, the enantioselectivity of the reductions also increased with increasing FMN concentration, while inversion of the enantioselectivity occurred at FMN concentrations below 10 µM. This inversion of enantioselectivity was suspected to be due to activity of an endogeneous *E. coli* NemA, an enoate reductase with possibly higher affinity towards (or faster electron transfer rates with) FMN compared to YqjM (Grau *et al.*, 2009).
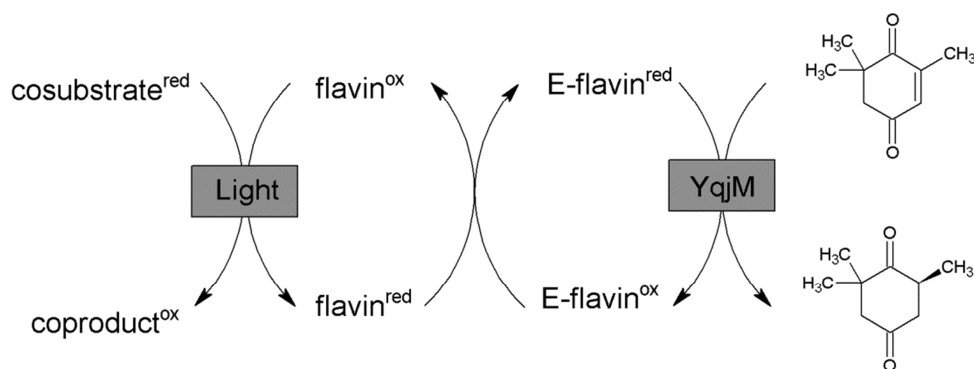
Figure 1.15: Photoenzymatic reduction of ketoisophorone to (*R*)-levodione by YqjM. E-flavin refers to enzyme-bound flavin. (Adapted from Grau *et al.*, 2009; Taglieber *et al.*, 2008).

1.4.2. Examples of asymmetric bioreductions catalysed by OYEs

The utilization of biological catalysts in asymmetric synthetic processes is becoming an increasingly popular route in the production of pharmaceuticals and other fine chemicals, due to the often high efficiency and regio-, stereo- and enanteoselectivity (Toogood *et al.*, 2010). The OYE family of biocatalysts are capable of reducing an extensive library of α,β-unsaturated activated alkene compounds that are of commercial importance, many of which will be discussed in the following section.

1.4.2.1. *α,β-unsaturated carbonyl compounds*

a) Aldehydes

Many acyclic α,β-unsaturated alkyl aldehydes serve as substrates for OYEs (figure 1.16) and a variety of biotransformations employing enals have shed light on the substrate preferences of these OYEs (Fryszkowska *et al.*, 2009). Both R_1 and R_2 substituted acyclic enone compounds tend to act as substrates, as long as only one of the substituents is large (Stuermer *et al.*, 2007). Substituents on the Cα (R_3) must be smaller, but may be electron-withdrawing (such as halogens or azides) (Stuermer *et al.*, 2007).

One almost ubiquitous OYE reaction - catalysed by (among other OYEs) PETN reductase, OYE from *Thermonanaerobacter pseudethanolicus* E39, OYE1-3, and NCR from *Zymonas mobilis* - is the reduction of 2-methylpent-2-enal (figure 1.16b), albeit at poor *ee* (Adalbjörnsson *et al.,* 2010; Fryszkowska *et al.*, 2009; Müller *et al.*, 2007). PETN reductase serves as an example, producing (*S*)-2-methylpentanal with >99% conversion (Fryszkowska *et al.*, 2009). In this case, *ee* was poor at 66%, possibly due to unidentified side-products formed by non-enzymatic substrate, or product decomposition and product racemisation. Longer chain acyclic enals such as *trans*-2-hexenal have also been shown to act as substrates for a variety of OYEs. *Le*OPR, an OYE from *Lycopersicon esculentum* (tomato), exhibits the highest substrate specificity for *trans*-2-dodecenal and a 23-fold decrease in specificity for *trans*-2-hexenal (Straßner *et al.*, 1999). This substrate preference is thought to be attributed to the larger size and hydrophobicity of the Cβ substituent (Straßner *et al.*, 1999).

The reduction of cinnamaldehyde, a naturally occurring aldehyde found in essential oil, is catalysed by a large number of OYEs, including xenobiotic reductase (XenA) from *Pseudomonas putida*, KYE1 from *Kluyveromyces lactis* (an OYE with high homology to OYE1) and Yers-Er from *Yersinia bercovieri*, an OYE sharing high sequence homology with PETN reductase (Chaparro-Riggers *et al.*, 2007). The reduction of α-substituted cinnamaldehydes by baker's yeast (figure 1.16c) was demonstrated as an essential step in the production of L-phenylalaninol, an important intermediate in the synthesis of some HIV-1-protease inhibitors (Fardelone *et al.*, 2004; Fronza *et al.*, 2009). Bioreduction of α-bromo-, α-azido- and α-methyl-cinnamaldehyde by baker's yeast yielded (*S*)-2-bromo-, (*S*)-2-azido and (*S*)-2-methyl-3-phenyl-1-propanol with high enantioselectivity and high yields (Fardelone *et al.*, 2004; Fronza *et al.*, 2009).
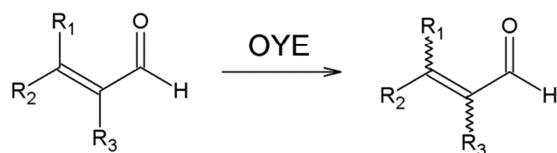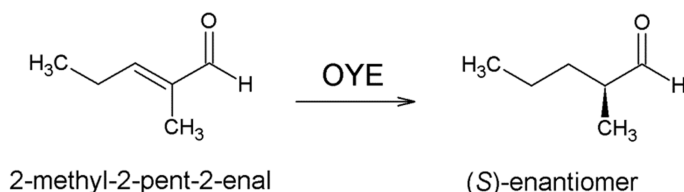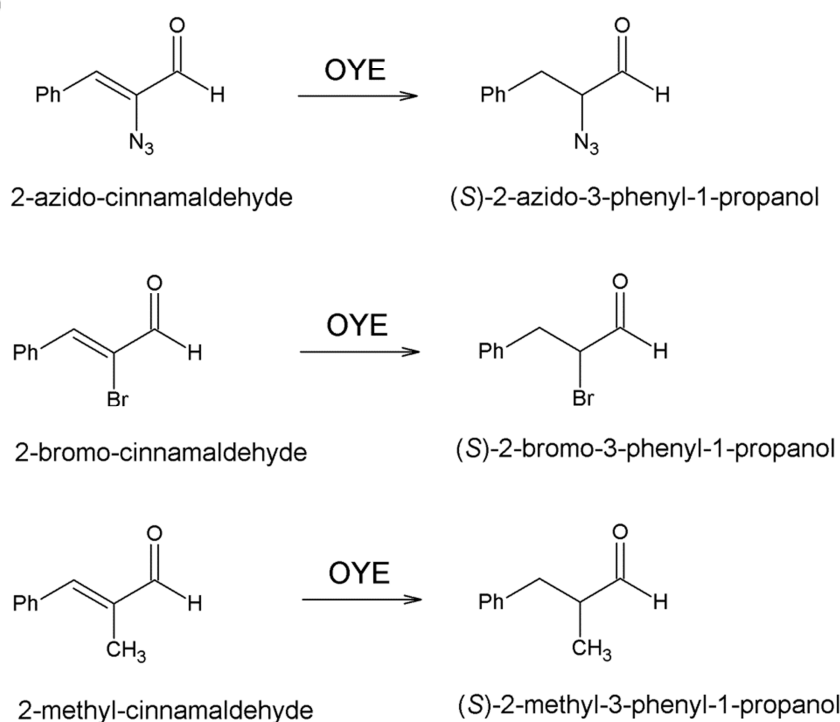
Figure 1.16: **a)** Examples of asymmetric bioreductions attributed to OYEs where α,β-unsaturated alkenes with aldehye activating groups act as substrates. **b)** Reduction of 2-methylpent-2-enal by PETN reductase, OYE from *Thermonanaerobacter pseudethanolicus* E39, OYE1-3 from brewer's bottom yeast and baker's yeast, as well as NCR from *Zymonas mobilis.* **c)** Reduction of α-substitued cinnamaldehydes by baker's yeast (Adapted from Fardelone *et al.,* 2004; Fryszkowska *et al.*, 2009; Stuermer *et al*, 2007).

b) Ketones

Ketones play an important role in industrial processes, acting as solvents or precursors in the synthesis of polymers and pharmaceuticals (Toogood *et al.*, 2010). A rather wide variety of ketones act as substrates for members of the OYE family (figure 1.17). For whole-cell biotransformations with baker's yeast, substrate preferences for acyclic enones (especially in terms of substituent requirements) appear to be similar to that described for aldehydes (Kawai *et al.*, 2001; Stuermer *et al.*, 2007). Stereoselectivity improves drastically with elongation of α-substituents, while in cases where a β-phenyl ring is present, stereoselectivity is also drastically influenced by substituents on the ring structure. Enones having no substituent or a *para*-substituent present on the β-phenyl ring are reduced to the corresponding saturated (*S*)-ketones with moderate stereoselectivity, while stereoselectivity is drastically improved by the presence of a *meta*-substituent, regardless of its type (Kawai *et al.*, 2001; Stuermer *et al.*, 2007).

The reduction of 2-cyclohexenone to cyclohexanone is a "classical" and ubiquitous reaction catalysed by members of the OYE family (Adalbjörnsson *et al.,* 2010; Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.*, 2008b; Mueller *et al.*, 2010) . Cyclohexanone is an important ketone produced industrially to a large scale, thus providing an opportunity for OYEs to be used in the industrial process. While unsubstituted 2-cyclohexenone appears to be a ubiquitous OYE substrate, alkyl substitutions at the α- and β- carbons seem to have a significant impact on the substrate binding, activity and enantioselectivity. For example, *Beauveria sulfurescens* has been shown to reduce α-alkyl substituted unsaturated cyclohexenones to produce saturated ketones at high yields (90%) and high enantioselectivity, while α,β-alkyl-unsaturated cyclohexenones are not reduced (Dauphin *et al.*, 1992).

The α- and β-substituents of cyclic enone substrates tend to be limited to small chain alkyl groups (Swiderska & Stewart, 2006), which is likely due to steric hindrances within the enzyme active site. An example can be found in the reduction of cyclic enones by the cyanobacterium *Synechoccus* sp. PCC 7942 (Shimoda *et al.*, 2004). Under illumination, the cyanobacterium is capable of reducing the endocyclic C=C bond of (*S*)-*trans*-enones and the exocyclic C=C bond of (*S*)-*cis*-enones to the corresponding (*S*)-ketones with high enantioselectivity (and with minor formation of saturated alcohols). As far as endocyclic C=C bonds are concerned, reduction of a range of 2-cyclohexenone-type compounds by *Synechoccus* sp. PCC 7942 revealed a preference for substrates with a H at the Cβ and small substituents at the Cα. Substrates with exocyclic C=C were converted to the equivalent (*S*)-ketone at comparably lower yields and reasonable enantioselectivity. This cyanobacterium is not alone in displaying selectivity between reduction of

endocyclic and exocyclic C=C double bonds. Two enoate reductases from *Nicotiana tabacum* have been described: (*S*)-*trans*-enoate reductase which reduces the endocyclic C=C bonds of enones with a hydrogen substituent at the position β to the carbonyl, and (*S*)-*cis*-enoate reductase which reduces the exocyclic C=C bonds of enones (Tang & Suga, 1992).

While there appear to be a few general rules among the OYEs regarding substrate preferences and enantioselectivity with respect to substituted cyclohexenone and cyclopentenone compounds, yields, reaction rates and optical purity of products vary among OYEs (Adalbjörnsson *et al.,* 2010; Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.*, 2008b; Mueller *et al.*, 2010). These variations are often influenced by the type of cofactor and cofactor source. Generally, the reduction of 3-methyl substituted cyclopentenone and 3-methyl substituted cyclohexenone results in the formation of the equivalent (*S*)-enantiomeric product at high optical purity but at lower yields (Mueller *et al.*, 2010; Toogood *et al.*, 2010). This was demonstrated by the reduction of 3-methyl-2-cylohexenone and 3-methyl-2-cyclopentenone by crude extracts of the OYE EBP1 (estrogen binding protein) (Mueller *et al.*, 2010), reductions that the OYEs MR (Mueller *et al.*, 2010), OPR1 (Hall *et al.*, 2008b) and YqjM failed to catalyse. EBP1 reduced 3-methyl-2-cyclohexenone to the corresponding (*S*)-product at >99% yield and 96% *ee*, while 3-methyl-2-cyclopentenone was reduced at 76% conversion at >99% *ee* (Mueller *et al.*, 2010). In contrast, the reduction of 2-methyl substituted cyclopentenone and cyclohexenone (figure 1.17c) generally proceeds with higher yields, but often at poorer product optical purity (Adalbjörnsson *et al.,* 2010; Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.*, 2008b; Mueller *et al.*, 2010). For example, EBP1 reduces the 2-methyl substituted compounds at 99% yield, but produces racemic mixtures of each corresponding product (Mueller *et al.*, 2010).

While the cofactors and cofactor regeneration systems employed affect the enantioselectivity, poor *ee* values are often attributed to the tendency of substrates containing Cα substituents to undergo nonenzymatic racemisation (Adalbjörnsson *et al.,* 2010). In addition, the ability of OYEs to bind these substrates in multiple binding modes may also contribute to the lower optical purity of the enzymes (Fryszkowska *et al.*, 2009). When more bulky 2- and 3-alkyl substituents are present on 2-cyclohexenones, the size and position of the substituent greatly affected product yield, but not the enantiospecificity of the optical purity of the product during whole-cell biotransformations, as shown with OYE1 from *S. pastorianus* (Swiderska & Stewart, 2006). When OYE1 from *S. pastorianus* was expressed in *E. coli* and whole-cell biotransformations performed, substrates with bulky substituents at the Cα resulted in poor conversion.   These whole cells converted only 16% of 2-ethyl-2-cyclohexenone to product, compared to 100% conversion observed for 2-methyl-2-cyclohexenone. More bulky substituents at the Cβ proved less problematic, with ethyl-, *n*-propyl and *iso*-propyl

substituents resulting in higher conversion rates. Despite differences in yields, products of high optical purities were obtained for all the 2- and 3-alkyl substituted 2-cyclohexenones (Swiderska & Stewart, 2006).

While these variations in yields and enantioselectivities for 3-alkyl-substituted 2-cyclohexenones and 2-cyclopentenones are problematic for the industry, mutagenesis studies with the OYE YqjM have revealed that small numbers of specific mutations can greatly improve the substrate tolerance of the enzyme, albeit with varying enantioselectivities (Bougioukou *et al.*, 2009). Twenty residues were targeted and mutated using iterative saturation mutagenesis, generating a library of less than 5000 mutants which were subsequently screened by performing whole-cell biotransformations. One of the most promising mutants, C26D YqjM, catalysed the reduction of 3-methyl-2-cyclopentenone to the (*S*)-enantiomeric product, a reaction not catalysed significantly by the wild type. In another mutant, C26G, the mutation caused a switch in enantioselectivity, producing the (*S*)-enantiomric product (compared to the (*R*)-product formed by the wild type) during the reduction of 3-methylester substituted 2-cyclopentenone at 100% yield and 93% *ee* (Bougioukou *et al.*, 2009).

a)
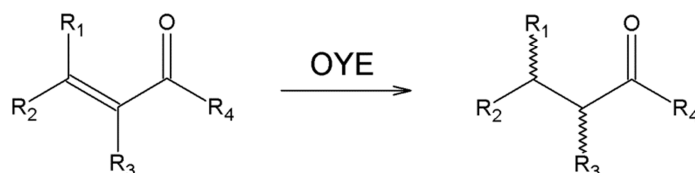
R₁ = large where R₂ = small (or vice versa); R₃ = not too large

R₁, R₂ = H, alkyl, haloalkyl, (hetero)aryl, arylalkenyl

R₃ = H, CH₃, CF₃, F, Cl, Br, N₃

R₄ = CH₃, CH₂CH₃

b)
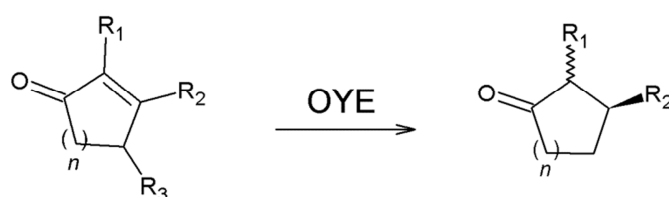
R₁, R₂ = H, CH, OAc, fluoroaryl, short-chain alkyl

R₃ = H, O(sp²), (CH₂)₇COOH, alkyl, short-chain alkyl

$n = 1, 2$

c)

(*R*)-enantiomer

2-methyl-2-cyclohexen-1-one

(*S*)-enantiomer
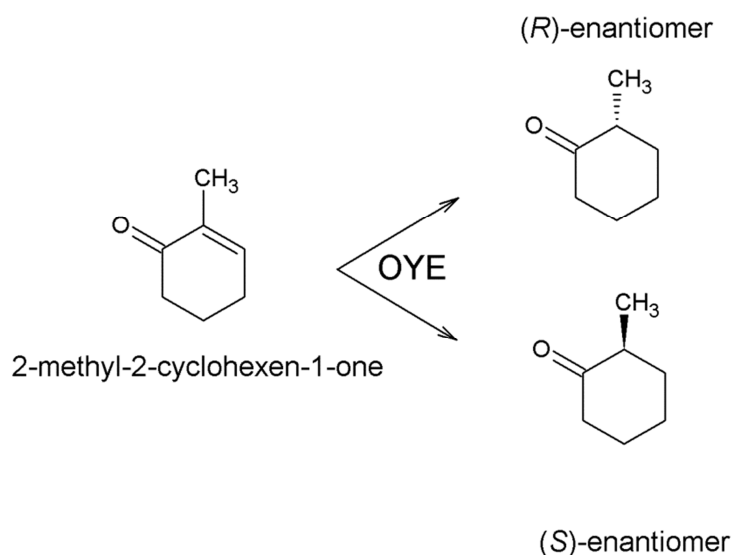
Figure 1.17: Examples of asymmetric bioreductions attributed to OYEs where **a)** acyclic and **b)** cyclic α,β-unsaturated alkenes with ketone activating groups act as substrates. **c)** Reduction of 2-methyl-2-cyclohexenone. (Adapted from Adalbjörnsson *et al.*, 2010; Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.*, 2008b; Mueller *et al.*, 2010; Stuermer *et al*, 2007; Toogood *et al.*, 2010).

34

The reduction of ketoisophorone to levodione is an industrially important reaction, as leviodone is used as a building block in the production of carotanoids, such as zeazanthin, an important feed additive for salmon and trout (Hall *et al.*, 2008a; Hall *et al.*, 2008b). Whole-cell asymmetric bioreduction of ketoisophorone on a large scale has been achieved with the production of (*R*)-levodione using baker's yeast (Stuermer *et al.,* 2007). As is often the case with whole-cell bioreductions, the process was plagued by poor chemoselectivity with reduction of both C=C and C=O bonds occurring due to the presence of both enoate reductases and competing alcohol dehydrogenases (figure 1.18). Careful reaction control (involving removal of the desired product from the bioreactor via an external crystallization loop) was employed instead to produce (*R*)-levodione as the main product at >80% yield, with the undesired *sec*-alcohol dehydrogenase activity kept at a minimum (Buque-Taboada *et al.*, 2005).

Another solution involves the cloning of a yeast OYE into *E. coli* and employing a NAD$^+$/glucose dehydrogenase cofactor recycling system, for the regeneration of NADH, by coexpression of a suitable glucose dehydrogenase in the same host (Kataoka *et al.*, 2004; Wada *et al.*, 2003). Employment of OYE2 from *S. cerevisiae* in this method resulted in 96% conversion at 95.5% *ee*, while similar use of an OYE from *Candida macedoniensis* resulted in a 100-fold reactivity increase compared to the original reduction using baker's yeast. In both cases, the competing carbonyl reduction reaction was eliminated completely (Stuermer *et al.*, 2007).

The asymmetric bioreduction of ketoisophorone using isolated OYEs has also been successful. In conjunction with a cofactor recycling system, the OYEs OPR1 and OPR3 from *Lycopersicon esculentum* (tomato) and YqjM from *Bacillus subtilis* achieved conversion of ketoisophorone to (*R*)-levodione in up to 99% *ee*, while reduction with the NCR enoate reductase from *Zymomonas mobilis*, OYE1 from *S. pastorianus* and OYE2-3 from *S. cerevisiae* led to formation of the desired product up to 98% *ee* (Hall *et al.*, 2008a; Hall *et al.*, 2008b). In all cases employing isolated enzymes, OYEs displayed different activities depending on the cofactor and the recycling system employed, with the system employing the recycling enzyme glucose-6-phosphate dehydrogenase resulting in the best conversion an *ee* in most cases. One problem often encountered in the production of (*R*)-levodione from ketoisophoropne is the occurrence of water-mediated product racemization, leading to poor *ee* values (Fryszkowska *et al.*, 2009). Reactions using PETN reductase revealed that product racemization can be prevented through control of reaction conditions – *ee* values greater than 98% were obtained within a 1.5 hour reaction timeframe, when pH ranged between 5.0 and 7.5 and at high enzyme concentrations (Fryszkowska *et al.*, 2009).
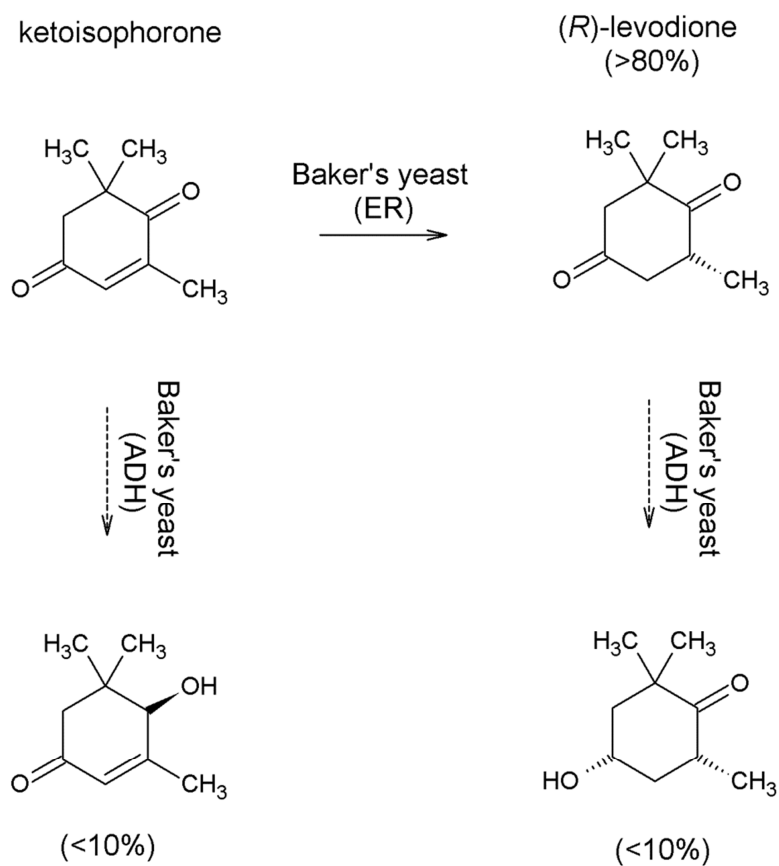
Figure 1.18: Industrial-scale asymmetric bioreduction of ketoisophorone by baker's yeast, showing competing enoate reductase (ER) and alcohol dehydrogenase (ADH) reactions (Adapted from Stuermer *et al.*, 2007).

Ketoisophorone has also been used as substrate in two-step enzymatic asymmetric reduction to yield the doubly chiral compound (4R,6R)-4-hydroxy-2,2,6-trimethylcyclohexanone, or actinol (figure 1.19), another precursor in the synthesis of naturally occurring, optically active hydroxylated carotenoids such as zeaxanthin and cryptoxanthin (Wada *et al.*, 2003). Although this conversion had been achieved using bacterial cells, the whole-cell biotransformations resulted in mixtures of isomers (4R,6S), (4S,6R), (4R,6R) and (4S,6S) being produced, with low productivity. However, employing the isolated enzymes OYE2 from *S. cerevisiae* and levodione reductase from *Corynebacterium aquaticum* M-13 (expressed in *E. coli*) in conjunction with a NAD$^+$/glucose dehydrogenase cofactor regeneration system, proved successful. In the reaction, chirality is first introduced at the C-6 position through the stereoselective enzymatic hydrogenation of the double bond by OYE2. This is followed by selective and stereospecific reduction of the carbonyl group at the C-4 position by levodione reductase to yield the corresponding alcohol. The two-step enzymatic reduction system resulted in the near-stochiometric production of (4R,6R)-4-hydroxy-2,2,6-trimethylcyclohexanone at >94% *ee* (Wada *et al.*, 2003).
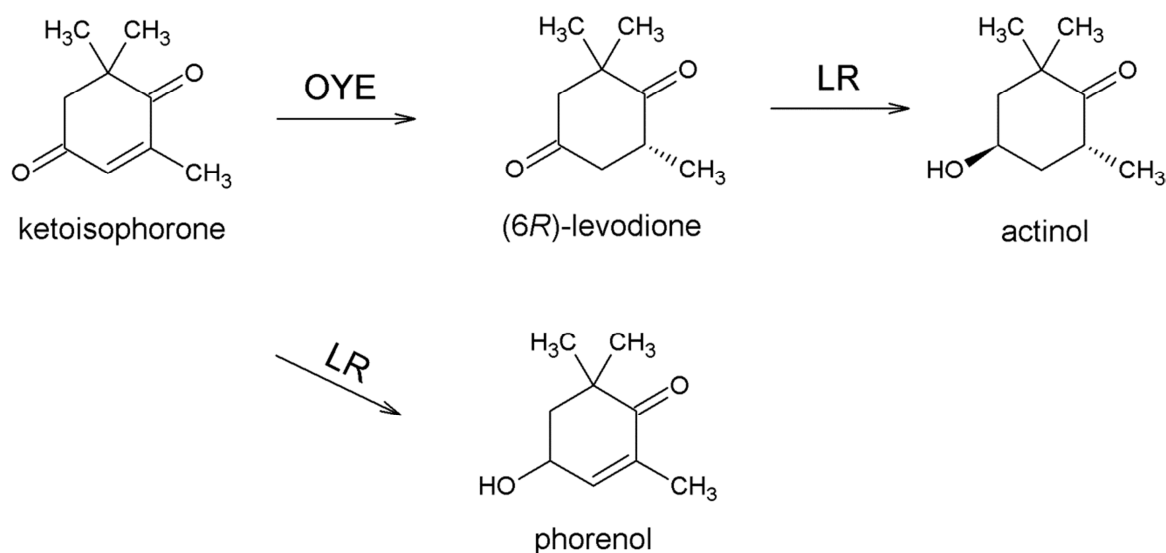
Figure 1.19: Two-step conversion of ketoisophorone to (4R,6R)-actinol using OYE2 (OYE) and levodione reductase (LR). (Adapted from Wada *et al.*, 2003).

c) Maleimides, maleic acid derivatives and coumarins

Maleimides (or unsaturated imides), are important in the synthesis of organic compounds due to high reactivity of the double bond with hydroxy, amine and thiol functional groups to generate stable C-C, C-N and C-S bonds respectively (Toogood *et al.*, 2010). α,β-unsaturated carboxylic acid derivatives (acid anhydrides) and maleimides have been shown to act as good substrates for whole-cell biotransformations and isolated OYEs (figure 1.20) (Adalbjörnsson *et al.*, 2010; Chaparro-Riggers *et al.*, 2007; Hall *et al.*, 2007) . For example, the NADH-dependent enoate reductase (p68 reductase) from *Marchantia polymorpha* (liverwort) catalysis the reduction of citraconic anhydride to the (*R*)-enantiomer (Shimoda & Kubota, 2004). In another example, whole-cell biotransformations using *N. tabacum* cells reduce *N*-phenyl-2-methylmaleimide to generate (*R*)-*N*-phenyl-2-methylsuccinimide with 99% *ee* (Hirata *et al.*, 2004).

Generally, the reduction of *N*-substituted maleimides by isolated OYEs such as OPR1, OPR3, YqjM and PETN reductase produces high yields of the (*R*)-enantiomeric product with high enantiopurity (Adalbjörnsson *et al.,* 2010; Chaparro-Riggers *et al.*, 2007; Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.*, 2008b). Morphinone reductase and purified estrogen binding protein (EBP1) appear to be exceptions, catalysing the reduction of 2-methyl maleimide and *N*-phenyl-2-methylmaleimide at low yields, but high enantioselectivity (Mueller *et al.*, 2010). Studies with PETN reductase have indicated that conversion of *N*-substituted maleimides occur at near-physiological reaction rates (Fryszkowska *et al.*, 2009). These substrate appear to be highly specific for PETN reductase, with tight binding of the maleimides attributed to the interaction between one carbonyl group and the active site His181/184 pair, as well as between the second carbonyl group and Y351. Xenobiotic reductase A (XenA) from *P. putida* IIB has been shown to catalyse the reduction of 8-hydroxycoumarin to produce 8-hydroxy-3,4-dihydrocoumarin and the reduction of coumarin to 3,4-dihydro-coumarin as part of the degradation of quinoline by the organism (Griese *et al.*, 2006). Opening of the lactone ring as a result of the reduction reaction allows 8-hydroxy-3,4-dihydrocoumarin to be further degraded by *P. putida* (Griese *et al.*, 2006).

a)

R_1 = ethyl, aryl, H

R_2 = (=O), H

R_3 = H, CH_3

R = N,O

b)

2-methyl maleimide   (*R*)-enantiomer

c)

citraconic anhydride   (*R*)-enantiomer

d)

coumarin or
8-hydroxycoumarin

3,4-dihydrocoumarin or
8-hydroxy-3,4-dihydrocoumarin

R = H, OH
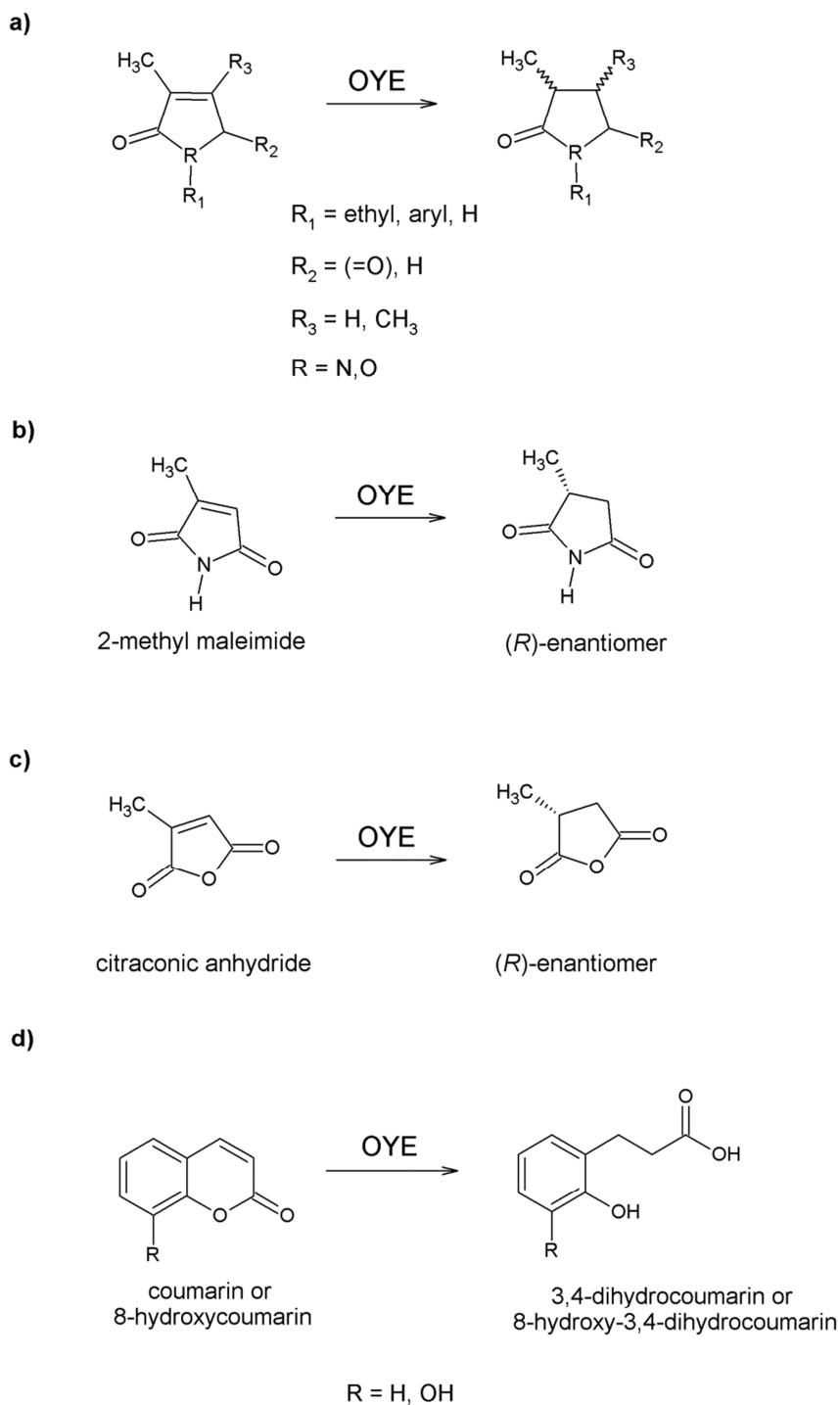
Figure 1.20: Examples of symmetric bioreduction of **a-c)** α,β-unsaturated maleimides and maleic acid derivatives and **d)** coumarins by members of the OYE family. **b)** Reduction of 2-methyl maleimide by OYEs such PETN reductase, OPR1, OPR3 and YqjM. **c)** Reduction of citraconic anhydride by p68 reductase from liverwort. **d)** Reduction of coumarins by the OYE XenA. (Adapted from Griese *et al.*, 2006; Hall *et al.*, 2007; Shimoda & Kubota, 2004; Stuermer *et al.*, 2007).

d) Carboxylic acids and esters

Carboxylic acids are widely used in the industry in synthetic processed involving the production of polymers, solvents, food additives and pharmaceuticals (Stuermer *et al.*, 2007). In addition, carboxylic acids can be used to produce esters, amides and thioesters. *Clostridium* sp. La1 and *Proteus* spp. have been shown to perform biotransformations with a wide variety of carboxylic acid substrates with a broad range of C=C substituents (Eck & Simon, 1994; Simon *et al.*, 1985). Generally (figure 1.21), satisfactory yields and enantioselectivities are observed for substrates with which the R_2 substituent must not be too large (Stuermer *et al.*, 2007). In contrast, a wide variety of R_3 substituents are tolerated. Halogens are well tolerated at the Cα, but not at the Cβ position. Activity was restricted in the case of branching in the Cβ position, with (*E*)-isomers of the substrates generating (*R*)-products and (*Z*)-isomers of the substrates generating (*S*)-products (Kurata *et al.*, 2004; Utaka *et al.*, 1987). With substrates containing two C-C double bonds, only the α,β-unsaturated bond was reduced to yield products with high stereospecificity (Stuermer *et al.*, 2007). Compared to their disubstituted counterparts, α,β-unsaturated monocarboxylic acids and esters appear to be less popular substrates among the known and characterised OYEs. PETN reductase, *N*-ethylmaleimide reductase (NEMR), morphinone reductase and EBP1 (Mueller *et al.*, 2010), as well as OPR1, OPR3 and YqjM (Stueckler *et al.*, 2007) all catalysed the reduction of (*E*)- and (*Z*)-2-methylmaleic acid and their corresponding methyl esters, showing clear preference for the (*Z*)-isomer.
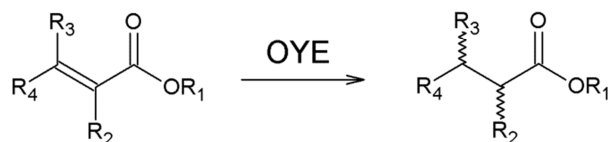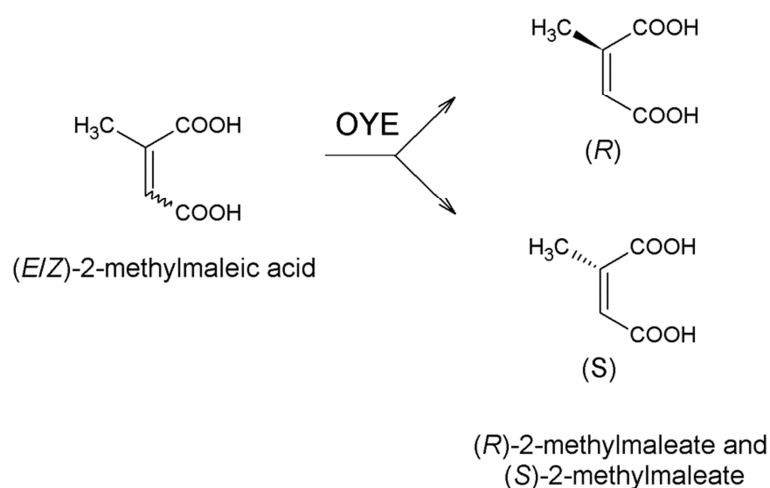
Figure 1.21: **a)** Examples of asymmetric bioreductions attributed to OYEs where α,β-unsaturated alkenes with carboxylic acid or ester activating groups act as substrates. **b)** Reduction of 2-methylmaleic acid by PETN reductase, *N*-ethylmaleimide reductase (NEMR), morphinone reductase, OPR1, OPR3 and YqjM and EBP1 (Adapted from Mueller *et al.*, 2010; Stueckler *et al.*, 2007; Stuermer *et al.*, 2007; Toogood *et al.*, 2010).

1.4.2.2. *Terpenoids*

Terpenoids are naturally occurring cyclic and acyclic molecules, derived from five-carbon isoprene units (Goretti *et al.*, 2009; Toogood *et al.*, 2010). Terpenoids from plants are widely used for their aromatic properties (examples include the terpenoids occurring in eucalyptus oil and the essential oils of ginger). Examples of terpenoids applied extensively in the industry include citral and carvone (*R* and *S*). Carvone is used as a fragrance and flavour, as well as a potato sprouting inhibitor and an antimicrobial agent, among other uses (Goretti *et al.*, 2009).

A recent study screening 146 yeasts led to the discovery of 16 strains (14 species, belonging to the genera *Candida, Cryptococcus, Hansensiaspora, Kluyveromyces, Pichia* and *Saccharomyces*) that are able to biotransform (4*S*)-(+)-carvone (figure 1.22) to (predominantly) (1*S*,4*S*)- and (1*R*,4*S*)-dihydrocarvone (Goretti *et al.*, 2009). Side-products of the reaction included (1*S*,2*S*,4*S*)-, (1*S*,2*R*,4*S*)-, (1*R*,2*S*,4*S*)-, and traces of (1*R*,2*R*,4*S*)-dihydrocarveol. Yields varied considerably, depending on the strain used, as well as the state of the cells (growing, resting, or lyophilized). The formation of side products was attributed to the presence of carbonyl reductase (CR, which catalyses C=O reduction) and was combated by, firstly, selecting strains that exhibit high enoate reductase activity and secondly, optimizing the reaction conditions (with respect to use of solvents, inhibitors and temperature regulation) (Goretti *et al.*, 2009).

Isolated OYEs PETN reductase, OYE1 and the OYE from *Thermonanaerobacter pseudethanolicus* E39 have been shown to reduce (*S*/*R*)-carvone (figure 1.23) to yield (2*R*,5*S*)- and (2*R*,5*R*)-dihydrocarvone (Adalbjörnsson *et al.*, 2010; Fryszkowska *et al.*, 2009; Padhi *et al.*, 2009). The consistency in the stereochemical outcome was attributed to the presence of the bulky substituent that results in a common binding mode for both substrates. Numerous OYEs have shown the ability to reduce citral (neral and geraninal isomers) to *S*-citronellal at high yields and high product enantiopurity (figure 1.23) (Adalbjörnsson *et al.,* 2010; Hall *et al.*, 2007; Hall *et al.*, 2008b). However, PETN reductase serves as an exception (Fryszkowska *et al.*, 2009), forming the expected (*S*)-product at low yields (57%) and reasonable enatiopurity (87% *ee*). Reduction of citral by OYE1 was greatly influenced by the source of cofactor: in the presence of NADPH, NADH or the NADP⁺/G6PDH recycling system, the reduction yielded the (*S*)-enantiomeric product at 15-20% *ee* (Hall *et al.*, 2008a). Reductions in the presence of the NAD⁺/GDH recycling system resulted in the formation of the (*R*)-enantiomeric product at 77% *ee* (Hall *et al.*, 2008a).

(1S,2S,4S)-dihydrocarveol

(1R,2S,4S)-dihydrocarveol

(1S,4S)-dihydrocarvone

(4S)-(+)-carvone

(1R,4S)-dihydrocarvone

(1S,2R,4S)-dihydrocarveol

(1R,2R,4S)-dihydrocarveol
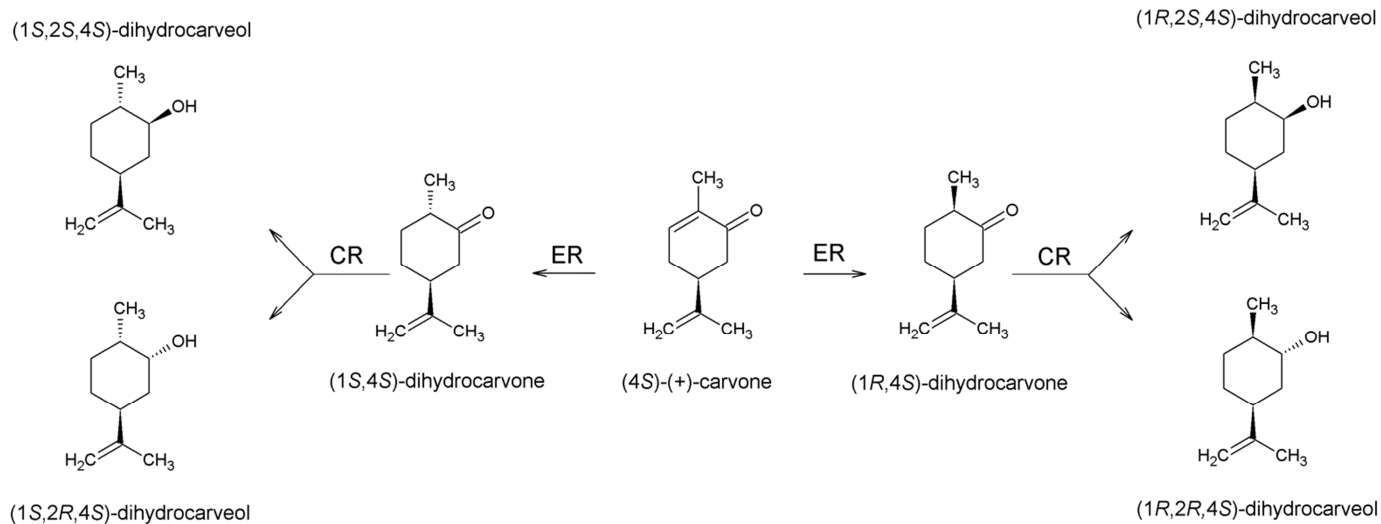
Figure 1.22: Proposed pathway for the reduction of (4S)-(+)-carvone by enoate reductase (ER) and carbonyl reductase (CR) from yeasts (Adapted from Goretti *et al.*, 2009).

a)

(5R)-carvone or (5S)-carvone

(2R,5R)-dihydrocarvone and (2R,5S)-dihydrocarvone

b)

(Z/E) citral

(S)-citronellal

(R)-citronellal
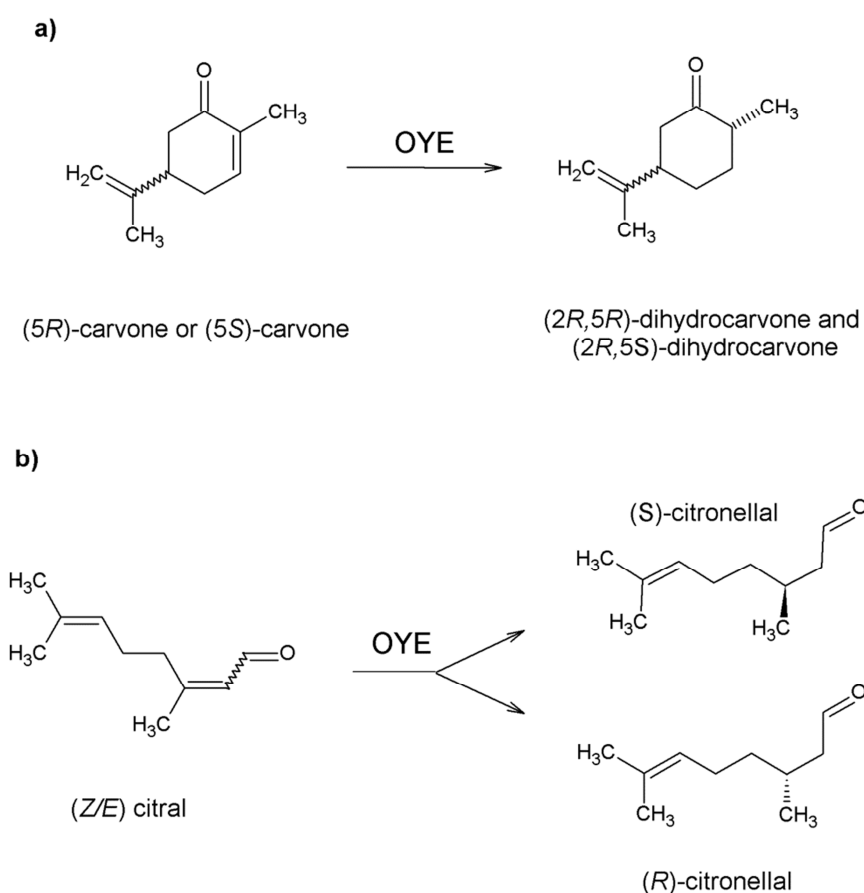
Figure 1.23: **a)** Asymmetric bioreduction of (S/R)-carvone to (2R,5R)-dihydrocarvone and to (2R,5S)-dihydrocarvone. **b)** Asymmetric bioreduction of (Z/E)-citral to the (S)- and (R)-enantiomeric products (Adapted from Adalbjörnsson *et al.*, 2010).

43

1.5. TOWARDS UNDERSTANDING THE PHYSIOLOGICAL FUNCTION OF THE OYE FAMILY

Extensive research into the OYE family of flavoenzymes has revealed a vast and diverse range of reactions catalysed by these enzymes (Williams & Bruce, 2002), yet the physiological function of the OYE remains elusive. Even so, a few studies have begun to shed light on these enzymes' physiological function. One such study revealed the role played by OYEs in protecting *S. cerevisiae* cells against acrolein toxicity (Trotter *et al.*, 2006).

Acrolein, a ubiquitous reactive aldehyde in biological systems, is formed as a product of lipid peroxidation and has a number of cytotoxic effects, including inhibition of a variety of enzymes (Jamieson, 1998; Trotter *et al.*, 2006). Lipid peroxidation occurs when unsaturated lipids serve as major targets for reactive oxygen species (ROS) and results in the formation of reactive lipid and fatty acid radicals. These radicals can attack other fatty acids, leading to oxidative chain reactions that proceed until they come into contact with an antioxidant (such as vitamin E). The oxidative reactions result in the formation of various lipid hydroperoxides which are relatively unstable and are converted into breakdown products which can be toxic (Esterbauer, 1993). Major breakdown products include the highly toxic reactive aldehydes. Among the α,β-unsaturated aldehydes present in these biological systems, acrolein is one of the strongest electrophiles, reacting with nucleophiles such as the sulfhydryl group of cystein, the imidazole group of histidine and the amino group of lysine (Uchida *et al.* 1998).

When performing genome-wide screening to identify *S. cerevisiae* mutants sensitive to acroline, a wide range of gene functions were identified as playing a role in protecting the yeast cells against acrolein stress (Trotter *et al.*, 2006). However, OYE2 was shown to be a requirement for acrolein tolerance in *S. cerevisiae*, with mutants lacking the *OYE2* gene showing poor growth on plates containing 3 mM acrolein. In contrast, loss of *OYE3* did not affect acrolein tolerance. Overexpression of either *OYE2* or *OYE3* resulted in a marked increase in resistance to acrolein. These results, along with the observation that propionaldehyde (the product of acrolein reduction by OYE) is non-toxic to yeast, indicate that *OYE2* and *OYE3* play a role in protecting the yeast cells form acrolein. Also observed in the study was the inability of the OYEs to protect against larger active aldehydes, as increasing alkyl substitution to the β carbon of the aldehydes resulted in decreased reduction rates by the OYEs. While *in vitro* analysis has shown that OYE can reduce the α,β-unsaturated ketones methyl vinyl ketone (MVK), 3-penten-2-one, as well as crotonaldehyde, overexpression of *OYE2* and *OYE3* in *S. cerevisiae* did not result in increased resistance towards 3-penten-2-one or crotonaldehyde, but did result in increased resistance to MVK (Trotter *et al.*, 2006). One possible explanation lies with the sizes of these molecules – the increased sizes of crotonaldehyde and 3-penten-2-one could lead to decreased solubility (compared to acrolein and

MVK). As a result, acrolein and MVK can move freely into the cells, while crotonaldehyde and 3-penten-2-one may not fully penetrate the cell membranes, making them inaccessible to OYE2 and OYE3.   While shedding light on a possible physiological function of the OYE of *S. cerevisiae*, the study also demonstrated that care must be taken when extending *in vitro* findings based on purified enzymes to the activities under physiological conditions in the cell (Trotter *et al.*, 2006).

Another possible physiological role of OYE was observed in plants, where the enzyme appears to play a role in the biosynthesis of jasmonates (Wasternack & Kombrink, 2010). The plant hormone jasmonic acid (JA) and its metabolites (collectively called jasmonates) are lipid-derived signals involved in stress responses to insects and pathogens, wounding, UV light and other abiotic stresses.   In addition, jasmonites modulate root growth, flower development and other developmental processes of healthy plants.   Jasmonites are generated in the chloroplast and peroxisomes *via* one specific branch of oxylipin biosynthesis – the allene oxide synthase (AOS) branch of the so-called lipoxygenase (Lox) pathway. Of particular interest in this pathway is the role played by oxophytodienic acid reductase (OPR) in the synthesis of the jasmonate-precurser 3-oxo-2-(2'[Z]-pentenyl)-cyclopentan-1-octanoic acid (OPC8) in the peroxisomes.   Mutagenesis studies performed using *Arabidopsis* and tomato plants confirmed that *opr3* mutants are JA-deficient (Wasternack & Kombrink, 2010).

1.6. CONCLUSIONS AND INTRODUCTION TO THE STUDY

Despite the failure to conclusively solve the mystery of the OYE's physiological function, the broad spectrum of OYE substrates identified over the years have provided numerous opportunities for industrial and bioremedial application of these biotransformations (Toogood *et al.*, 2010; Williams & Bruce, 2002). Indeed, the OYE family has been revealed as an important source of potential novel biocatalysts. The wealth of knowledge regarding the structural characteristics and binding modes of substrates and inhibitors in the OYE active site can assist not only in the rationalisation of the stereochemical outcomes of known reactions, but also serve as an aid in targeting and predicting the industrial viability of new OYEs as biocatalysts (Toogood *et al.*, 2010). As imperative as *in vitro* substrate specificity studies and sequence-based information are, care must be taken when drawing conclusions regarding the enzyme family's activity and functions in living cells (Trotter *et al.*, 2006).

This study demonstrates and investigates the principle of identifying potential candidates for industrial biocatalysis among the OYEs by utilizing current structural knowledge combined with insights into structure-function relationships. In addition, functional and structural characterisation of two target OYEs will be performed in an attempt to broaden knowledge regarding substrate specificity and scope, while also serving as a method to discover new potential biocatalysts. Lastly, the study will provide insight into the value of predictions made regarding enzyme function, from sequence and structural data.

SEQUENCE-BASED ANALYSIS OF OYE HOMOLOGUES FROM BACTERIA AND ARCHAEA

2.1. INTRODUCTION

With the development of molecular techniques and the advent of genome projects, the amount of sequence data available is growing rapidly (Shenoy & Jayaram, 2010). One of the greatest challenges is using the available genomic and structural data to decipher the structural, functional and evolutionary characteristics of proteins (Shenoy & Jayaram, 2010). The wealth of sequence and structural data of OYE homologues available has proven to be a vital resource in the study of the OYE family and their biocatalytic abilities (Adalbjörnsson *et al.*, 2010; Oberdorfer *et al.* 2011; Toogood *et al.*, 2010; Williams & Bruce, 2002). Availability of this data has allowed great progress to be made in the characterisation of OYEs and the study of the flavoenzymes' viability for biocatalytic implementation (Toogood *et al.*, 2010).

One of the most important tools in the functional annotation of proteins from sequence data is the use of phylogenetic methods to elucidate the evolution of protein homologues (Shenoy & Jayaram, 2010). This chapter describes the sequence-based analysis of a vast number of OYE homologues from three classes in the phylum Proteobacteria, the phylum Firmicutes and the domain Archaea through phylogenetic methods. Whereas by no means an exhaustive evolutionary analysis, this chapter investigates the conservation of residues implicated in substrate binding and modulation of the bound flavin's electropotential among OYE homologues.

2.2. MATERIALS AND METHODS

2.2.1. Similarity searches and multiple alignments

Similarity searches were performed using tBLASTn (Altschul *et al.*, 1997) against the non-redundant nucleotide database available from the National Center for Biotechnology Information (NCBI). The amino acid sequence of the OYE homologue, enoate reductase from *Thermus scotoductus* SA-01 (Opperman *et al.*, 2010) was used as query. tBLASTn searches were performed separately for each selected class of the Proteobacteria. Similarly, a separate tBLASTn search was performed limited to the Archaea and the Firmicutes. Nucleotide and translated amino acid sequences of old yellow enzyme homologues identified were retrieved from the NCBI database. Sequences of OYE homologues were numbered as retrieved and given names combining a three letter code for the organism of origin, along with the assigned number. Multiple sequence alignments of the amino acid sequences were performed using the MUSCLE EBI web tool ttp://www.ebi.ac.uk/Tools/msa/muscle) with the default parameters (Edgar, 2004). Alignments were manually inspected and partial sequences, as well as sequences with visible frame-shifts, were removed. In addition, sequences clearly belonging to fusion proteins were also removed.

2.2.2. Construction of maximum likelihood trees

The best amino acid substitution model was estimated for tree building using the Molecular Evolutionary Genetics Analysis version 5 (MEGA 5) software (Tamura *et al.*, 2011). The Whelan and Goldman (WAG) model (Whelan & Goldman, 2001) was selected. Evolutionary rate differences among sites were modelled according to a discrete Gamma distribution with 5 rate categories and by assuming that a certain fraction of sites are evolutionary invariable (G+I). An un-rooted maximum likelihood tree was constructed from multiple alignments of amino acid sequences of OYE homologues using MEGA 5. The phylogenetic tree was inferred using Nearest-Neighbour-Interchange (NNI) with bootstrap support for individual nodes calculated on 500 replicates.

2.2.3. Homology modelling

Homology modelling was performed using Multi-Sources Threader (MUSTER) (http://zhanglab.ccmb.med.umich.edu/MUSTER; Wu & Zhang, 2008). Flavin cofactors were inserted using Yasara. Energy minimisation was performed in Yasara using the YASARA2 force field (Krieger *et al.*, 2009).

2.3. Results and Discussions

2.3.1. Similarity searches and multiple alignments

In order to identify OYE homologues for evolution studies, tBLASTn similarity searches were performed against the non-redundant (nr) nucleotide database available from the National Center for Biotechnology Information (NCBI). The amino acid sequence of enoate reductase from *Thermus scotoductus* SA-01 was used as query sequence. tBLASTn acts by comparing a protein query sequence against the six-frame translations of a nucleotide sequence database (Wheeler & Bhagwat, 2007). As the translated searches make comparisons at the level of protein sequence, translated BLAST searches are more sensitive than direct nucleotide sequence searches. In addition, tBLASTn is useful for identifying, and searching among, unannotated coding regions (Wheeler & Bhagwat, 2007).

Three separate similarity searches were performed: the first limited only to the class betaproteobacteria, the second to the class gammaproteobacteria and the third to the class alphaproteobacteria. The fourth similarity search was performed limited to the Firmicutes and the fifth to the Archaea. The Proteobacteria and Archaea were selected as targets, as the Proteobacteria is a large and physiologically diverse group of bacteria (Imhoff, 1995; Stackebrandt *et al.*, 1988) while the Archaea and Firmicutes include a number of extremophile organisms (Horikoshi, 2010).

OYE sequences that were annotated as OYEs (NADPH/NADH-dependent oxidoreductases) were selected. However, due to the presence of a number of putative proteins in tBLASTn hits and the known diversity of the OYE family (Toogood *et al.*, 2010), sequences with above 30% identity to the query and with above 90% coverage, were also selected. However if, after construction of maximum likelihood trees in 2.2.2, sequences were revealed as vast outliers, these sequences were discarded. In addition, alignments were inspected for sequences bearing frame-shift mutations, which were subsequently discarded, along with any visible partial sequences or sequences clearly belonging to fusion proteins.

For each group of OYEs (from the alphaproteobacteria, betaproteobacteria, gammaproteobacteria, Firmicutes and Archaea, respectively), multiple alignments were performed using the MUSCLE algorithm and employing the default parameters which have been shown to give the best average accuracy (Edgar, 2004). The MUSCLE algorithm was selected, as it has been reported to achieve both better average accuracy and at better speeds, compared to CLUSTALW of T-Coffee (Edgar,

2004). Duplicated sequences visible in multiple alignments were identified and duplicates deleted where applicable.

2.3.2. Construction of maximum likelihood trees

The best amino acid substitution model was estimated for tree building using the MEGA 5 software (Tamura *et al.*, 2011). MEGA 5 allows the evaluation of major models of nucleotide and amino acid substitutions frequently desired by researchers. For amino acid substitution, six models, with and without empirical frequencies (+F), can be evaluated for amino acid substitution. These models include the Poisson, Equal-Input, Dayhoff (+F), Jones-Taylor-Thornton (+F) and Whelan and Goldman (+F) models. In addition, the goodness-of-fit of the substitution models can be evaluated with and without assuming the existence of evolutionary rate variation among sites. The evolutionary rate variation among sites is modelled by a discrete Gamma distribution (+G) (Yang, 1994) and/or the allowance for the presence of invariant sites (+I) (Fitch & Margoliash, 1967; Fitch 1986; Shoemaker & Fitch 1989). Evaluation of the different models allows quick examination of the robustness of the estimates of evolutionary parameters under different models of substitutions and assumptions about the distribution of evolutionary rates among sites (Tamura et al., 2011). For each model, a goodness-of-fit to the data is measured by the Bayesian information criterion (BIC; Schwarz, 1978). Models with the lowest BIC scores are considered to describe the substitution pattern the best. In this study, analysis of the numerous models revealed the WAG+I+G model to be best. As a result, a bootstrap consensus un-rooted maximum-likelihood (ML) tree was inferred for each group of OYE homologues using the Whelan and Goldman (WAG) model (Whelan & Goldman, 2001) with a discrete Gamma distribution with 5 rate categories and by assuming that a certain fraction of sites are evolutionary invariable. The maximum likelihood (ML) methods (Adachi & Hasegawa, 1996; Adachi *et al.*, 2000) have been shown to be superior to the counting methods such as maximum parsimony, with the drawback that it is computationally intensive (Whelan & Goldman, 2001).

2.3.3. Analysis of OYE homologues of the gammaproteobacteria

The number of amino acid sequences of OYE homologues identified and used for the gammaproteobacteria totalled 199, with all duplicates removed. Figure 2.1 shows the bootstrap consensus unrooted ML tree obtained for OYE homologues identified among organisms of the class gammaproteobacteria. As discussed in 1.2.1.4., members of the "classical" subclass of OYE, such as OYE1 and morphinone reductase, have conserved His and Asn residues (H191 and N194, OYE1 numbering) present in the active site that have been shown to play an important role in the binding and positioning of substrates (Adalbjörnsson *et al.*, 2010; Fox & Karplus, 1994). In

members of the "thermophilic-like" OYE subclass, such as YqjM and XenA, these two residues are replaced by a conserved His pair (His164 and His167, YqjM numbering) (Adalbjörnsson *et al.*, 2010; Kitzing *et al.*, 2005). It was thus decided to analyse the obtained tree and multiple alignments first by identifying where OYE homologues with the corresponding His and Asn residues in the catalytic site have been placed in the ML tree. Similarly, the placement in the ML tree of OYE homologues with the corresponding His-pair in the catalytic site was identified. Figure 2.2 shows the placement of the respective OYE homologues on the bootstrap consensus ML tree. By targeting these pairs of catalytic residues, the ML tree was clearly 'divided' into two groups, branching at the node labelled x. The larger group (group B_γ), consisting of approximately 60% of the OYE homologues analysed, exhibited the His-pair arrangement in the active site, whereas the smaller group (group A_γ, approximately 40% of the OYE homologues analysed) were found to have the His and Asn arrangement of residues in the catalytic site. It was clear that, despite the great sequence diversity among OYE homologues of the gammaproteobacteria, OYEs that have the same catalytic residues in this target area of the catalytic site are grouped together by the ML analysis performed.

Figure 2.1: Phylogenetic relationship of OYE homologues from gammaproteobacteria through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model. Percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches.

Figure 2.2: Phylogenetic relationship of OYE homologues from gammaproteobacteria through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the presence of the His and Asn residues in the active site (group A$_\gamma$, approximately 40% of OYE homologues) and the presence of the His-pair in the catalytic site (group B$_\gamma$, approximately 60% of homologues, respectively).

Further analysis of the bootstrap consensus tree obtained for OYE homologues among the gammaproteobacteria was done targeting the catalytic residue that has been implicated in modulating the reduction potential of the flavin, as discussed in 1.2.1.3. In the "classical" OYEs, such as OYE1, Thr37 (OYE1 numbering) has been shown to modulate the redox potential of the enzyme, as the hydrogen bond formed between O4 of the FMN and the side-chain hydroxyl group of Thr37 causes electrons to be withdrawn from the flavin ring (Xu *et al.*, 1999). This not only enhances the electrophilicity of the oxidised enzyme, but also serves to stabilise the negatively charged flavin. In "thermophilic-like" OYEs, such as YqjM, the corresponding flavin-modulating residue has been identified as a Cys residue (Cys26 in YqjM; Kitzing *et al.*, 2005). It was thus decided to further analyse the placement of OYE homologues in the ML tree according to the identity of the flavin modulator residue, in conjunction with the identity of the two substrate-binding catalytic residues. Figure 2.3 shows the result of this analysis on the bootstrap consensus ML tree. By targeting these catalytic residues, the ML tree was, rather surprisingly, clearly 'divided' into 4 subgroups. The largest group corresponds to group A_γ obtained in the first analysis, consisting of approximately 40% and branching from the other three groups at the node labelled x. As mentioned, OYE homologues in group A_γ exhibited the His with Asn arrangement in the active site, but combined with a Thr as flavin modulator. This arrangement corresponds with that reported for OYEs such as morphinone reductase (MR) from *Pseudomonas putida* (Xu *et al.*, 1999; Toogood *et al.*, 2010).

The remaining OYE homologues, all clustered together in group B_γ in the first analysis step, appeared subdivided further into 3 subgroups. The second largest group (group $B2_\gamma$, branching at the node labelled y), consists of approximately 30% of the OYE homologues analysed. This group exhibited the His-pair arrangement in the active site, combined with a Cys residue as flavin modulator. This arrangement corresponds with that reported for the OYE described for XenA from *Pseudomonas putida* (Spiegelhauer *et al.*, 2010). The smaller, remaining groups (groups $B1_\gamma$ and $B3_\gamma$) consist of approximately 15% of the OYE homologues each and both exhibited the His-pair arrangement. However, group $B3_\gamma$ has a Thr residue as flavin modulator, whereas group $B1_\gamma$ has either a Glu or a Ser residue as flavin modulator. Group B3 corresponds with the structure described for PETN reductase from *Enterobacter cloacae* (Barna *et al.*, 2001).
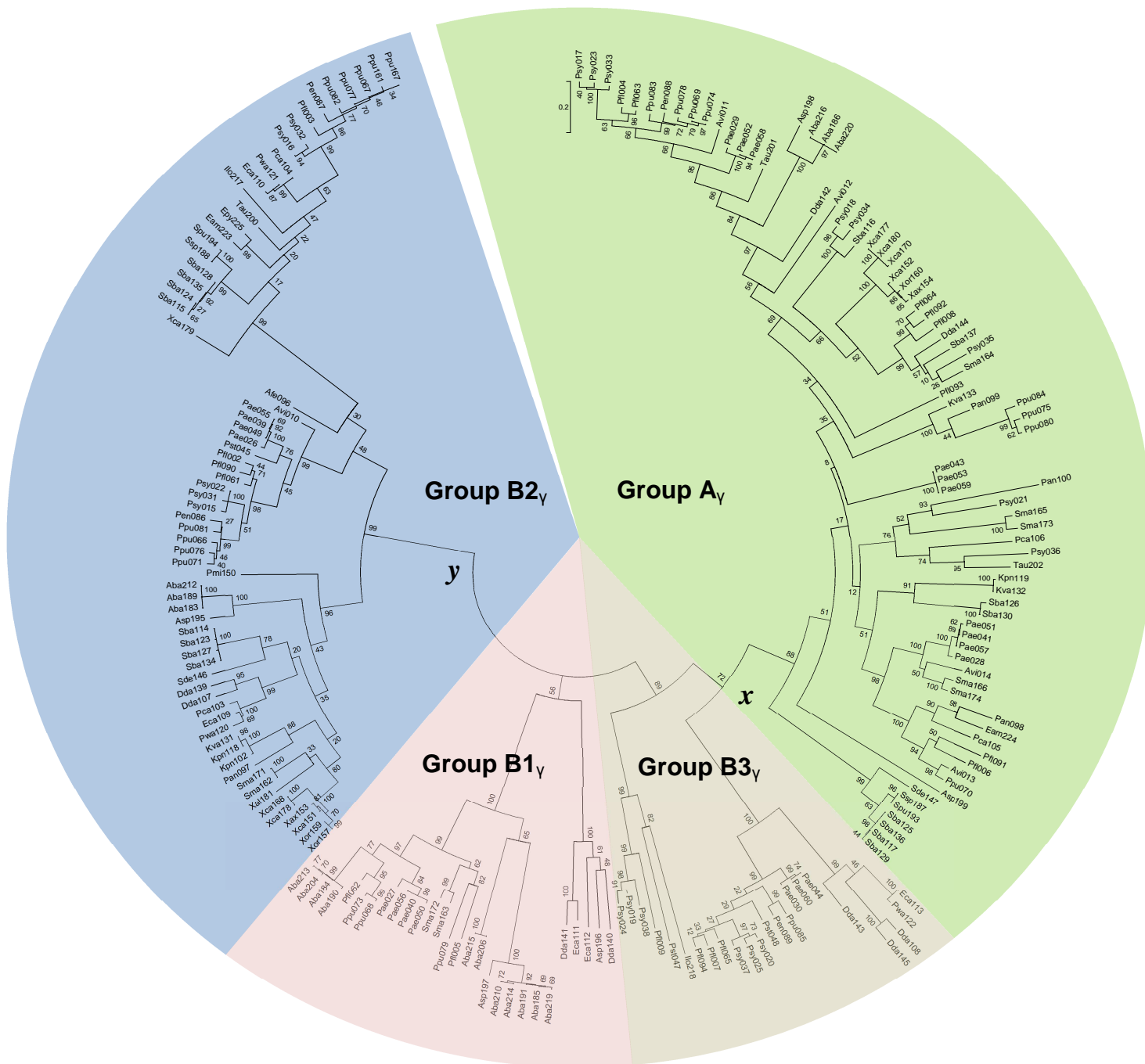
Figure 2.3: Phylogenetic relationship of OYE homologues from gammaproteobacteria through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the identities of the previously-target substrate-binding residues (His-pair or His and Asn residue), combined with the identity of the flavin modulator residue. Group A1$_\gamma$ (approximately 40% of OYE homologues) consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. Group B1$_\gamma$ (approximately 15% of the OYE homologues) consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Glu or Ser as flavin modulator. Groups B2$_\gamma$ (approximately 30% of the OYE homologues) and B3$_\gamma$ (approximately 15% of the OYE homologues) both have the His-pair in the catalytic site, but combined with a Cys and Thr residue as flavin modulator, respectively.

Figure 2.4 is a graphical representation of the multiple alignment obtained containing all the OYE homologues analysed from the gammaproteobacteria, as well as figures illustrating the active site architecture of OYE homologues in each subgroup. For groups A_γ, $B2_\gamma$ and $B3_\gamma$, cloning and characterisation for at least one OYE homologue in the group has been reported: MR for group A_γ, XenA for group $B2_\gamma$ and PETN reductase for group $B3_\gamma$. Illustrations of the active site architecture for these three groups were thus derived from the crystal structures available. As no cloning or characterisation has been reported for any OYE homologues in group B1, a homology model was constructed from the translated nucleotide sequence of the OYE homologue labelled Ppu086 from *Pseudomonas putida*. Homology modelling was performed using MUSTER (Wu & Zhang, 2008). Shown on the alignment are the four separate groups obtained through the above analysis. Illustrations of the active site architecture show the two substrate-binding residues in the active site, as well as the flavin modulator residue. Also notable is a highly conserved Tyr residue (Tyr186 in PETN reductase) which has been shown to play the role of proton donor during catalysis in many OYEs (section 1.2.1.4). Apparent in the homology model obtained for Ppu086 is the incorrect orientation of the His-pair in the catalytic site. As a result, it is unclear whether interaction between the Glu flavin modulator and the flavin will occur. This model, however, is purely for illustrative purposes.

Grouping together of the diverse OYE homologues of the gammaproteobacteria according to the identity of the flavin modulator, combined with the identity of the two target residues in the active site, results in a grouping that corresponds to that obtained from evolutionary analysis of the OYE homologues using the selected ML method.
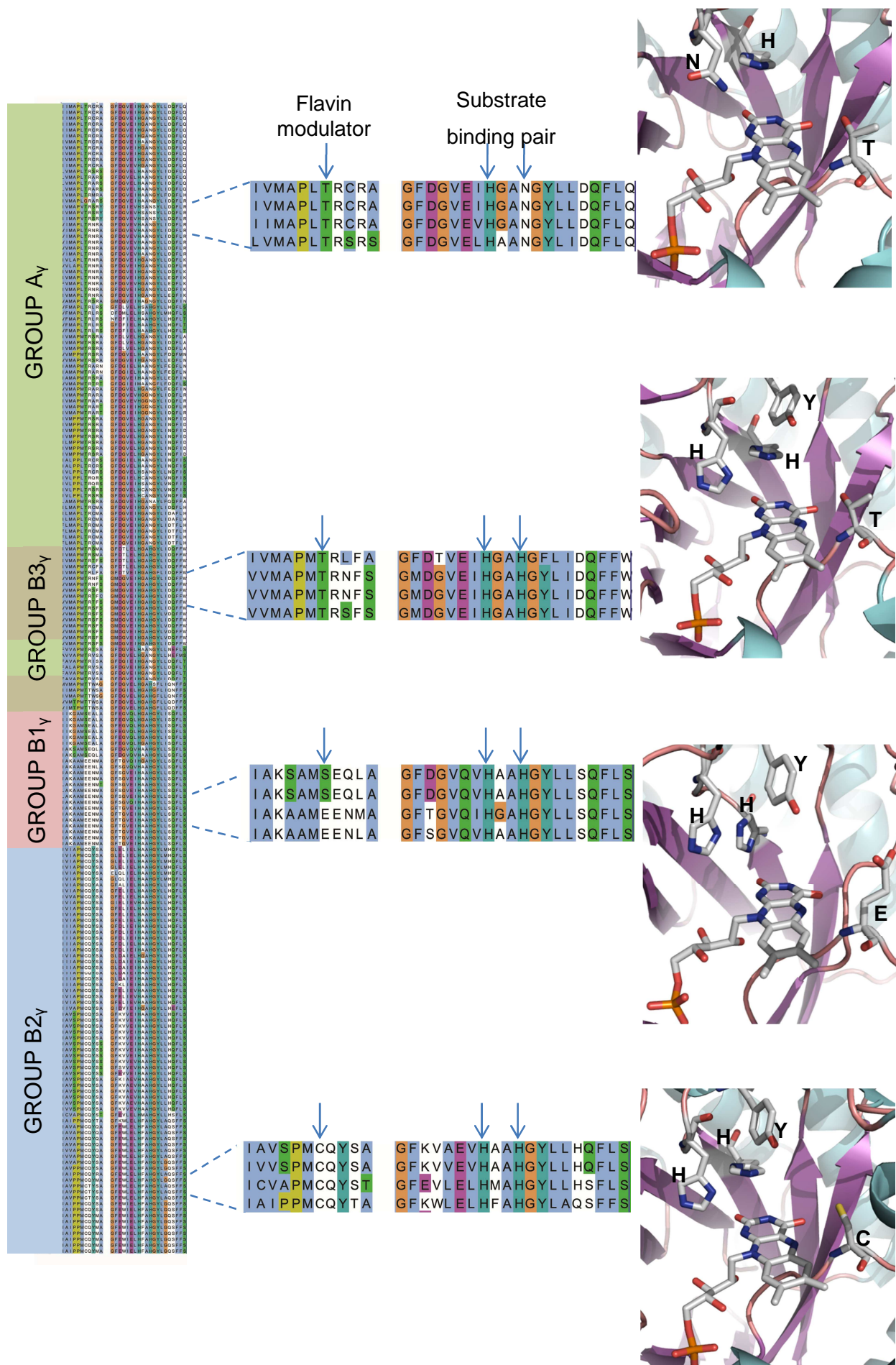
Figure 2.4: Graphical representation of the multiple alignment of OYE homologues from the gammaproteobacteria, showing the four identified subgroups, as well as the typical active site architecture of homologues for each group.

2.3.4. Analysis of OYE homologues of the beta- and alphaproteobacteria

The number of OYE homologues from organisms of the class betaproteobacteria analysed totalled to 160, with duplicates removed. Analysis of the bootstrap consensus tree performed was as described for the gammaproteobacteria OYE homologues. As observed for gammaproteobacteria, targeting of OYE homologues according to the identity of substrate-binding residues (His-pair or His and Asn) appeared to divide the ML tree into two groups. However, the two groups were more similar in size than observed for the gammaproteobacteria. Group A_β (approximately 47% of total OYE homologues) consists of the OYE homologues exhibiting the His-Asn arrangement of the targeted substrate-binding residues, whereas group B_β (approximately 53% of total OYE homologues) consists of OYE homologues with the His-pair as substrate binding residues. Identifying the placement on the ML tree of OYE homologues with identical flavin modulator residues resulted in division of the ML tree into only 3 subgroups. As observed among the OYE homologues of the gammaproteobacteria, the first subgroup corresponds to group A_β obtained from the first analysis step, consisting of OYE homologues with a His and Asn residue as substrate-binding residues, combined with a Thr as flavin modulator residue. However, clustered in group A_β are two OYE homologues with His-pair in the active site, combined with the Thr flavin modulator. The second subgroup (Group $B2_\beta$, approximately 34% of total OYE homologues) consists solely of OYE homologues that combine a His-pair as substrate-binding residues with a cysteine as flavin modulator. The final subgroup (Group $B1_\beta$, approximately 19% the OYE homologues) combined the His pair in the catalytic site with a Ser, Glu or Asn residue. The result of the analysis on the ML tree is shown in figure 2.5.
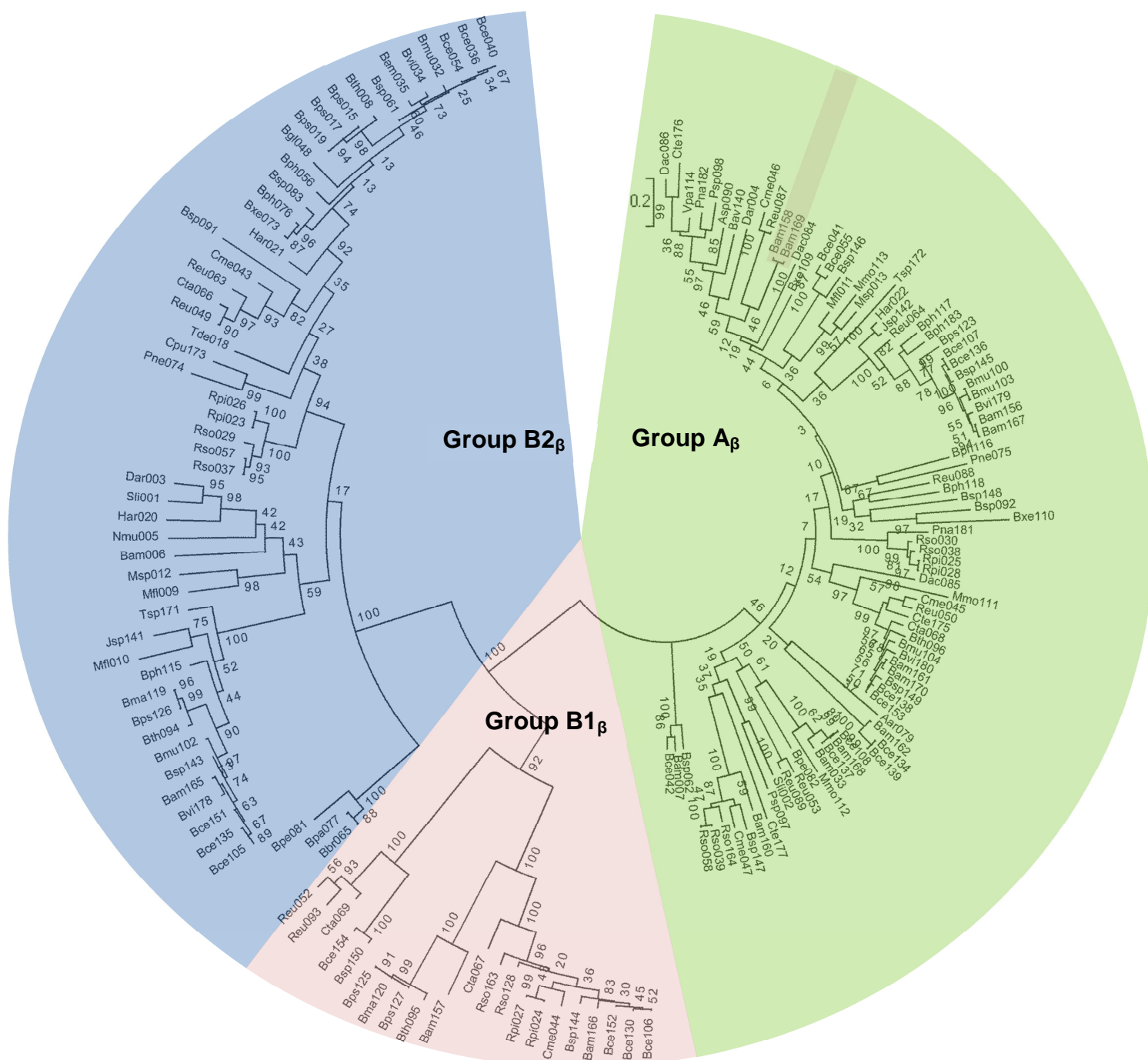
Figure 2.5: Phylogenetic relationship of OYE homologues from betaproteobacteria through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the identities of the target substrate-binding residues (His-pair or His and Asn residue), combined with the identity of the flavin modulator residue. Group A_β (approximately 47% of OYE homologues) consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. The two homologues in the group with His-pair in the catalytic site are highlighted in brown. Group $B2_\beta$ (approximately 34% of the OYE homologues) consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Cys as flavin modulator. Group $B1_\beta$ (approximately 19% of the OYE homologues) have the His-pair in the catalytic site, but combined with a Ser, Glu or Asn residue as flavin modulator.

59

Analysis of OYE homologues from the alphaproteobacteria gave similar results to those from the betaproteobacteria. The number sequences analysed totalled to 102, without duplicates. Analysis of the ML tree obtained by identifying the placement of OYEs with the His-Asn arrangement resulted in 'division' of the ML tree into two groups almost equal in size (figure 2.6). Group A_α (approximately 49% of OYE homologues) was identified as consisting of OYE homologues with the His and Asn residues in the catalytic site, where group B_α (approximately 51% of OYE homologues) was identified as consisting of OYE homologues with the His-pair arrangement in the catalytic site. Identification of OYEs according to the identity of the flavin regulator residue gave the result comparable to that found for OYEs of the betaproteobacteria. Group A_α was confirmed as consisting of OYEs with the presence of Thr as flavin modulator. YqjM-like OYEs were found to make up group $B2_\alpha$, exhibiting the characteristic His-pair arrangement with Cys as flavin modulator residue. A small group ($B3_\alpha$, approximately 4% of the OYE homologues) consisted of PETN reductase-like OYEs with the His-pair combined with a Thr as flavin modulator. The final group (group $B1_\alpha$, approximately 12% of OYE homologues) was found to consist only of OYE homologues with the His-pair arrangement combined with a Glu as flavin modulator residue.

Figure 2.6: Phylogenetic relationship of OYE homologues from alphaproteobacteria through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the identities of the target substrate-binding residues (His-pair or His and Asn residue), combined with the identity of the flavin modulator residue. Group A$_\alpha$ (approximately 49% of OYE homologues) consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. Group B2$_\alpha$ (approximately 35% of the OYE homologues) consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Cys as flavin modulator. Group B3$_\alpha$ (approximately 4% of OYE homologues) consists of OYEs with His-pair combined with a Thr as flavin modulator. Group B1$_\alpha$ (approximately 12% of the OYE homologues) have the His-pair in the catalytic site, but combined with a Glu residue as flavin modulator.

61

2.3.5. Analysis of OYE homologues of the Firmicutes

For the Firmicutes, 91 OYE homologues were analysed. Construction and analysis of the bootstrap consensus tree was performed as described for the Proteobacteria OYE homologues. Targeting of OYE homologues according to the identity of substrate-binding residues (His-pair or His and Asn) once again appeared to divide the ML tree into two groups, as observed with all the Proteobacterial classes analysed. However, Group A_F (consisting of the OYE homologues exhibiting the His-Asn arrangement) constituted only 12% of the total number of homologues analysed. The remaining 81% of OYE homologues fell within group B_F (consisting of OYE homologues with the His-pair as substrate binding residues). Targeting OYE homologues with identical flavin modulator residues resulted in division of the ML tree into 5 groups (figure 2.7). As observed for all the Proteobacteria analysed, the first group corresponded to group A_F, consisting of OYE homologues with a His and Asn residue in the catalytic site, combined with a Thr as flavin modulator. However, one OYE homologue (Bli009, from *Bacillus lichenformis*), was grouped into A_F, despite having a Glu as flavin modulator residue. In addition, as seen with the betaproteobacteria, an OYE homologue (Bwe059, from *Bacillus weihenstephanensis*), grouped into group A_F despite having a His-pair in the active site, combined with a Thr residue. The presence of errors in the whole genome sequencing of the organism must not be ruled out as a possible cause.

18% of the OYE homologues analysed grouped together to form group $B3_F$ – OYEs combining the His-pair arrangement with a Thr as flavin modulator residue. The largest group (52% of OYE homologues) was identified as group $B2_F$, containing OYEs that combine the His-pair in the active site with a Cys residue as flavin modulator. Included in this group are OYEs for which structural and functional characterisation has been reported: Tps084 (OYE from the thermophile *Thermoanaerobacter pseudethanolicus*; Adalbjörnsson *et al.*, 2010), Bsu021 (YqjM from *Bacillus subtilis*; Kitzing *et al.*, 2005) and Gka001 (OYE from the thermophile *Geobacillus kaustophilus*; Schittmayer *et al.*, 2011). A third, new group (group $B5_F$) was apparent, consisting of twelve OYEs combining the His-pair in the catalytic site with an Arg, His, Val or Trp residue in the position aligning with the flavin modulator residue. An additional five OYEs with Val residues aligning with the flavin modulator residues clustered together in a separate group, group $B4_F$. However, ten of the twelve OYEs in group $B5_F$ and two of the homologues in group $B4_F$ have either a Thr, Cys, Glu or a Ser residue adjacent to the residue aligned in the position as flavin modulator. As all these residues have been identified as flavin modulator residues in some OYE homologues from the Proteobacteria, it is possible that these adjacent residues rather serve the flavin modulators.
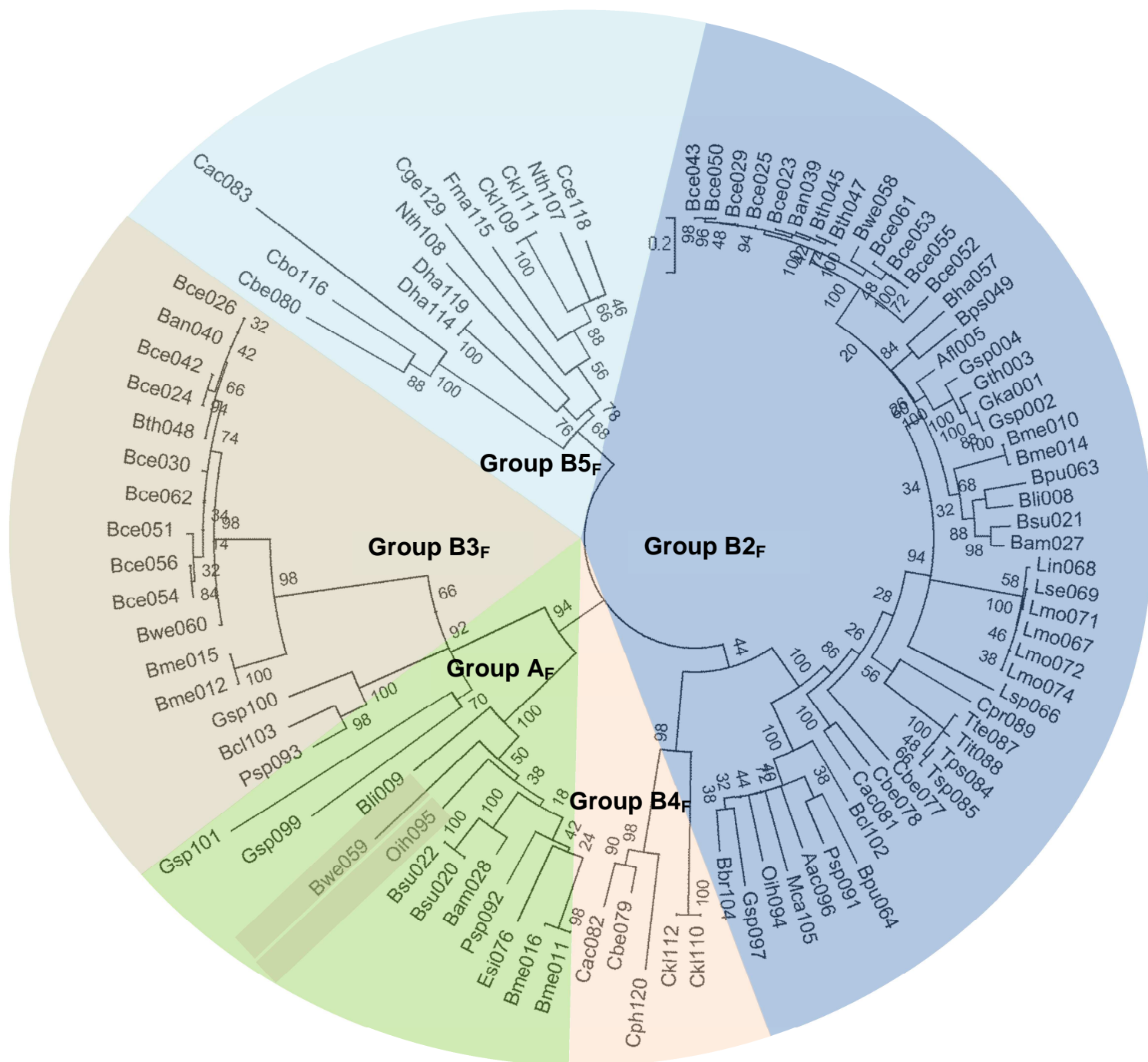
Figure 2.7: Phylogenetic relationship of OYE homologues from Firmicutes through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the identities of the target substrate-binding residues (His-pair or His and Asn residue), combined with the identity of the flavin modulator residue. Group A_F (approximately 12% of OYE homologues) consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. Two 'incorrectly' grouped homologues in group A_F are highlighted. Group $B2_F$ (approximately 52% of the OYE homologues) consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Cys as flavin modulator. 18% of the OYE homologues belong to group $B3_F$, having a Thr as flavin modulator combined with a His-pair. Group $B4_F$ (approximately 5% of OYE homologues) consists of OYEs with His-pair combined with a Val in the position of flavin modulator. Group $B5_F$ (approximately 13% of the OYE homologues) have the His-pair in the catalytic site, but combined with an Arg, His or Val in the position of flavin modulator residue.

63

2.3.6. Analysis of OYE homologues of the Archaea

The number of OYE homologues from Archaea was significantly fewer than observed for the different classes of Proteobacteria and Firmicutes, totalling to only 39, with duplicates removed. The limited number of OYE sequences available among the Archaea seems to be simply due to fewer Archaeal genomes being subjected to whole-genome sequencing compared to the Proteobacteria and Firmicutes. Targeting of OYE homologues according to the identity of substrate-binding residues (His-pair or His and Asn) again appeared to divide the ML tree into two groups – one small and one large. The smaller group, group A_A (approximately 15% of total OYE homologues) consists of the OYE homologues exhibiting the His-Asn arrangement of the targeted substrate-binding residues, whereas group B_A (approximately 85% of total OYE homologues) consists of OYE homologues with the His-pair as substrate binding residues. Identifying the placement on the ML tree of OYE homologues with identical flavin modulator residues resulted in division of the ML tree into 4 subgroups, as observed for the gammaproteobacteria. The first subgroup again corresponds to group A_A obtained from the first analysis step, consisting of MR-like OYE homologues with a His and Asn residue as substrate-binding residues, combined with a Thr as flavin modulator residue. The second subgroup (Group $B2_A$, approximately 26% of total OYE homologues) consists of YqjM-like OYE homologues that combine a His-pair as substrate-binding residues with a cysteine as flavin modulator. The third subgroup is the largest (Group $B4_A$, approximately 39% the OYE homologues) and combined the His pair in the catalytic site with (rather surprisingly) an Ile residue. Even more surprisingly, the final group (Group $B5_A$) consists of OYE homologues exhibiting the His-pair combined with an Arg, Asn or His residue. The result of the analysis on the ML tree is shown in figure 2.8.
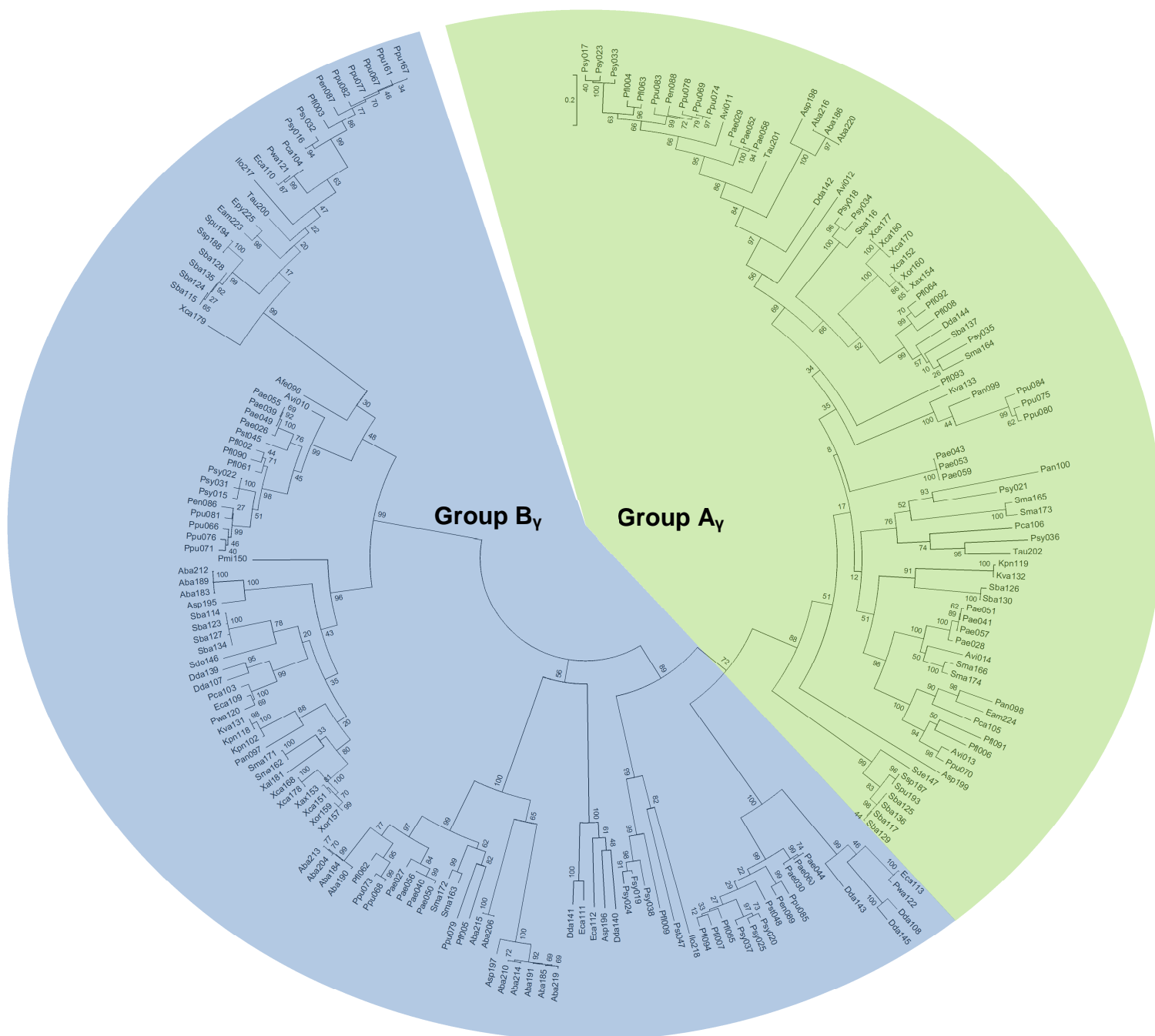
Figure 2.8: Phylogenetic relationship of OYE homologues from Archaea through a bootstrap consensus un-rooted maximum-likelihood (ML) tree inferred from the WAG+I+G model, showing grouping of OYE homologues according to the identities of the target substrate-binding residues (His-pair or His and Asn residue), combined with the identity of the flavin modulator residue. Group A$_A$ consists of OYE homologues with His and Asn as target residues in the catalytic site, combined with a Thr as flavin modulator. Group B2$_A$ consists of OYE homologues with the His-pair as target catalytic site residues, combined with a Cys as flavin modulator. Group B4$_A$ has the His-pair in the catalytic site, but combined with an Ile residue as flavin modulator. The final group, group B5$_A$, combines the His-pair with either an Asn, Arg or His in the position aligning with the flavin modulator residues of other OYE homologues.

2.3.7. General discussion

In a recent study, the correlation between protein function and protein structure was examined by comparison of the 12 unique OYEs for which structures are currently available from the protein data bank (PDB) (Oberdorfer *et al.*, 2011). Comparisons of the OYE structures and their reported enantioselectivity was performed by focussing mainly on two 'bounding residues' (Tyr, Phe or Ile; Tyr78 and Tyr 358 in OPR-1) involved in binding of substrate in the catalytic site of the enzymes. A strong correlation was found between the identity of these bounding residues and the enantioselectivity of the enzyme. However, it was only through examination of available crystal structures that these two residues (that occur at a distance from each other in the linear amino acid sequence) could be identified as possible role-players in the binding of substrate and thus the enantioselectivity of the enzymes. It was revealed that the identification of sequence motifs that allow the prediction of the functionality of OYEs would be a valuable tool in the search for biocatalytic application of OYEs. The need for such sequence motifs becomes truly apparent when comparing the small number of OYE crystal structures available to the available wealth of sequence data available. Table 2.1 shows all the OYE homologues for which crystal structures are available. It is clear that the problem is exacerbated by the fact that OYEs for which crystal structures are available form part of subgroups A, B2 or B3, with no cloning, characterisation or crystal structure data reported for OYE homologues in the remaining subgroups identified (Adelbjörnsson, *et al.,* 2010; Oberdorfer *et al.*, 2011; Toogood *et al.*, 2010).

The presence of conserved regions of primary (and tertiary) structure across the diverse OYE family points to members of the OYE family possibly being orthologous (Williams & Bruce, 2002). The discovery of the highly conserved three-residue fingerprint targeted in this study points to early divergent evolution in the OYE family. In addition, the slight differences observed in this sequence fingerprint among OYE homologs might be crucial for catalysis of their natural substrates. Directed evolution studies investigating the role of the two substrate-binding catalytic residues in OYE1 (His191 and Asn194) revealed that replacement of Asn194 with a His residue greatly reduced the ability of the enzyme to bind 2-cyclohexenone (Brown *et al.*, 1998). Yet numerous YqjM-like OYEs catalyse the reduction of 2-cyclohexenone with the presence of a His-pair as binding residues in the catalytic site (Brown *et al.*, 1998; Kitzing *et al.*, 2005). Clues to the significance of these differences may be found by studying OYEs that fall in different subgroups, but occur in the same species, such as OYEs from *Shewanella oneidensis* (Brigé *et al.*, 2006) and *Pseudomonas putida* (Spiegelhauer *et al.*, 2010). The genome of *Shewanella oneidensis* contains four different OYEs, three with the His and Asn residues combined with the Thr flavin modulator, and the fourth with the His-pair combined with Glu as flavin modulator residue (Brigé *et al.*, 2006). Similarly, the genome of *P.*

putida contains the genes for multiple OYEs, containing either Cys, Thr or Ala as flavin modulator residues (Spiegelhauer *et al.*, 2010).

While a study attempting comparisons among the four OYEs of *S. oneidensis* failed to result in the successful cloning and expression of the fourth OYE from the yet uncharacterised subgroup (Brigé *et al.*, 2006), mutagenic studies targeting the Cys flavin modulator of XenA (Cys25) demonstrated the value of targeted mutagenic studies in understanding the relationship between primary structure and biocatalytic function of OYEs (Spiegelhauer *et al.*, 2010). The study revealed that replacement of Cys25 with Ala had very little effect on the reduction potential of XenA, whereas replacement of Cys25 with Ser resulted in a +82 mV difference in the reduction potential compared to wild type XenA. Analysis of the reductive and oxidative half-reactions of the mutants towards the substrates 2-cyclohexenone and coumarin revealed that replacement of Cys25 with Ala results in a two-fold decreased rate constant for the reductive half-reaction and a 2-fold increase in the rate constant of the oxidative half-reaction compared to the wild type XenA. Replacement of Cys25 with Ser, however, accelerated the reductive half-reaction and decelerated the oxidation half-reaction four-fold compared to wild type XenA. As a result, when utilizing 2-cyclohexenone as substrate and NADH as reducing substrate, the reductive half-reaction is the rate-limiting step for both wild type XenA and the XenA-C25A mutant, while the rates of both half-reactions are balanced for the XenA-C25S mutant (resulting in XenA-C25S achieving the highest turnover for the substrate). When utilizing NADPH as reducing substrate, however, the XenA-C25S mutant is severely limited by the slow oxidative half-reaction, while the XenC25A mutant and wild type XenA have high turnover of the substrate with none of the two half-reactions acting as rate-limiting steps (Spiegelhauer *et al.*, 2010).

The effect of the identity of the flavin modulator residue has been demonstrated further through saturation mutagenesis studies using YqjM (Cys26) and 3-methylcyclohexenone as substrate (Bougioukou *et al.*, 2009). While wild type YqjM shows approximately 3% conversion of the substrate to the (*R*)-product with 76% *ee*, replacement of the Cys26 with Trp, Gly and Asp resulted in greatly improved conversions (above 60%, and up to 100% conversion for mutants replacing Cys26 with Asp). The three mutants also exhibited activity towards substituted 2-cyclohexenone substrates (with ethyl, isopropyl or CO_2Me substituents on C3) which are not utilised as substrates by wild-type YqjM. In addition, replacement of Cys26 with Trp or Asp resulted in great improvement of the enzyme's enantioselectivity, converting 3-methylcyclohexenone to the (*R*)-product at above 80% *ee*. More surprisingly, however, was the reversal of the enzyme's enantioselectivity upon replacement of Cys26 with Gly, with the (*S*)-product obtained at above 99% *ee* (Bougioukou *et al.*, 2009).

As mentioned, the genome of *P. putida* contains several copies of the XenA gene that differ in the identity of the flavin modulator residue. This has been suggested to be an adaptation allowing the enzymes to utilize different classes of substrates, ensuring that neither of the two half-reactions severely limit the overall catalytic rates (Spiegelhauer *et al.*, 2010). For YqjM, the vast improvement in conversion and enantioselectivity in previously uncatalysed reactions upon replacement of Cys26 with Glu, Asp or Trp residues strongly indicates that the 'unconventional' identity of the flavin modulator is an adaption in OYE homologues to catalyse a broader range of substrates with higher selectivity (Bougioukou *et al.*, 2009).

Table 2.1: Summary of all OYE homologues for which crystal structures are available, with identities of the two substrate-binding residues and flavin modulator residue targeted in the analysis of above ML trees, given.

Enzyme	PDB code	Organism	Substrate binding	Flavin modulator	Domain/Phylum/Class	Reference
OYE1	1OYB	*Saccharomyces pastorianus*	His and Asn	Thr	Eukaryotes	Fox & Karplus, 1994
OPR-1	1ICP	*Solanum lycopersicum*	His-pair	Thr	Eukaryotes	Breithaupt *et al.*, 2001
OPR-3	2HSA	*Arabidopsis thaliana*	His-pair	Thr	Eukaryotes	Breithaupt *et al.*, 2006
YqjM	1Z41	*Bacillus subtilis*	His-pair	Cys	Firmicutes	Kitzing *et al.*, 2005
XenA	2H8X	*Pseudomonas putida*	His-pair	Cys	Gammaproteobacteria	Griese *et al.*, 2006
TOYE	3KRU	*Thermoanaerobacter pseudethanolicus*	His-pair	Cys	Firmicutes	Adalbjörnsson *et al.*, 2010
TsOYE	3HF3	*Thermus scotoductus*	His-pair	Cys	Deinococci	Opperman *et al.*, 2010
GkOYE	3GR7	*Gebacillus kaustophilus*	His-pair	Cys	Firmicutes	Schittmayer *et al.*, 2011
PETN	1H50	*Enterobacter cloacae*	His-pair	Thr	Gammaproteobacteria	Barna *et al.*, 2001
SYE-1	2GOU	*Shewanella oneidensis*	His and Asn	Thr	Gammaproteobacteria	Van den Hemel *et al.*, 2006
N-ethylmaleimide reductase	3GKA	*Burkholderia pseudomallei*	His and Asn	Thr	Betaproteobacteria	Edwards & Staker, 2009
Morphinone reductase	1GWJ	*Pseudomonas putida*	His-pair	Thr	Gammaproteobacteria	Barna *et al.*, 2002

2.4. Conclusions

Conservation of the two substrate-binding residues, as well as the flavin monitoring residue seems highly significant in the evolution of OYE homologues. By targeting merely the identity of these three residues, OYE homologues of the alpha-, beta- and gammaproteobacteria, Firmicutes and Archaea are clustered in a similar manner as obtained by complex maximum likelihood methods. Clearly visible in this clustering are the two known subclasses of OYE (Oberdorfer *et al.*¸ 2011; Toogood *et al.*, 2010) "classical" and "thermophilic-like". However, two additional subclasses were apparent among OYEs of the Proteobacteria. Both subclasses consist of OYEs with the His-pair arrangement, but one combines the arrangement with Thr as flavin modulator while the other employs Ser, Glu or Asn as flavin modulator. Similarly, OYE homologues from the Firmicutes revealed two more possible subclasses of OYE, with Val, Trp, His or Arg as flavin modulator residue. Analysis of OYE homologues from Archaea revealed the clustering of OYE homologues combining the His-pair with an Ile residue as well as Asn, Arg or His in the position aligning with the flavin modulator residues of all the analysed OYE homologues.

While the presence of the highly conserved three-residue fingerprint indicates possible early divergent evolution among members of the OYE family (Williams & Bruce, 2002), mutagenesis studies targeting the flavin modulator residue of YqjM and XenA suggest that the different flavin modulator residues among OYE homologues could be adaptations enabling the enzymes to catalyse reactions utilizing their natural substrates (Bougioukou *et al.*, 2009; Spiegelhauer *et al.*, 2010).

Due to the limited structural information available for OYEs, the need for the ability to predict functionality of OYEs from sequence data is growing (Oberdorfer *et al.*, 2011). It was therefore decided to select two OYE homologues, each from a different subgroup, for expression and characterisation in an attempt to shed light on the correlation of function among OYE members belonging to different subclasses. The remaining section of the project describes the expression and characterisation of the two chosen OYE homologues.

CHAPTER 3

CLONING AND HETEROLOGOUS EXPRESSION OF NEW OYE GENES

3.1. INTRODUCTION

The synthesis of polymers, pharmaceuticals and speciality chemicals is often plagued by high cost processes, low selectivity and the production of undesired by-products (Demirjian *et al.*, 2001). Biocatalysts provide the industry with elegant solutions to these problems, and members of the OYE family are no exception (Demirjian *et al.,* 2001; Williams & Bruce, 2002). The discovery of extremophilic microorganisms and their enzymes (often referred to as extremozymes) has greatly impacted the field of biocatalysis, with the industrial applications of these resilient, often novel enzymes having increased significantly over the past two decades. Indeed, extremophiles are now regarded in enzyme discovery as vital, almost routine, sources of novel enzymes for biocatalysis (Demirjian *et al.,* 2001). OYEs from thermophiles for which cloning, expression and characterisation have been reported include the OYEs from *Thermus scotoductus* SA-01 (Opperman *et al.*, 2010), *Geobacillus kaustophilus* DSM 7263 (Schittmayer *et al.,* 2011) and *Thermoanaerobacter pseudethanolicus* E39 (Adalbjörnsson *et al.*, 2010).

This chapter describes the cloning and expression of two as yet uncharacterised OYEs from different extremophiles, in a mesophilic host. The first OYE gene is from *Capriavidus metallidurans* CH34 (formerly *Ralstonia metallidurans* CH34), a gram-negative organism (from the class betaproteobacteria) capable of flourishing in millimolar concentrations of heavy metals (Goris *et al.,* 2001; Vandamme & Coenye, 2004). The reference strain, CH34, was first discovered in 1976 and has the capacity to grow on minimal salt medium of low complexity containing cobalt, zinc and cadmium ions. Whole-genome sequencing of the organism revealed the presence of two large megaplasmids that contain genes for multiple resistances to the heavy metals Zn, Cd, Co, Pb, Cu, Hg, Ni and Cr. Due to its ability to withstand toxic metals, *C. metallidurans* is yet another candidate for use in bioremediation of heavy metal-contaminated environments (Goris *et al*, 2001).

The second OYE is from *Sulfolobus solfataricus* P2, an aerobic, hyperthermophilic crenarchaeon (She *et al.*, 2001). The organism grows optimally at 80℃ and pH 2-4. It is the most widely studied organism of the crenarchaeal branch of the Archaea and serves as a model for researching DNA

replication, the cell cycle, chromosomal integration, transcription, RNA processing and translation (She *et al*, 2001).

The aim of this chapter is the successful heterologous expression of soluble, active OYEs for downstream purification and characterisation. Successful expression of the two OYE homologues will allow the functional and structural characterisation of the enzymes in an attempt to broaden the available knowledge of the activities of yet uncharacterised OYE homologues, as well as add to the range of OYEs with potential for use as industrial biocatalysts.

3.2. MATERIALS AND METHODS

3.2.1. Identification of target OYEs for cloning and expression

Candidates for expression and characterisation were identified by means of tBLASTn (Basic Local Alignment Search Tool; Altschul *et al.*, 1997) with searches performed using the amino acid sequence of the enoate reductase from *Thermus scotoductus* SA-01 as query. Nucleotide sequences of targeted OYEs were download from the NCBI GenBank database and used as templates for primer design.

3.2.2. Homology modelling

Homology modelling of the OYE from *S. solfataricus* P2 was performed using Multi-Sources Threader (MUSTER) (http://zhanglab.ccmb.med.umich.edu/MUSTER; Wu & Zhang, 2008). Flavin cofactors and the iron-sulphur cluster were inserted using Yasara. Energy minimisation was performed in Yasara using the YASARA2 force field (Krieger *et al.,* 2009).

3.2.3. Bacterial strains and culture conditions

All bacterial strains and plasmids used in this study are listed in table 3.1. *Cupriavidus metallidurans* CH34 and *Sulfolobus solfataricus* P2 strains used throughout this study were obtained from the Leibniz Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) and cultured aerobically as per instruction (table 3.2).

Escherichia coli strain TOP10 (Invitrogen) was used as host for genetic manipulation, while strain BL21(DE3) (Lucigen) was used as host for expression of proteins. Rosetta-gami™ 2(DE3)pLysS strain served as a source of plasmid pLysSRARE2. *E. coli* strains were grown in Luria-Bertani (LB) medium at 37℃ with shaking (200 rpm). Ampicillin (100 µg.ml^{-1}), kanamycin (30 µg.ml^{-1}) or chloramphenicol (34 µg.ml^{-1}) was added to growth media for plasmid selection when required.

pGEM-T Easy (Promega) and pSMART (Lucigen) vectors were used for genetic manipulation. Plasmids pET-22b(+) and pET-28b(+) (Novagen) were used for expression of the proteins in *E. coli* strain BL21(DE3).

Table 3.1:   Strains and plasmids used in the study

Strain or plasmid	Description	Reference
Cupriavidus metallidurans CH34		ATCC 43123; DSM 2839
Sulfolobus solfataricus P2		ATCC 35092; DSM 1617
Escherichia coli TOP10	One Shot TOP 10 chemically competent cells F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec* A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K (Str^R) *end*A1 *nup*G	Invitrogen
Escherichia coli BL21(DE3)	*E. cloni* EXPRESS BL21(DE3) chemically competent cells F⁻ *omp*T *hsd*S$_B$(r$_{B-}$, m$_B^-$) *gal dcm* (DE3)	Lucigen
Escherichia coli Rosetta-gami™ 2 (DE3)pLysS	Rosetta-gami™ 2(DE3)pLysS chemically competent cells Δ(*ara-leu*)7697 Δ*lacX74* Δ*phoA Pvu*II *phoR araD139 ahpC galE galK rpsL* (DE3) F'[*lac⁺ lacI^q pro*] *gor522*::Tn*10 trxB* pLysSRARE2 (Cam^R, Str^R, Tet^R)	Novagen
pGEM®-T easy	Amp^R, *Lac*Z, f1 ori	Promega
pSMART	Kan^R; ori	Lucigen
pLysSRARE2	Cam^R, T7 lysozyme	Novagen
pET-22b(+)	Amp^R, T7 promoter/terminator, *Lac*I, f1 ori	Novagen
pET-28b(+)	Kan^R, T7 promoter/terminator, *Lac*I, N-terminal His-Tag and Thrombin configuration, f1ori	Novagen

Table 3.2:   Culture conditions of strains used in the study

Strain	Media	Culture conditions
Cupriavidus metallidurans CH34	DSMZ medium 1	30℃; 200 rpm shaking
Sulfolobus solfataricus P2	DSMZ medium 182	70℃; 200 rpm shaking

3.2.4. Construction of expression vectors

3.2.4.1.   *Total genomic DNA isolation*

Total genomic DNA (gDNA) of each strain in table 3.2 was isolated from 1 mL overnight culture using the Aquapure Genomic DNA kit (Bio-Rad). DNA pellets were resuspended in 50 µL elution buffer. DNA concentrations were determined using a NanoDrop spectrophotometer (NanoDrop Technologies).

3.2.4.2.   *DNA electrophoresis*

DNA integrity was evaluated on a horizontal agarose slab gel [0.8% (w/v)] containing 1.5 µg.mL^{-1} ethidium bromide in TAE buffer [Tris (40 mM), acetic acid (20 mM) and EDTA (1 mM)]. Electrophoresis was done for approximately 1 h at 90 V. DNA was visualised using a ChemiDox XRS Gel Documentation system (Bio-Rad). DNA fragment size was estimated based on electrophoretic mobility with the aid of GeneRuler™ DNA Ladder (Fermentas), MassRuler™ DNA Ladder Mix (Fermentas) or O'GeneRuler™ DNA Ladder Mix (Fermentas) standards during the same electrophoretic run.

3.2.4.3.   *Polymerase chain reaction (PCR) amplification of OYEs*

The complete OYE gene selected from *C. metallidurans* CH34 (*Cm*OYE) was amplified by PCR from genomic DNA using the Expand High Fidelity PCR System (Roche). The selected OYE gene from *S. solfataricus* P2 (*Ss*OYE) was amplified using the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific). All PCR reactions were performed in total reaction volume of 50 µL using a Thermal Cycler (PxE 0.2, Thermo Electron Corporation). For the Expand High Fidelity PCR System (Roche), reaction mixtures (50 µL) consisted of 10X Expand High Fidelity buffer (5 µL) with 15 mM MgCl$_2$ (5 µL), deoxynucleoside triphosphates (dNTP's) (0.2 mM each), Expand High Fidelity enzyme mix (3.5 U), 50 ng of gDNA, and 0.2 µM both the forward and reverse primers (primer sets for each

OYE gene is given in table 3.3). PCR reaction conditions consisted of an initial denaturing step at 95°C for 5 min, followed by 25 cycles of denaturing at 95°C (30 s), annealing for 40 s (see table 3.3 for annealing temperatures), and elongation at 72°C (1.5 min), with a final extension at 72°C for 10 min. For the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific), reaction mixtures (50 µL) consisted of 5X Phusion CG Buffer (5 µL) with 17.5 mM $MgCl_2$ (10 µL), dNTP's (0.8 mM), Phusion DNA polymerase (1U), 50 ng of gDNA and 0.3 µM of both the forward and the reverse primers (primer sets for each OYE gene is given in table 3.3). Reaction conditions consisted of an initial denaturing step at 94°C for 1 min, followed by 10 cycles of denaturing at 94°C (10 sec), annealing at 55°C (30 sec) and elongation at 68°C (1.5 min). This was followed by 25 cycles of denaturing at 94°C (15 sec), annealing at 55°C (30s) and elongation at 68°C (1.5 min + 20 sec for each successive cycle). A final elongation step of 7 min at 68°C was added to ensure the complete elongation of the amplified products.

PCR products were purified from agarose gel through excision of the gel bands and extraction using the Biospin Gel Extraction kit (BioFlux) according to the manufacturer's instructions.

Table 3.3:   Primer sequences used for PCR amplification of the OYEs

Amplicon/Primer	Sequence	T_m	Annealing temperature
CM_F1_Nde	5' <u>CAT ATG</u> CCT CAT CTC TTC GAT CCG TAC C 3'	59.9°C	62°C
CM_R1_Eco	5' <u>GAA TTC</u> TCA ACG CTG GCC GAA GTG CGC GT 3'	67.6°C	
SS_F2_Nde	5' <u>CAT ATG</u> GAT CTA TCG AAG CTA CTG GTT CCA ATA AG 3'	59.2°C	55°C
SS_R2_Sal	5' <u>GTC GAC</u> TTA ATA TGA CTT AAA GCT CCT ACC TAA CC 3'	59.1°C	

Underlined sequences indicate introduced restriction sites for *Nde*I, *Eco*RI and *Sal*I

CM_F1/R1   -   OYE from *C. metallidurans* CH34

SS_F2/R2   -   OYE from *S. solfataricus* P2

3.2.4.4. Ligations and transformations

The purified PCR products generated by means of the Expand High Fidelity PCR System (Roche) were ligated into the pGEM-T Easy vector (Promega, figure 3.1), while PCR products generated using the Phusion® High-Fidelity DNA Polymerase (Thermo Scientific) were ligated into the pSMART vector (Lucigen, figure 3.2) as per manufacturer's instructions. Constructs were subsequently proliferated in One Shot TOP10 *E. coli* competent cells (Invitrogen): 50 µL of *E. coli* One Shot TOP10 competent cells were allowed to thaw on ice. Two µL of the ligated plasmid mixture reactions were added to the cells and incubated on ice for 30 min. The cell-suspension was heat shocked at 42°C for 40 s and immediately returned to ice-water for 2 min. Two hundred and fifty µL of SOC medium (20 g.L^{-1} Bacto-tryptone, 5 g.L^{-1} Bacto-yeast extract, 0.5 g.L^{-1} NaCl, 2.5 mM KCl and 20 mM glucose) was added to the transformation reactions and incubated at 37°C for 1 h with gentle shaking. Aliquots were plated out on separate LB plates containing ampicillin (100 µg.ml^{-1}) or kanamycin (30 µg.mL^{-1}) as applicable and incubated overnight at 37°C.

Xmnl 2009
Scal 1890
Nael 2707
f1 ori
Amp^r
pGEM®-T Easy Vector
(3015bp)
lacZ
ori

T7 ↓	1 start
Apal	14
AatlI	20
Sphl	26
BstZI	31
Ncol	37
BstZI	43
Notl	43
SacII	49
EcoRI	52

Spel	64
EcoRI	70
Notl	77
BstZI	77
Pstl	88
Sall	90
Ndel	97
SacI	109
BstXI	118
Nsil	127
	141
↑ SP6	

T7 Transcription Start

5′ . . . TGTAA TACGA CTCAC TATAG GGCGA ATTGG GCCCG ACGTC GCATG CTCCC GGCCG CCATG
3′ . . . ACATT ATGCT GAGTG ATATC CCGCT TAACC CGGGC TGCAG CGTAC GAGGG CCGGC GGTAC
T7 Promoter
Apal   AatlI   Sphl   BstZI   Ncol

GCGGC CGCGG GAATT CGAT**T**3′(cloned insert) ATCAC TAGTG AATTC GCGGC CGCCT GCAGG TCGAC
CGCCG GCGCC CTTAA GCTA 3′**T**TAGTG ATCAC TTAAG CGCCG GCGGA CGTCC AGCTG
Notl   SacII   EcoRI   Spel   EcoRI   Notl   Pstl   Sall
BstZI   BstZI

SP6 Transcription Start

CATAT GGGA GAGCT CCCAA CGCGT TGGAT GCATA GCTTG AGTAT TCTAT AGTGT CACCT AAAT . . . 3′
GTATA CCCT CTCGA GGGTT GCGCA ACCTA CGTAT CGAAC TCATA AGATA TCACA GTGGA TTTA . . . 5′
Ndel   SacI   BstXI   Nsil   SP6 Promoter

Figure 3.1: Vector map and multiple cloning site of pGEM-T Easy indicating (among others) the ampicillin resistance gene, f1 origin of replication, *lacZ* coding sequence and the multiple cloning site under the T7 and SP6 promoters (Marcus *et al*, 1996; diagram from manufacturer's manual).
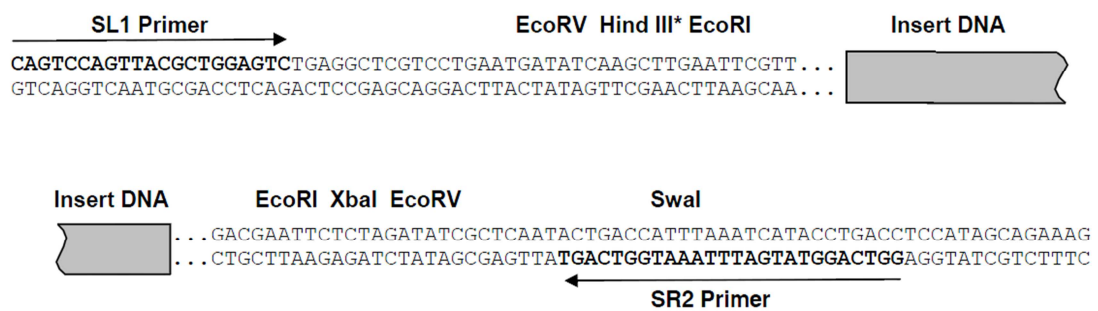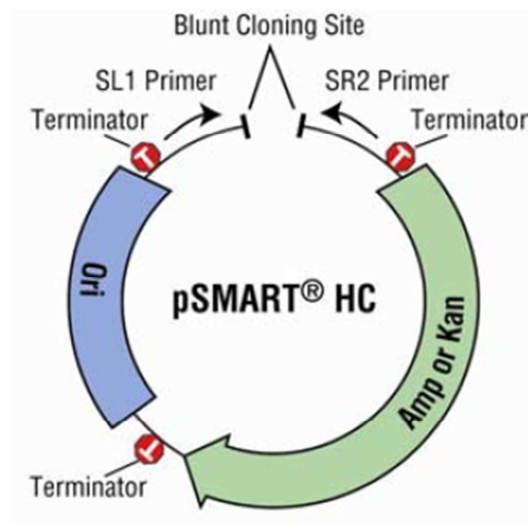
Figure 3.2: Vector map of pSMART (1788 bp) indicating the kanamycin resistance gene, origin of replication and the multiple cloning site. Sequence of the cloning region shows the primer binding sites and restriction enzyme sites (Godiska *et al.*, 2001; diagram from manufacturer's manual).

3.2.4.5.   Constructs for expression in E. coli

Single colonies of the above-mentioned transformed *E. coli* One Shot TOP10 competent cells were inoculated into 5 mL LB media containing ampicillin (100 µg.ml^{-1}) or kanamycin (30 µg.mL^{-1}) as applicable and incubated overnight at 37°C. Plasmids were isolated using the BioSpin Plasmid DNA Extraction kit (BioFlux) and double digested with restriction enzymes - *Nde*I (0.5 U.µL^{-1}, Fermentas) and *Eco*RI (0.5 U.µL^{-1}, Fermentas) for *Cm*OYE; *Nde*I and *Sal*I (0.5 U.µL^{-1}, Fermentas) for *Ss*OYE - at 37°C (Buffer O, 3 h) for ligation into the vectors pET22b(+) (figure 3.3) and pET28b(+) (figure 3.4) (similarly digested). Double-digested constructs were evaluated by agarose gel electrophoresis as previously described. Suitable inserts and digested pET22b(+) and pET28b(+) vectors were cleaned from agarose gel using the Biospin Gel Extraction kit. Cohesive end ligations were performed in 1:1 molar ratio, overnight at 16 °C with T4 DNA ligase (1.5 Weiss U.µL^{-1}, Fermentas).   To proliferate pET22/28-OYE constructs, ligation mixtures were again transformed into TOP10 *E. coli* and positive clones were identified through plasmid isolation, restriction digestion and evaluation by agarose gel electrophoresis as described above.

3.2.4.6.   Sequencing

Plasmid inserts in pGEM-T Easy, pET22b(+) and pET28b(+) were sequenced using a 3130*xl* Genetic Analyser (Hitachi). Sequencing reactions were performed using the BigDye (v3.1) dye terminator cycle sequencing kit (Applied Biosystems) as per manufacturer's instructions, using the universal T7 promoter and SP6 promoter (pGEM-T Easy) or T7 terminator primers (pET vectors).

Figure 3.3: Vector map of pET22b(+) indicating (among others) the ampicillin resistance gene, ColE1 origin of replication, *lacI* coding sequence and the multiple cloning site under the T7 promoter. Sequence of the cloning region shows the primer binding sites and restriction enzyme sites (obtained from pET System Manual 11th Edition; diagram constructed from sequence data using Geneious).

Figure 3.4: Vector map of pET28b(+) indicating (among others) the kanamycin resistance gene, ColE1 origin of replication, *lacI* coding sequence and the multiple cloning site under the T7 promoter. Sequence of the pET28b(+) cloning region shows the ribosome binding site and the configuration for the N-terminal His-Tag and thrombin cleavage site fusion (obtained from pET System Manual 11th Edition; diagram constructed from sequence data using Geneious).

3.2.5. Expression of the selected OYEs

The pET22/28-OYE constructs containing OYE gene from *C. metallidurans* CH34 was transformed into *E. coli* BL21(DE3) competent cells (Lucigen) for expression. For expression of the OYE from *S. solfataricus* P2, pET22/28-OYE constructs were transformed into *E. coli* BL21(DE3) competent cells (Lucigen) previously transformed with plasmid pLysSRARE2. pLysSRARE2 had been isolated from *Escherichia coli* Rosetta-gami™ 2 (DE3)pLysS (Novagen) using the BioSpin Plasmid DNA Extraction kit (BioFlux). Positive OYE clones for expression were identified through selection on LB-plates containing kanamycin (30 µg.mL^{-1}) or ampicillin (100 µg.mL^{-1}) in addition to 34 µg.mL^{-1} chloramphenicol for selection of the pLysSRARE2 plasmid where applicable, and inoculated into 100 mL volumes of LB-medium also containing the appropriate antibiotic. Cells were incubated with shaking at 200 rpm until an OD$_{600nm}$ of 0.8 – 1 was reached. IPTG (1 mM final concentration) was added as inducer, after which cells were incubated for an additional 4 h. One mL samples of the cultures were taken 0, 2 and 4 hours after induction of expression. Cells were harvested through centrifugation (5 000 x g, 15 min) and washed using 20 mM MOPS-NaOH (pH 7).

3.2.6. Analysis of expression

3.2.6.1. Harvesting and cell disruption

Harvested cells were resuspended in 20 mM MOPS-NaOH (pH 7), with approximately 1 g cells (wet weight) in 10 mL. Cells were broken using a One Shot Cell Disruption system (Constant Systems Ltd), with 7 mL aliquots of cell suspension subjected to 3 cycles of cell disruption at 30 KPSI. Unbroken cells and debris were removed through centrifugation (5 000 x g for 20 min). The soluble fraction (cytoplasm) was separated from the insoluble fraction (membranes) through ultracentrifugation (100 000 x g for 90 min).

3.2.6.2. Analysis of expression through gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using the "Mighty Small" miniature slab gel electrophoresis unit, SE 200 (Hoefer Scientific Instruments). The protocol used was as described by Laemmli (1970), using a 10% resolving gel and a 4% stacking gel. Precision Plus Protein All Blue Standards (Bio-Rad) and PageRuler™ Unstained Broad Range Protein Ladder (Fermentas) was used as molecular weight markers in SDS-PAGE and proteins were visualized by staining polyacrylamide gels with Coomasie brilliant blue R250 (Bio-Rad) and destained as described by Fairbanks *et al.* (1971).

3.2.6.3. Analysis of activity of expressed OYE

Integrity of enzyme preparations was determined by continuous assay by measuring oxidation of NADPH in the presence of 2-cyclohexenone, using the above-mentioned soluble fractions of crude cell extracts, for 5 minutes at 23°C. Continuous assays were performed aerobically in 1 mL reaction mixtures containing 20 mM MOPS-NaOH buffer (pH 7), 0.3 mM NADPH, 10 mM 2-cyclohexenone and 0.05 mL of the crude enzyme preparation. Absorbance was measured using a Cary 300 Bio UV-Visible spectrophotometer (Varian) at 340 nm. An extinction coefficient of 6.22 mM.cm^{-1} was used for NADPH at 340 nm (Jablonski & DeLuca, 1977).

3.3. RESULTS AND DISCUSSION

3.3.1. Identification of OYEs for cloning and expression

tBLASTn (Altschul *et al.,*1997) searches returned two OYE genes each from the whole genome sequences of *C. metallidurans* CH34 and *S. solfataricus* P2. From these pairs of OYE genes, one OYE gene from each strain was selected for characterisation.

Reasons for selection of the two OYE homologues were twofold: firstly, the cloning and expression (and subsequent characterisation) of new, uncharacterised OYE homologues is an attempt to provide the industry with new OYE homologues that exhibit potential for use as biocatalysts. Secondly, characterisation of these two OYE homologues will provide insight into the functionalities of OYE homologues with respect to the subclass in which they belong, according to sequence-based analysis performed in chapter 2.

The OYE gene from *C. metallidurans* CH34 (*Cm*OYE) was selected, as this enzyme has a His-pair in the catalytic site responsible for substrate binding, combined with a Cys residue corresponding to the flavin modulator. According to the subclasses identified in chapter 2, these structural characteristics group *Cm*OYE as a member of the subclass B2. This subclass is dominated by OYEs from thermophiles, including the OYEs from *T. scotoductus* SA-01 (CrS; Opperman, *et al.*, 2010) and *Geobacillus kaustophilus* (Schittmayer *et al.*, 2011), as well as YqjM from *Bacillus subtilis* (Kitzing *et al.*, 2005) and XenA from *Pseudomonas putida* (Griese *et al.*, 2006). BLAST searches revealed that *Cm*OYE shares 43% amino acid sequence identity with XenA, 40% sequence identity with YqjM and 35% sequence identity with CrS. Characterisation of *Cm*OYE will be valuable, as C*m*OYE will add to the small number of mesophilic OYE homologues characterised in this subclass.

The second OYE, from the hyperthermophilic archaeon *S. solfataricus* P2 (*Ss*OYE), was chosen both for its stability at high temperatures and for its non-classical amino acid sequence. As with *Cm*OYE, *Ss*OYE has a pair of histidine residues responsible for the binding of substrate in the catalytic site. However, this His-pair arrangement is combined with an Ile residue corresponding to the flavin modulator residue. According to the subclasses identified in chapter 2, *Ss*OYE groups within the subclass B4$_A$. The amino acid sequence of *Ss*OYE shares 35% sequence identity with XenA, 38% sequence identity with YqjM and 36% sequence identity to CrS. Cloning, expression and characterisation of *Ss*OYE will be valuable, as no such information has been reported for any members of this subclass.

The unconventional structure of *Ss*OYE, however, goes beyond the presence of the Ile residue in the place of the flavin modulator. Construction of a homology model (figure 3.6) from the translated nucleotide sequence in MUSTER (Wu & Zang, 2008) revealed that the OYE homologue exhibits a two-domain architecture, bearing a striking resemblance to the crystal structure of 2,4-dienoyl-CoA reductase from *E. coli* (figure 3.5) (Hubbard *et al.*, 2003). 2,4-dienoyl-CoA reductase is an iron-sulphur flavoenzyme necessary for the beta-oxidation of unsaturated fatty acids with even-numbered double bonds. The enzyme contains both FMN and FAD, as well as a 4Fe-4S cluster and requires NADPH as an electron source. The reaction is thought to involve an electron transfer mechanism involving both flavins and the iron-sulphur cluster: the reaction is initiated by hydride transfer from NADPH to FAD from which electrons are subsequently transferred (one at a time) to the FMN cofactor via the 4Fe-4S cluster, allowing the fully reduced FMN to be oxidised via hydride transfer to the substrate (Hubbard *et al.*, 2003). In both *Ss*OYE and 2,4-dienoyl-CoA reductase, the structure of the N-terminal domain corresponds to a TIM-barrel structure with bound FMN, as is characteristic of OYEs (Barna *et al.*, 2001; Breithaupt *et al.*, 2001; Kitzing *et al.*, 2005; Fox & Karplus, 1994; Hubbard *et al.*, 2003). It must be noted that, although the N-terminal domain of 2,4-dienoyl-CoA exhibits the TIM-barrel structure, the active site architecture of the enzyme differs vastly from that of OYE homologues (Hubbard *et al.,* 2003). Also similar to both enzymes is the presence of the 4Fe-4S cluster (Hubbard *et al.,* 2003). The C-terminal domain of 2,4-dienoyl-CoA reductase is a flavodoxin-like domain which is responsible for the binding of the FAD cofactor. The homology model of *Ss*OYE revealed a C-terminal domain corresponding to a truncated 2,4-dienoyl-CoA reductase C-terminal domain, binding FAD.
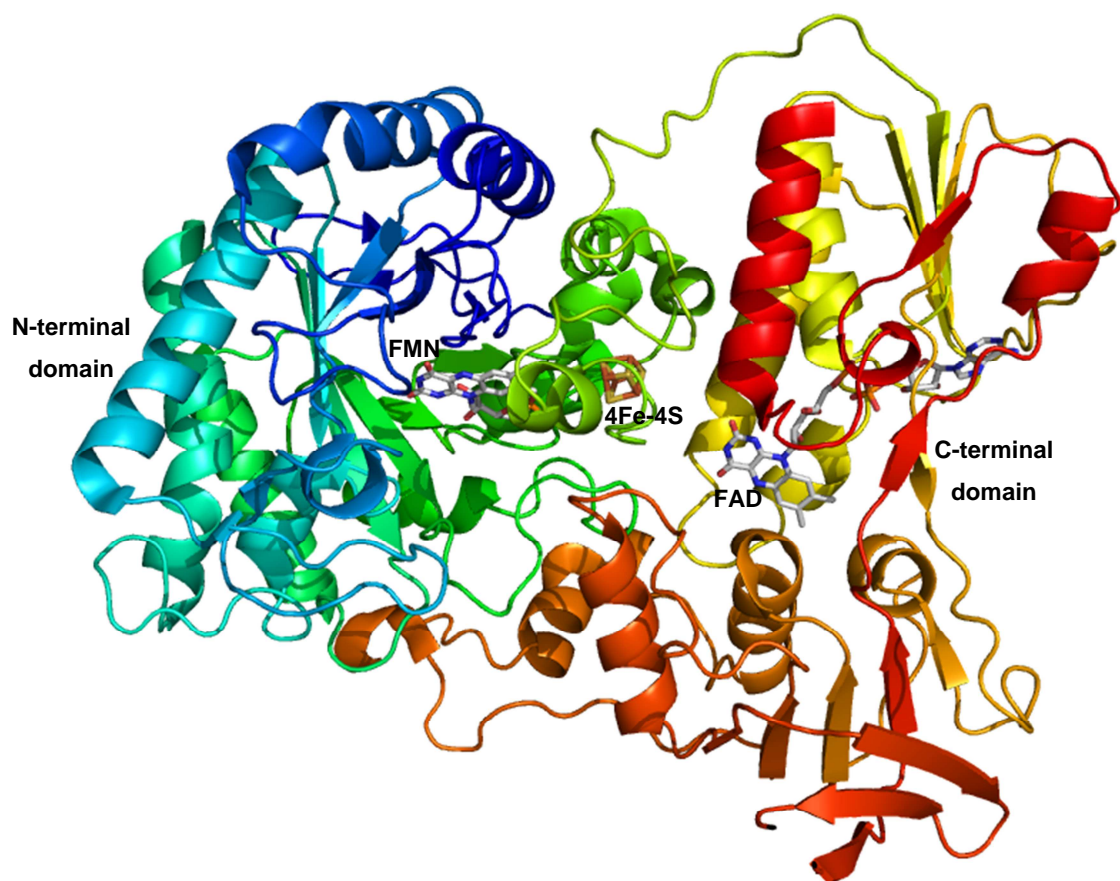
Figure 3.5: Diagram of 2,4-dienoyl-CoA reductase showing the N-terminal, 4Fe-4S cluster and C-terminal domains. Diagram constructed in PyMOL from PDB file of the crystal structure (Hubbard *et al.*, 2003).
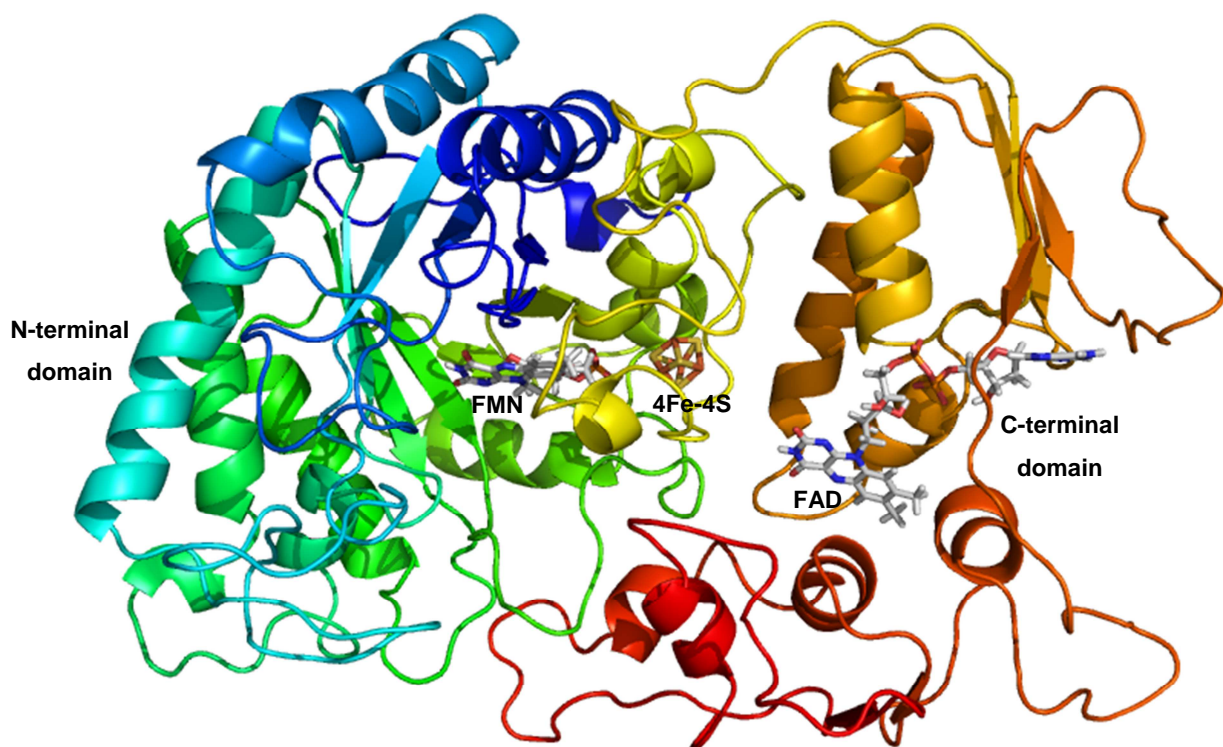
Figure 3.6: Homology model for *Ss*OYE showing the N-terminal, 4Fe-4S cluster and C-terminal domains. Model constructed using MUSTER. Diagram constructed in PyMOL.

3.3.2. Construction of expression vectors

Total genomic DNA was isolated successfully from *C. metallidurans* CH34 and *S. solfataricus* P2 using the Aquapure Genomic DNA kit. Agarose gel electrophoresis (figure 3.7) revealed clear bands larger than 10k bp in size, indicating the presence of high-quality gDNA with very little shearing. Nanodrop spectrophotometer readings confirmed the integrity and high concentrations of the isolated gDNA (table 3.4), with A_{260}/A_{280} ratios indicative of pure DNA with little contaminating protein or other agents.
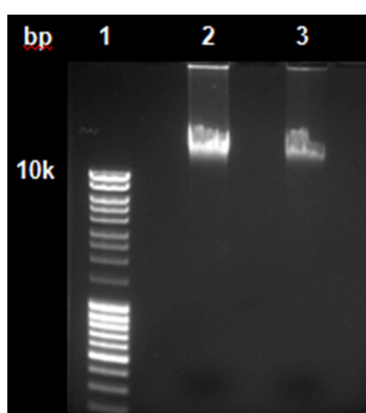
Figure 3.7: Total genomic DNA from *S. solfataricus* P2 (lane 2) and *C. metallidurans* CH34 (lane 3). Lane 1, MassRuler DNA ladder mix.

Table 3.4: Concentration and integrity of total genomic DNA isolated from strains

Strain	Genomic DNA concentration	$A_{260}{:}A_{280}$
Cupriavidus metallidurans CH34	415 ng.μL^{-1}	1.98
Sulfolobus solfataricus P2	230 ng.μL^{-1}	1.90

The complete ORFs of both OYEs were amplified from total genomic DNA using the Expand High Fidelity PCR System (Roche) and Phusion® High-Fidelity DNA Polymerase (Thermo Scientific). For the later directional cloning of amplified ORFs into pET-vectors, primers (table 3.3) were designed to incorporate a *Nde*I restriction site at the 5' end of both OYE ORFs, with an *EcoR*I restriction site at the 3' end of *Cm*OYE ORF and a *Sal*I restriction site at the 3' end in the case of *Ss*OYE. Agarose gel electrophoresis confirmed that PCR amplification had yielded products of the expected sizes (figure 3.8). Amplification of the *Cm*OYE gene yielded clear bands of ~1 000 bp, corresponding to the expected 1 116 bp gene. Similarly, clear bands of ~1 600 bp were obtained for the amplification of the *Ss*OYE gene (corresponding to the 1 668 bp gene).
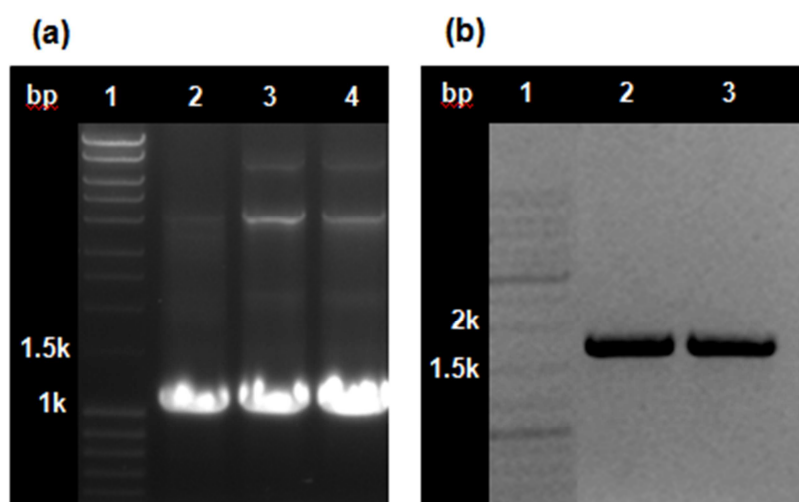
Figure 3.8:   Agarose gel electrophoresis of the PCR amplified OYEs from **a)** *C. metallidurans* CH3 and **b)** *S. solfataricus* P2. Lanes 2-4 PCR amplified products. **a)** Lane 1, MassRuler DNA ladder mix and **b)** lane 1, GeneRuler DNA ladder mix.

Expand High Fidelity PCR System, being a mixture of thermostable Taq DNA polymerase and Tgo DNA polymerase (a thermostable polymerase with 3'-5' exonuclease or "proofreading" activity), yields PCR fragments with 3' dA overhangs. As a result, products obtained by amplification of the *Cm*OYE gene were ligated into vector pGEM-T Easy after excision and purification from agarose gel. In contrast, PCR products obtained using the Phusion® High-Fidelity DNA Polymerase (a polymerase with 5'-3' DNA polymerase activity and 3'-5' exonuclease activity) are blunt-ended in nature. As a result, PCR products obtained for the amplification of the *Ss*OYE gene were excised and purified from agarose gel and subsequently ligated into vector pSMART through the blunt-end cloning protocol described by the manufacturer.

pGEM-OYE and pSMART-OYE constructs were evaluated by agarose gel electrophoresis (figure 3.9) after being subjected to double digests with *Nde*I and *EcoR*I (pGEM-OYE constructs, for *Cm*OYE) or *Nde*I and *Sal*I (pSMART-OYE constructs, for *Ss*OYE). In the case of pGEM-OYE constructs, two bands were clearly visible: one band, between 1 000 bp and 1 200 bp, corresponding to the amplified C*m*OYE gene and a second, ~3 000 bp band corresponding to the pGEM-T Easy vector backbone. The presence of these two clear bands confirmed that ligation into pGEM-T Easy vector was successful. Evaluation of pSMART-OYE constructs also confirmed successful ligation, with agarose gel electrophoresis revealing two bands between 1 500 bp and 2 000 bp – the larger corresponding to the 1 788 bp pSMART vector backbone and the other to the 1 668 bp OYE gene.
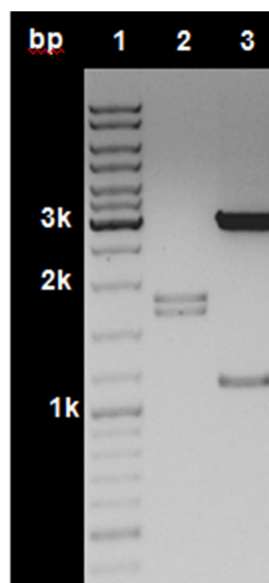
Figure 3.9: Agarose gel electrophoresis of double-digested pSMART-*Ss*OYE (lane 2) and pGEM-*Cm*OYE (lane 3) constructs. Lane 1, GeneRuler DNA ladder mix.

Following confirmation of successful ligation, inserts were excised from the pGEM-OYE and pSMART-OYE constructs using the same restriction enzyme combinations as before. pET22b(+)and pET28b(+) were digested with the same restriction enzyme combinations to yield compatible cohesive ends for cloning of OYE genes into the expression vectors. When introduced, pET22b(+) would yield unmodified proteins, while pET28b(+) would yield a poly(6)histidine tag and a thrombin cleavage site fused to the N-terminus of the expressed protein.

As with the pGEM-OYE and pSMART-OYE constructs, pET22/28-OYE constructs were evaluated by means of double digestion with the appropriate restriction enzymes and subsequent agarose gel electrophoresis (figure 3.10). Released inserts and digested pET vectors were subjected to agarose

gel electrophoresis, followed by excision and purification from agarose gel and subsequent ligation. For each pET22/28-OYE construct, agarose gel electrophoresis confirmed successful ligation into pET vectors as revealed by two clearly visible bands. Bands between 5 000 bp and 6 000 bp corresponded to the expression vector backbone. In each case, the second band corresponded to the expected OYE gene: an approximately 1 700 bp band in the case of *Ss*OYE and a band between 1 000 bp and 1 200bp in the case of *Cm*OYE. In addition, sequence analysis of plasmid inserts in pET22b(+) and pET28b(+) revealed that no mutations were present and that directional cloning into pET vectors was successful.
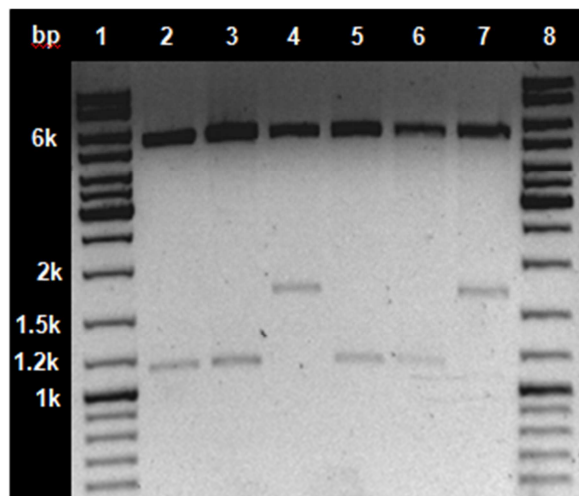
Figure 3.10: Agarose gel electrophoresis of the pET22/28-OYE constructs. Lanes 2-4 represent pET22-OYE constructs for *Cm*OYE (lanes 2-3) and *Ss*OYE (lane 4). Lanes 5-7 represent pET28-OYE constructs for *Cm*OYE (lanes 5-6) and *Ss*OYE (lane 7). Lanes 1 and 8, MassRuler DNA ladder mix.

3.3.3. Heterologous expression of OYEs

pET22/28-OYE constructs containing *Cm*OYE and *Ss*OYE were expressed in *E. coli* BL21(DE3) chemically competent cells. The B strain of *E. coli* is suitable for use in the expression of non-toxic heterologous genes. This strain lacks two key proteases (*lon* protease and the outer membrane protease OmpT), reducing the degradation of heterologous proteins expressed in the cells (Studier, 1991). The strain contains the DE3 lysogen that carries the gene for T7 RNA polymerase under control of the *lac*UV5 promoter, allowing expression of the T7 RNA polymerase to be induced in the presence of IPTG.

SDS-PAGE analysis (figures 3.11-3.12) revealed that heterologous expression of both modified and unmodified protein was successful for *Cm*OYE, in both total crude cell extracts and soluble fractions. In each case, SDS-PAGE analysis revealed bands between 37 kDa and 50 kDa not visible in cell extracts of *E. coli* cells transformed with expression vectors containing no inserts. These visible bands correspond to the expected ~40 kDa unmodified and ~42 kDa modified (with addition of ~2 kDa His$_6$-tag and thrombin cleavage site) OYEs. In the case of *Ss*OYE, no expression of either expected unmodified ~62 kDa or modified ~65 kDa protein was observed.
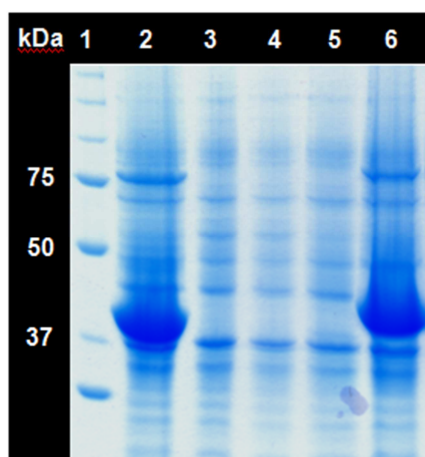
Figure 3.11:   SDS-PAGE analysis of total crude extracts showing overproduction of *Cm*OYE (lane 2) and *Ss*OYE (lane 3) using pET22-OYE constructs, 4 hours after induction with IPTG. Lanes 4-6 represent total crude extracts showing overproduction of OYEs in *E. coli* using pET28-OYE constructs, 4 hours after induction with IPTG: *Ss*OYE (lane 5) and *Cm*OYE (lane 6). Lane 4 represents crude extracts of *E. coli* transformed with pET28b(+) plasmid without inserts. Lane 1, Precision Plus protein standard.
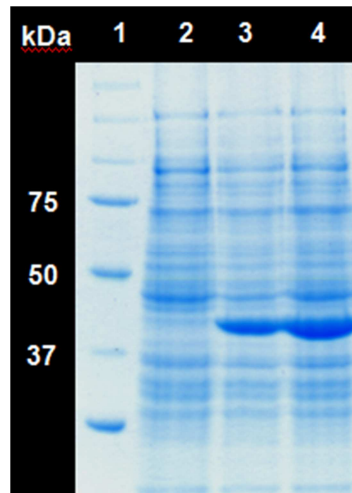
Figure 3.12: SDS-PAGE analysis showing overproduction of *Cm*OYE in soluble cell fractions of *E. coli* extracts 4 hours after induction with IPTG. *E. coli* cells were transformed with pET22-OYE (lane 3) and pET28-OYE (lane 4) constructs. Lane 2 represents soluble cell fractions of *E. coli* transformed with pET28b(+) plasmid without inserts. Lane 1, Precision Plus protein standard.

Closer examination of the nucleotide sequence of the ORF of *Ss*OYE was necessary due to the unsuccessful expression of the gene in *E. coli* BL21(DE3) cells. Since sequence analysis of plasmid inserts had previously confirmed that no mutations were present in the ORFs, attention was directed towards the possibility of rare codons in the gene sequence. Analysis of the nucleotide sequence using the Codon Usage online tool (http://www.bioinformatics.org/SMS/index.html) revealed rare codons in the gene sequence (figure 3.13) coding for leucine (CTA), isoleucine (ATA), arginine (AGA) and glycine (GGA).

```
>Sulfolobus solfataricus P2 OYE2
ATGGATCTATCGAAGCTACTGGTTCCAATAAGAGTTGGTGATGTGGTCTTAAAGAACAGA
ATAGCCATGTCACCAATGATTAGCAATCTAGGTACACCTGAGGGTTACCCGAGCGATGCT
CATATAGCTTACTTAGCTGAGAGAGCTAAGGGAGGAGTTGGGTTAATAATAACGGAGTAC
ACTTACGTTAACCACGTAGATGCTAGGGGGTCTGTTAACGAATTGGGGATGTATTCAGAT
GAGCTAACACCAAAGTTTACGAGGTTGACTGAGTTAATTCACGCTTTGGGAAGTAAGATT
TTCGTGCAGTTAGTTCACGTTGGCAGGAAGACTAGGAAGGATATAATTTGGGGTAACAAA
CCCATAGCTCCTTCACCAATTCCAATAATGGACGAAGTAAGGGAAATGACTAAAGAGGAT
ATAGAGAGGGTTAAAAACGATTTCATAAACGCAGCAATAAGGGCTAAGAGGGCTGGATTT
```

Figure 3.1: Nucleotide sequence of the ORF of *Ss*OYE showing rare codons coding for leucine (blue), isoleucine (red), arginine (green) and glycine (yellow).

Due to the presence of these rare codons, expression of *Ss*OYE was repeated, with *E. coli* BL21(DE3) competent cells previously transformed with plasmid pLysSRARE2 serving as expression host. pLysSRARE2, a chloramphenicol-resistant plasmid, consists of a CamR plasmid (pACYC184) that carries the gene for T7 lysozyme (for inhibiting T7 RNA polymerase), as well as tRNA genes for codons rarely used in *E. coli* - AUA, AGG, AGA, CUA, CCC, GGA and CGG (Novy *et al.*, 2001). Expression was induced through the addition of IPTG, as previously described. SDS-PAGE analysis (figure 3.14) of cell extracts revealed that heterologous expression of both modified and unmodified protein was successful for *Ss*OYE. For both total and soluble cell fractions, clear bands between 50 kDa and 70 kDa were observed, corresponding to the expected ~63 kDa unmodified and ~65 kDa modified *Ss*OYE.
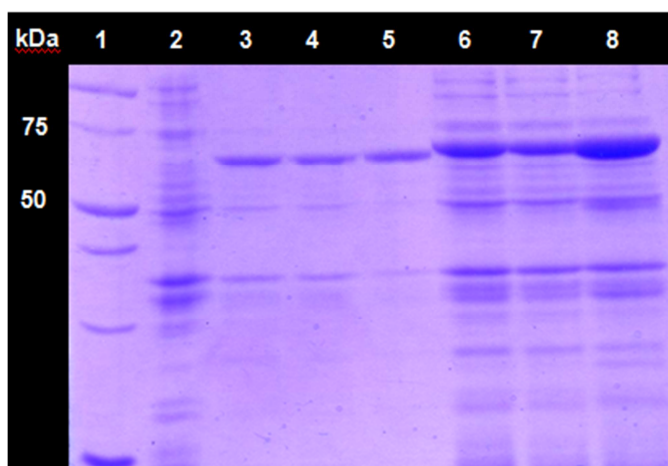
Figure 3.14: SDS-PAGE analysis showing overproduction of *Ss*OYE in total and soluble cell fractions of *E. coli* extracts 4 after induction with IPTG. *E. coli* cells were transformed with pET22-OYE (lanes 3-5) and pET28-OYE (lanes 6-8) constructs. Expression was visible for both pET22-OYE and pET28-OYE constructs 2 hours (lanes 3 and 6, respectively) and 4 hours (lanes 4 and 7, respectively) after induction with IPTG. Overproduction of *Ss*OYE was also visible in the soluble cell fraction of *E. coli* for both pET22-OYE (lane 5) and pET28-OYE (lane 8) constructs. Lane 2 represents soluble cell fraction of *E. coli* transformed with pET28b(+) plasmid without inserts 4 hours after induction with IPTG. Lane 1, Precision Plus protein standard.

3.3.4. Analysis of activity of expressed OYEs

The integrity of the overexpressed OYEs present in the soluble cell fractions obtained in 3.3.3 was determined by investigating the oxidation of NADPH aerobically in the presence of the ubiquitous OYE substrate 2-cyclohexenone. Oxidation of NADPH was monitored by measuring the absorbance of reaction mixtures spectrophotometrically at 340 nm. Negative controls consisted of reactions performed in the absence of crude extract, as well as reactions containing crude extracts of *E. coli* transformed with pET28b(+) plasmids. Since molecular oxygen also acts as an electron acceptor, re-oxidising the enzyme-bound flavin (Karplus *et al.*, 1995), blank rates for the oxidation of NADPH in the absence of 2-cyclohexenone were also determined. These blank rates represent the oxidation of the flavoenzyme by molecular oxygen in the aerobic environment.

Figure 3.15 shows the progress curves obtained during the continuous assay for modified *Ss*OYE from pET28-OYE constructs, using crude cell extracts. The change in absorbance of reaction mixtures at 340 nm is shown over a period of 5 minutes. Gradients for the linear section of each curve were used to determine the rate of NADPH oxidation in each reaction per minute. Due to the use of crude enzyme extract, specific activities were not calculated and assays served merely as a preliminary test to confirm the integrity of the heterologously expressed enzymes.

Negative control reactions exhibited minimal oxidation of NADPH, with 6.17×10^{-4} µmol NADPH oxidised per minute in reactions without crude enzyme extract, and NADPH oxidation rate of 2.15×10^{-3} µmol.min^{-1} for reactions containing crude extracts of *E. coli* transformed with intact pET28 plasmid. The higher oxidation rate in the latter negative control reaction could be attributed to NADPH-oxidase activity of native *E. coli* proteins present in the crude enzyme extracts. Reactions representing the oxidation of heterologous *Ss*OYE by molecular oxygen in the absence of the substrate 2-cyclohexenone exhibited considerably higher NADPH oxidation rates compared to the negative controls (4.07×10^{-3} µmol.min^{-1}), confirming that heterologous *Ss*OYE catalyses the oxidation of NADPH with the employment of molecular oxygen as final electron acceptor. The enzyme's ability to accept 2-cyclohexenone as substrate was confirmed, as reactions containing modified *Ss*OYE exhibited oxidation of NADPH at a net rate of 1.08×10^{-2} µmol.min^{-1}.

The integrity of unmodified *Ss*OYE, as well as of modified and unmodified *Cm*OYE was confirmed in a similar manner, with reactions performed exhibiting rapid net oxidation of NADPH compared to negative controls and blank rates.
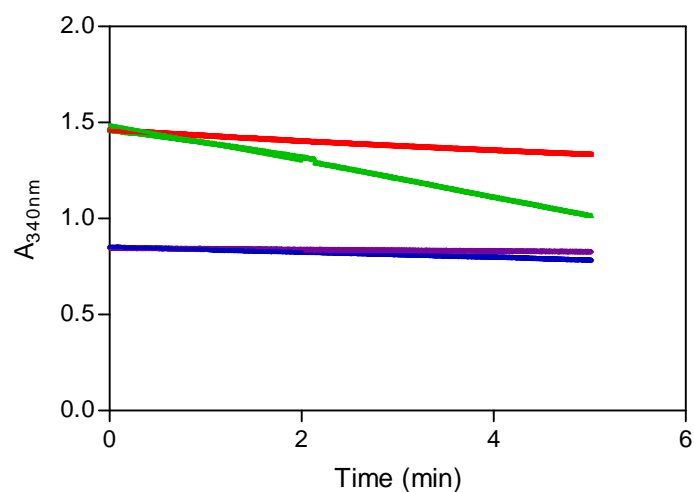
Figure 3.15: NADPH oxidation by cytoplasmic fractions of *E. coli* containing *Ss*OYE with 2-cyclohexenone as oxidant, involving the spectrophotometric monitoring of reaction mixtures at 340 nm. Reactions performed in the absence of 2-cyclohexenone (■), served as blanks. Negative control reactions performed in the absence of crude extract (♦), as well as with crude enzyme extract from *E.coli* cells transformed with intact pET28b(+) plasmids (●), are also shown. Activity towards 2-cyclohexenone monitored for OYE obtained from pET28-OYE constructs (▲) is also shown.

3.4. Conclusions

Two targeted OYEs were effectively identified using tBLASTn using the amino acid sequence of an OYE from a known thermophile – the enoate reductase from *T. scotoductus* SA-01 - as query. The first of these, OYE from *C. metallidurans* CH34, is an OYE from a mesophilic betaproteobacterium, and falls within the "thermophilic-like" subclass of OYEs, similar to YqjM and XenA. The second, an OYE from *S. solfataricus* P2, is an OYE from a hyperthermophilic archaeon. This OYE, belonging to a yet uncharacterised subclass of OYEs, was shown by means of homology modelling to have an unconventional structure.

With the use of high-fidelity, proofreading DNA polymerases, PCR amplification of the selected OYE genes from total genomic DNA of *C. metallidurans* CH34 and *S. solfataricus* P2 was successful, with complete ORFs correctly amplified in each case. Both unmodified and modified (through N-terminal poly-histidine fusion) OYE from *C. metallidurans* CH34 and *S. solfataricus* P2 were heterologously overproduced in *E. coli* as soluble, active enzymes. Activity of expressed protein was confirmed upon observed enzymatic oxidation of NADPH when utilizing the ubiquitous OYE substrate, 2-cyclohexenone, as final electron acceptor.

CHAPTER 4

PURIFICATION AND CHARACTERISATION OF OYES FROM
C. METALLIDURANS CH34 AND *S. SOLFATARICUS* P2

4.1. INTRODUCTION

The old yellow enzyme family has remained a promising candidate for use as industrial biocatalysts due to its ability to catalyse the reduction of numerous commercially useful substrates (Toogood *et al.*, 2010; Williams & Bruce, 2002). The *in vitro* functional and structural characterisation of the members of the OYE family have been revealed as vital tools in the search for new biocatalysts (Bougioukou *et al.*, 2009; Padhi *et al.*, 2009, Williams & Bruce, 2002).

Structural characterisation of OYEs is an important step in the search for new biocatalysts, providing significant insight into the catalytic mechanism and substrate specificity across the OYE family as a whole (Williams & Bruce, 2002). In addition, the wealth of structural information on the binding of substrates and inhibitors in the active site of OYEs play an important role in allowing the possible prediction of both substrate selectivity and enantiopreference of the product (Oberdorfer *et al.*, 2011; Adalbjörnsson *et al.*, 2010; Toogood *et al.*, 2010). A structure-driven approach also enables the targeting of specific residues for mutagenesis studies (Bougioukou *et al.*, 2009; Padhi *et al.*, 2009). Additions to the ever-increasing substrate range of the OYE family have often been as a result of attempts to pinpoint the enzyme family's physiological function (Williams & Bruce, 2002). While care must be taken when drawing parallels between *in vitro* enzyme function and physiological function (Jeffery, 2009), functional characterisation of these enzymes serves to continuously shed light on the potential of OYEs as industrial biocatalysts (Padhi *et al.*, 2009, Williams & Bruce, 2002; Toogood *et al.*, 2010).

This chapter describes the functional characterisation (optimum reaction conditions, kinetic parameters and substrate scope) of the OYEs from *C. metallidurans* CH34 and *S. solfataricus* P2. The efficacy of a light-driven cofactor regeneration approach in conjunction with these two OYEs will also be investigated. Lastly, the crystallization of OYE from *C. metallidurans* CH34 and subsequent collection of X-ray diffraction data will be described.

4.2. MATERIALS AND METHODS

4.2.1. Preparation of recombinant OYE for characterisation

4.2.1.1.   Heterologous expression

OYEs from *S. solfataricus* P2 (*Ss*OYE) and *C. metallidurans* CH34 *(Cm*OYE) were over-produced, using the pET28-OYE constructs that resulted in positive expression in chapter 3, to yield protein with a poly(6)histidine tag and a thrombin cleavage site fused to the N-terminus. As described in sections 3.2.4. and 3.3.3., OYE from *C. metallidurans* CH34 was expressed using *E. coli* BL21(DE3) (Lucigen) cells as expression host, whereas OYE from *S. solfataricus* P2 was expressed using *E. coli* BL21(DE3) cells containing the pLysSRARE2 plasmid (Novagen). For each OYE, expression was performed in thirty 500 ml Erlenmeyer flasks containing 100 mL cell culture each for a total culture volume of 3 L. *E. coli* strains were grown in ZYP5052 auto-induction media [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 25mM $(NH_4)_2SO_4$, 0.5% (w/v) glycerol, 0.05% (w/v) glucose, 0.2% (w/v) α-lactose, 2 mM $MgSO_4$; Studier, 2005]. Cultures were incubated at 25°C (200 rpm) for 24 h. Cells were harvested through centrifugation (5 000 x *g*, 15 min) and washed twice using 20 mM MOPS-NaOH (pH 7).

4.2.1.2.   Harvesting and cell disruption

Harvested cells were resuspended in 20 mM MOPS-NaOH (pH 7.4), with approximately 1 g cells (wet weight) in 10 mL buffer. Cells were broken using a One Shot Cell Disruption system (Constant Systems Ltd), with 7 mL aliquots of cell suspension subjected to 3 cycles of cell disruption at 30 KPSI. Unbroken cells and debris were removed through centrifugation (5 000 x *g* for 20 min). The soluble fraction (cytoplasm) was separated from the insoluble fraction (membranes) through ultracentrifugation (100 000 x *g* for 90 min).

4.2.2. Purification of heterologously expressed OYEs

4.2.2.1.   Immobilised metal-affinity chromatography (IMAC)

Recombinant N-terminally His_6-tagged OYEs were firstly purified through immobilized metal-affinity chromatography (IMAC). Soluble fractions obtained in 4.2.1.2 were loaded onto HisTrap FF columns (5 mL, GE Healthcare) equilibrated with 20 mM MOPS-NaOH (pH 7.4) containing imidazole (20 mM) and NaCl (0.5 M). Unbound proteins were eluted (5 mL.min^{-1}) using the same buffer. Bound proteins were then eluted in the same buffer using an increasing linear gradient (100

mL) of imidazole up to 0.5 M. Fractions containing the characteristic yellow colour (due to the presence of the flavin cofactor) were pooled for subsequent purification.

4.2.2.2. Size-exclusion chromatography and desalting

Recombinant N-terminally His$_6$-tagged OYEs were subjected to a second purification step by means of size-exclusion chromatography. To ensure complete saturation of the apoprotein with flavin cofactor, pooled samples from the Ni-affinity chromatography step were incubated with excess FMN (FMN and FAD for *Ss*OYE) overnight at 4°C before size-exclusion chromatography or desalting was performed. Samples were subsequently concentrated to approximately 3 mL by ultrafiltration (30 kDa MWCO, Millipore) and loaded onto a Sephacryl S-200HR column (65 x 2.5 cm, Sigma) equilibrated with 20 mM MOPS-NaOH (pH 7). Proteins were eluted with the same buffer at a flow speed of 1 mL.min^{-1}. Alternatively, samples from the Ni-affinity chromatography step were desalted using PD-10 desalting columns (GE Healthcare) into 20 mM MOPS-NaOH (pH 7).

Purification by size-exclusion chromatography also allowed the native oligomeric state of the OYEs to be determined. Biorad Gel Filtration Standard and bovine serum albumin (66 kDa)) were used as molecular weight standards, eluted with the same buffer and flow-rate (1 mL.min^{-1}) as the OYEs (figures 4.1 and 4.2).
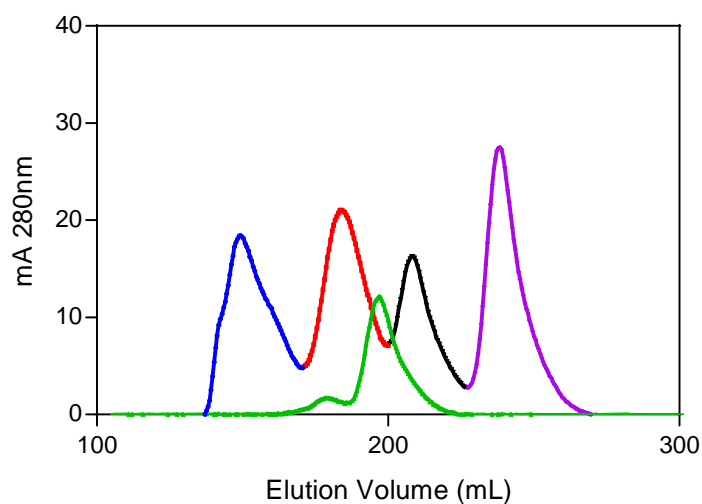
Figure 4.1: Elution profile of Sephacryl S-200HR column calibration using bovine serum albumin [66 kDa (▲)]and the Gel Filtration Standard (Biorad) consisting of bovine thyroglobulin [670 000 kDa (●)], bovine ɣ-globulin [158 kDa (■)], chicken ovalbumin [44 kDa (▼)] and horse myoglobin [17 kDa (◆)].
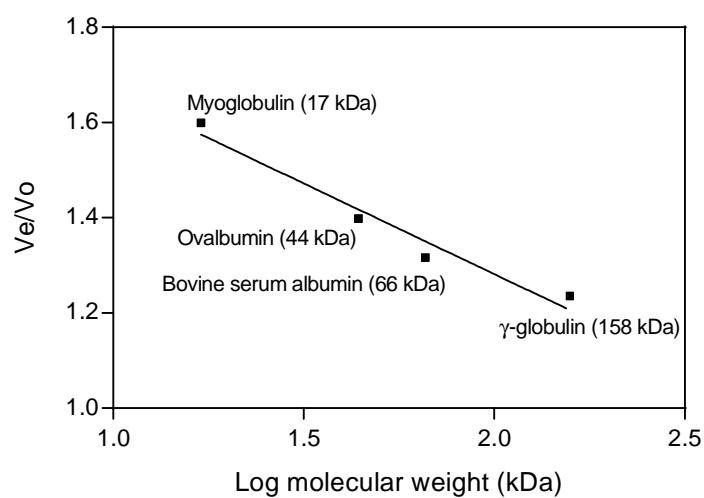
Figure 4.2: Calibration curve of Sephacryl S-200HR relating molecular weight to elution volume. The void volume (V_0) was calculated using the elution volume of bovine thyroglobulin (670 kDa).

4.2.2.3.   Analysis of protein purity through gel electrophoresis

Enzyme purity was evaluated using SDS-PAGE as described in 3.2.5.2. Precision Plus Protein All Blue Standards (Bio-Rad) and PageRuler™ Unstained Broad Range Protein Ladder (Fermentas) were used as molecular weight markers.

4.2.2.4.   Protein concentrations

Protein concentrations were determined using the bicinchoninic acid (BCA) method (Smith *et al.*, 1985). The BCA Protein Assay Kit (Pierce) was used according to the manufacturer's instructions with bovine serum albumin (BSA) (supplied with kit) as standard.
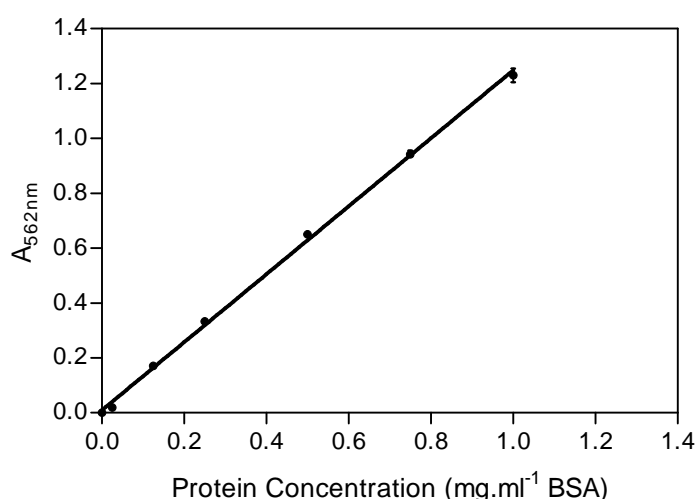
Figure 4.3:   Standard curve for the BCA protein assay kit (Pierce) at 37°C using BSA as protein standard. Error bars indicate standard deviation.

4.2.3. Characterization of purified OYEs

4.2.3.1.   Effect of pH on enzyme activity

The effect of pH on the activity of purified OYEs towards 2-cyclohexenone was measured by incubation of the reaction mixtures (at 22°C for *Cm*OYE and at 60°C for *Ss*OYE) at various pHs ranging from pH 5 to pH 9.   The pH range was constructed by adjusting the pH of an equimolar 20

mM MES-MOPS-Bicine buffer with HCl or NaOH. Reactions were performed aerobically in 1 mL reaction mixtures containing 20 mM MES-MOPS-Bicine buffer, 0.3 mM NADPH, 10 mM 2-cyclohexenone and 6 μg of the purified enzyme. Oxidation of NADPH was monitored as described in 3.2.5.3. Assays were done in triplicate and blanks (representing the oxidation of NADPH by molecular oxygen in the absence of 2-cyclohexenone) taken at each assayed pH value.

4.2.3.2. Effect of temperature on enzyme activity

The effect of temperature on the activity of purified OYEs towards 2-cyclohexenone was measured by incubation of the reaction mixtures at various temperatures at pH 7. Reactions were performed aerobically in 1 mL reaction mixtures containing 20 mM MOPS buffer (pH adjusted to be pH 7 for *Cm*OYE and pH 5.5 for *Ss*OYE at each temperature step), 0.3 mM NADPH, 10 mM 2-cyclohexenone and 6 μg purified enzyme. Oxidation of NADPH was monitored as described in 3.2.5.3. Assays were done in triplicate and blank rates (representing oxidation of NADPH in the absence of 2-cyclohexenone) taken at each assayed temperature value.

4.2.3.3. Steady-state kinetics

Kinetic constants were determined using purified enzyme by measurement of initial velocities at various concentrations of 2-cyclohexenone. Reactions were performed at pH 6.5 (MOPS-NaOH buffer) and 25°C for OYE from *C. metallidurans* CH34 and at pH 5.5 (MES-HCl buffer) and 55°C for OYE from *S. solfataricus* P2. Reactions were monitored by measuring NADPH oxidation at 340 nm as described in 3.2.5.3. Rates of NADPH oxidation were calculated with use of an extinction coefficient of 6.22 mM.cm^{-1}. Assays were performed in 1 ml reaction volumes containing NADPH (0.3 mM), 2-cyclohexenone and the purified protein (6 μg). Assays were performed under aerobic conditions and NADPH oxidation by the enzyme due to molecular oxygen was measured independently and subtracted from the total oxidation rates with substrates. Apparent maximum initial velocities (V_{max}) and apparent Michaelis constants (K_m) were determined through non-linear regressions of Michaelis-Menten plots using GraphPad Prism 5.04 (GraphPad Software, Inc).

4.2.3.4. Substrate scope

In order to investigate the substrate scope of the OYEs, biotransformations were performed using a variety of potential oxidants. Biotransformations were performed in 1 ml reaction volumes, consisting of equimolar NADH and substrate, reconstituted freeze-dried purified enzyme (5 μg) in 20 mM MES-NaOH buffer (pH 6.5 for *Cm*OYE and pH 5.5 for *Ss*OYE) at their respective optimal temperatures. Freeze-dried protein was reconstituted through addition of 10 mM MES pH 6.5 and

pH 5.5 for *Cm*OYE and *Ss*OYE, respectively. Substrate stocks were prepared in 10 mM MES buffer at pH 5.5 (for *Ss*OYE) and pH 6.5 (for *Cm*OYE). For dimethyl fumarate, the substrate and NADH concentrations in the reaction were 1.5 mM and for 2-methylmaleimide and N-phenylmaleimide the concentration was 1 mM. All other biotransformations were performed at 3 mM substrate and NADH concentrations. Samples for analysis were taken at 0 and 3 hours incubation and yields were determined after 3 h of incubation. For GC analysis, reaction mixtures were extracted using an equal volume of ethyl acetate whereas for HPLC analysis, reactions were stopped by the addition of 10 µL of concentrated HCl. Construction of calibration curves of each substrate and respective product allowed quantification of chromatography data. For calibration curves, dilution ranges consisted of 1 to 10 mM for each substrate and the corresponding product, with stocks prepared in 20 mM MES buffer pH 6.

4.2.3.5. Light-driven cofactor regeneration

The efficacy of the light-driven cofactor regeneration system was determined in conjunction with reduction of 2-cyclohexenone, (R)-(-)-carvone and (S)-(+)-carvone by OYE from *C. metallidurans* and dimethylmaleate by OYE from *S. solfataricus* P2. Reactions were performed in a 10 mL scale, anaerobically under N_2 flux at the optimum temperature and pH conditions for each enzyme, and consisted of 0.3 mM FMN, 5 mM EDTA, substrate (5 mM for 2-cyclohexenone, (R)-(-)-carvone and (S)-(+)-carvone; 10 mM for dimethylmaleate) and 0.2 mg.mL^{-1} enzyme. Reactions were performed under illumination (250 W bulb, 5 cm distance) and samples were taken for analysis at 0, 2, 6 and 8 hour intervals.

4.2.3.6. Reverse reactions

In addition to substrate scope, the enzymes' ability to catalyse the desaturation of dimethyl succinate, (+)-dihydrocarvone, cyclohexanone and (4R,6R)-4-hydroxy-2,6,6–trimethylcyclohexanone were also investigated. Reactions were performed in 1 mL reaction volumes, consisting of 0.3 mM FMN, 10 mM substrate, reconstituted freeze-dried purified enzyme (5 µg) and 20 mM MES-NaOH buffer (pH 6.5 for *Cm*OYE and pH 5.5 for *Ss*OYE) and incubated at their respective optimal temperatures for 18 hours. In some cases, reaction mixtures also contained 1.4 mg catalase. Control reactions were performed in the absence of enzyme. Reaction mixtures were extracted using an equal volume of ethyl acetate. Samples for analysis were taken at 0 and 18 hours.

For (R)-(-)-carvone, (S)-(+)-carvone, 3-methylcyclopentenone and citral, reactions were analysed using a Shimadzu 2010plus with an AOC-20i+s autosampler running Shimadzu GC solutions and an Astec Chiraldex GTA 50.0 m x 0.25 mm x 12 μm column.   For dimethylfumarate, dimethylmaleate, ketoisophorone, 2-methylcyclopentenone and 3-methylcyclohexenone, a Shimadzu GC-17A V3 with an AOC-20i+s autosampler running Shimadzu GC solutions and a Varian Cyclodex-B 62.5 m x 0.25 mm x 0.25 μm column.   HPLC analysis was done using a Shimadzu LC-20 HPLC-PDA system running Shimadzu LC solutions and a Waters Xterra RP18 5 μm x 4.6 mm x 1500 mm column.   For 2-cyclohexenone and dimethylmaleate, reactions were analysed using a Shimadzu GC-17A V3 with an AOC-20i+s autosampler running Shimadzu GC solutions and a Varian Cyclodex-B 62.5 m x 0.25 mm x 0.25 μm column.   Reactions for dimethyl succinate, ketoisophorone and cyclohexanone were analysed using a Shimadzu GC-17A V3 with an AOC-20i+s autosampler running Shimadzu GC solutions and a Varian Cyclodex-B 62.5 m x 0.25 mm x 0.25 μm column.   For (+)-dihydrocarvone, reactions were analysed using a Shimadzu 2010plus with an AOC-20i+s autosampler running Shimadzu GC solutions and an Astec Chiraldex GTA 50.0 m x 0.25 mm x 12 μm column.

4.2.4. Crystallization of OYEs from *C. metallidurans* CH34 and *S. solfataricus* P2 and X-ray diffraction

4.2.4.1. Crystallization of OYEs from C. metallidurans *CH34 and* S. solfataricus *P2*

For both proteins, recombinant N-terminally His$_6$-tagged OYE was prepared from BL21(DE3) (Lucigen) cells and subjected to IMAC and size-exclusion chromatography as previously described. Purified protein was concentrated to 8 mg.mL^{-1} using Amicon Ultra-15 centrifugal filter units (30 000 MWCO, Merck Millipore), with protein concentration determined using the BCA Protein Assay Kit (Pierce) as in 4.2.2.4.   Initial crystallization trials were performed by the sitting-drop vapour diffusion method in 96-well plates at 16°C, equilibrated against 150 μL reservoir solution using the PEGRx 1, PEGRx 2, PEG/Ion Screen, PEG/Ion 2 Screen, Crystal Screen, Crystal Screen 2, SaltRx 1, SaltRx 2, GridScreen Ammonium Sulphate and Grid Screen MPD kits (Hampton Research).   Each drop was prepared by mixing 1 μL protein solution prepared above (8 mg.mL^{-1}) and 1 μL crystallization (reservoir) solution.

Optimisation of the crystallization conditions was carried out using the hanging-drop vapour diffusion method at 16°C.   Drops were prepared by mixing 1 μL protein solution and 1 μL

crystallization (reservoir) solution and equilibrating against 1 mL reservoir solution. Conditions were optimised by varying pH, precipitant concentration and protein concentration (4 mg.mL^{-1}, 6 mg.mL^{-1} and 8 mg.mL^{-1}) until single crystals were obtained. Crystals were flash-frozen directly in liquid nitrogen using 30% (v/v) glycerol [30% (v/v) glycerol, 70% (v/v) reservoir solution] as cryoprotectant, enabling collection of diffraction data.

4.2.4.2. Data collection

X-ray diffraction data were collected at the Diamond, iO4-1 beamline (Oxfordshire, UK) at 0.9173 Å wavelength using single crystals and a Pilatus 2M hybrid pixel-array detector. Data reduction was performed using Xia2 (Winter, 2010).

4.3. RESULTS AND DISCUSSION

4.3.1. Protein expression and purification

Recombinant OYEs were purified from the soluble fraction of *E. coli* to near homogeneity through the use of immobilized metal affinity chromatography (IMAC) and size-exclusion chromatography. IMAC employs the interaction between a transition metal ion [in this case Ni^{2+}, with the column consisting of a nickel-nitrilotriacetic acid (Ni^{2+}-NTA) matrix] immobilized on a matrix and the short affinity tag consisting of polyhistidine residues (Bornhorst & Falke, 2000). Of all the amino acids, histidine exhibits the strongest interaction with immobilized metal ion matrices, due to the tendency of electron groups on the histidine imidazole ring to readily form coordination bonds with the immobilized transition metal. As a result, peptides containing sequences of consecutive histidine residues are effectively retained on IMAC column matrices and polyhistidine affinity-tagged proteins can be rapidly purified, resulting in up to 100-fold enrichments and up to 95% purity in a single purification step. In some cases, use of longer polyhistidine tags has resulted in increased purity due to the better binding and the ability to use more stringent washing steps (Grisshammer & Tucker, 1997). Yet it is advisable to use the smallest number of histidine residues required for efficient purification to minimize possible influences on protein function. A six histidine tag is generally accepted as being sufficient for protein purification (Bornhorst & Falke, 2000).

Polyhistidine tags can be placed on either the N or the C terminus of recombinant proteins, with optimal placement of the tag being protein-specific. Inability of the protein to interact with the column may be due to occlusion of the tag in the folded protein, resulting in inaccessibility of the tag to the immobilized metal (Bornhorst & Falke, 2000). The functionality of recombinant OYE has been

shown to be affected by the placement of the tag on the C-terminal of the protein (Fitzpatrick *et al.*, 2004). Comparison of the biochemical properties of untagged YqjM with that of a C-terminal histidine-tagged form and an N-terminal glutathione *S*-transferase (GST)-tagged form revealed critical differences in the catalytic activities and quaternary structure of the protein forms. Firstly, UV-visible absorption spectra revealed that the native (untagged) protein is isolated with FMN as a tightly bound cofactor. While the GST-form is almost indistinguishable from the native form, C-terminal His$_6$-tagging was found to significantly affect binding of the flavin cofactor to the apoprotein. These observations are perhaps not surprising, as the FMN cofactor is known to bind at the C-terminal end of the apoprotein (Fox & Karplus, 1994).

The catalytic activity of YqjM is also affected by the position of the chosen affinity tag (Fitzpatrick *et al.*, 2004). N-terminally GST-tagged YqjM catalyses the oxidation of NADPH in the presence of molecular oxygen, albeit at a lower rate than that observed for the native protein. GST-YqjM also exhibited the ability to utilize the same α/β-unsaturated carbonyl compounds as the native protein, but again at lower catalytic efficiency. In the case of C-terminally His-tagged YqjM, however, oxidation of NADPH was observed in the presence of oxygen, but at a 7-fold lower rate than that observed for the native protein. In addition, YqjM-His$_6$ exhibited an inability to utilise as substrates compounds previously shown to act as oxidants for both the native and the N-terminally GST-tagged YqjM (Fitzpatrick *et al.*, 2004). These observations were concurrent with the known important structural role played by C-terminal residues in the architecture of the catalytic site of the OYE (Fox & Karplus,1994). As the presence of an affinity tag at the C-terminal end of the protein appears to have a more adverse effect on the active site architecture and cofactor binding of the OYE than N-terminal tagging, OYEs from both *C. metallidurans* CH34 (*Cm*OYE) and *S. solfataricus* P2 (*Ss*OYE) were expressed as recombinant N-terminally His$_6$-tagged proteins.

Both recombinant OYEs interacted with the Ni-affinity column effectively as was evident by the yellow colour present on the columns after binding of the crude soluble fractions. While the characteristic yellow colour of OYEs was present for *Cm*OYE, columns containing bound *Ss*OYE exhibited a darker, yellow-brown colour, possibly due to the presence of the 4Fe-4S cluster in the protein. Following binding of the tagged OYE, the column was washed with buffer containing 20 mM imidazole to elute non-specifically bound proteins that bind weakly to the column, more effectively. These non-specifically bound proteins may include cellular proteins containing two or more adjacent histidine residues (Schmitt *et al.*, 1993). After the washing step, recombinant OYE bound strongly to the column was eluted by employing an increasing imidazole concentration gradient, up to 0.5 M imidazole. Imidazole, a histidine analogue, acts by competitively eluting the bound polyhistidine residues (Bornhorst & Falke, 2000). For both OYEs, elution profiles obtained through Ni-affinity chromatography (figure 4.4 a-b) exhibited a single absorbance peak appearing as the imidazole

concentration increases. In each case, the peak corresponds to the N-terminally histidine-tagged protein eluting from the column, the presence of which was confirmed by the characteristic yellow colour (as a result of the bound cofactors) observed in the fractions. Fractions containing the eluted recombinant OYE were selected using both the elution volumes of the emerging absorbance peak and the yellow colour of the eluted protein.

Pooled fractions from the above IMAC purification step were subjected to a second purification step by means of size-exclusion chromatography. Size-exclusion chromatography can be used to effectively separate small molecules from larger, complex polymers as an aid in sample clean-up (Barth *et al.*, 1994). In addition, the technique allows the determination of statistical average molecular weights of polymers. However, size-exclusion chromatography is a relative technique that depends on the relative size or hydrodynamic volume of a macromolecule with respect to the average pore size of the column matrix. Column calibration is required in order to determine the molecular weights of macromolecules (Barth *et al.*, 1994). As a result, the column was calibrated using cytochrome c (12.4 kDa), chymotrypsin (25 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (148 kDa) and Blue Dextran (2 000 kDa) as molecular weight standards. These standards were eluted with the same buffer and flow-rate as the OYEs and a standard curve correlating molecular weight and elution volume was constructed. Elution volume for Blue Dextran was regarded as the void volume of the column.

UV-visible spectral analyses with YqjM have shown that both the native and affinity-tagged proteins isolated are not fully saturated with the flavin cofactor (Fitzgerald *et al.*, 2004). As a result, pooled fractions were incubated overnight in the presence of excess FMN (FMN and FAD in the case of *Ss*OYE) to ensure the saturation of the recombinant OYE with the flavin cofactor, prior to size exclusion chromatography. For each OYE, the elution profiles (figures 4.4 c-d) obtained through the size-exclusion chromatography on a Sephacryl S-200HR column revealed a single absorbance peak corresponding to the recombinant OYE. *Cm*OYE eluted after approximately 202 mL, while *Ss*OYE eluted at approximately 190 mL.

SDS-PAGE analysis (figure 4.5) confirmed expression of *Cm*OYE and *Ss*OYE, with apparent monomer molecular weights of approximately 40 kDa and 60 kDa respectively. Using the elution profiles and the calibration curve of Sephacryl S-200HR (figure 4.2) it was established that the OYEs from both *C. metallidurans* CH34 and *S. solfataricus* P2 exhibit monomeric quaternary native structures in solution. Along with *Cm*OYE and *Ss*OYE, other monomeric OYEs include the "classical" PETN reductase (Barna *et al.,* 2001) and OPR (Breithaupt *et al.*, 2001), as well as the "thermophilic-like" XenA (Griese *et al.*, 2006). "Thermophilic-like" YqjM has been shown to exist in solution as a tetramer (dimer of homodimers) (Fitzpatrick *et al.*, 2003).

As mentioned in chapter 3, *Ss*OYE is thought to exhibit a non-classical OYE structure consisting of two domains – the N-terminal TIM-barrel of the OYEs, combined with the second, C-terminal domain highly homologous to that of 2,4-dienoyl-CoA reductase of *E. coli*. The presence of the second domain and the subsequent occupation of residues normally involved in the dimer interface of OYEs (Fox & Karplus, 1994; Kitzing *et al.*, 2005) may explain the monomeric state of *Ss*OYE in solution.

SDS-PAGE analysis (figure 4.5) confirmed the near-homogeneity of the protein after purification with IMAC and size-exclusion chromatography (or ultrafiltration). Analysis of soluble extract of *E. coli* expressing OYEs revealed numerous proteins present in addition to the approximately 40 kDa OYE (*Cm*OYE) and approximately 60 kDa OYE (*Ss*OYE). After purification steps, bands representing these native *E. coli* and other contaminating proteins were greatly reduced and clear bands indicated the presence of homogeneous OYE from *C. metallidurans* CH34 and *S. solfataricus* P2.
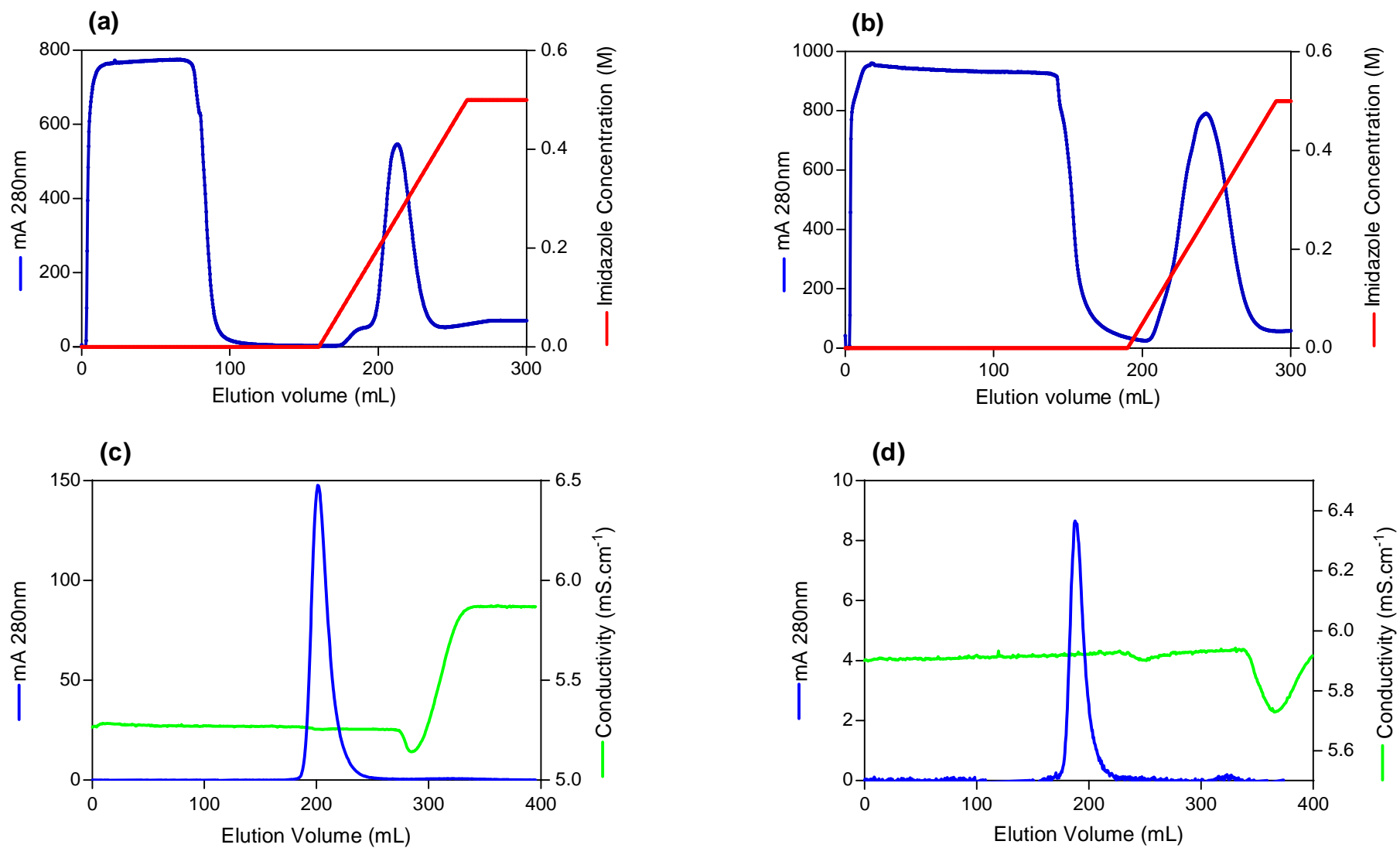
Figure 4.4: Purification of the recombinant OYEs from *C. metallidurans* CH34 **(a&c)** and *S. solfataricus* P2 **(b&d)** overproduced in *E. coli* through Ni-affinity **(a-b)** and size-exclusion **(c-d)** chromatography.
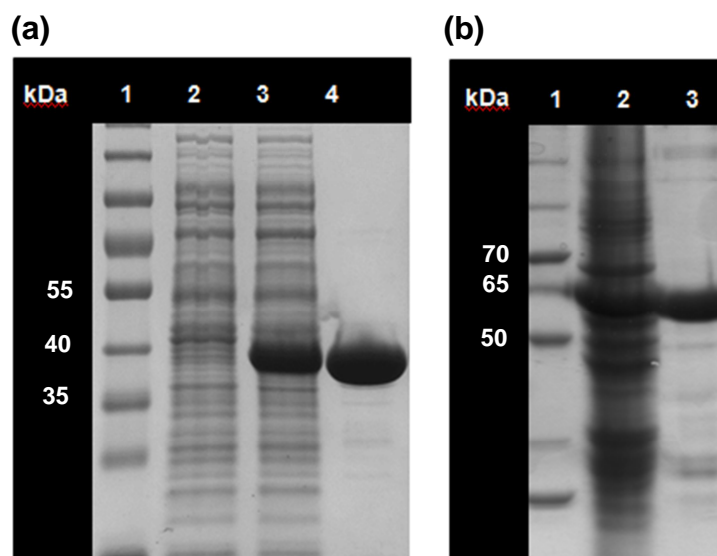
Figure 4.5: SDS-PAGE analysis of the expression and purification of OYE from **a)** *C. metallidurans* CH34 and **b)** *S. solfataricus* P2. **a)** Lane 2 represents crude extracts of *E. coli* transformed with pET28b(+) plasmid without insert. Lane 3 represents soluble extract of *E. coli* expressing *Cm*OYE. Lane 4 represents purified *Cm*OYE. **b)** Lane 2 represents crude extract of *Ss*OYE. Lane 3 represents purified *Ss*OYE. Lanes 1a and 1b, Precision Plus molecular weight marker.

4.3.2. Functional characterisation of purified recombinant OYEs

4.3.2.1. Effect of pH on enzyme activity

The effect of pH conditions on the activity of the two OYEs towards the ubiquitous OYE substrate 2-cyclohexenone was investigated by monitoring the oxidation of NADPH spectrophotometrically at different pH conditions. Figure 4.6 depicts the effect of pH on enzyme activity towards 2-cyclohexenone in a MES-MOPS-Bicine buffer. Relative activities are shown, with maximum activity taken as 100%. Maximal reductase activity was achieved at pH 6.5-7 for the OYE from *C. metallidurans* CH34, correlating with the optimum growth pH of the source organism of pH 7 (Goris *et al.*, 2001). At pH 5.5, the optimum pH of the OYE from *S. solfataricus* P2 is higher than the source organism's optimum growth pH of 2-4 (She *et al.,* 2001), but corresponds well with the reported intracellular pH (~6.5) of *S. solfataricus* (Baker-Austin & Dopson, 2007). OYE from *C. metallidurans* CH34 was more than 60% active between pH 5.5 and 8, while OYE from *S. solfataricus* P2 was more than 50% active between pH 4.5 and 7.
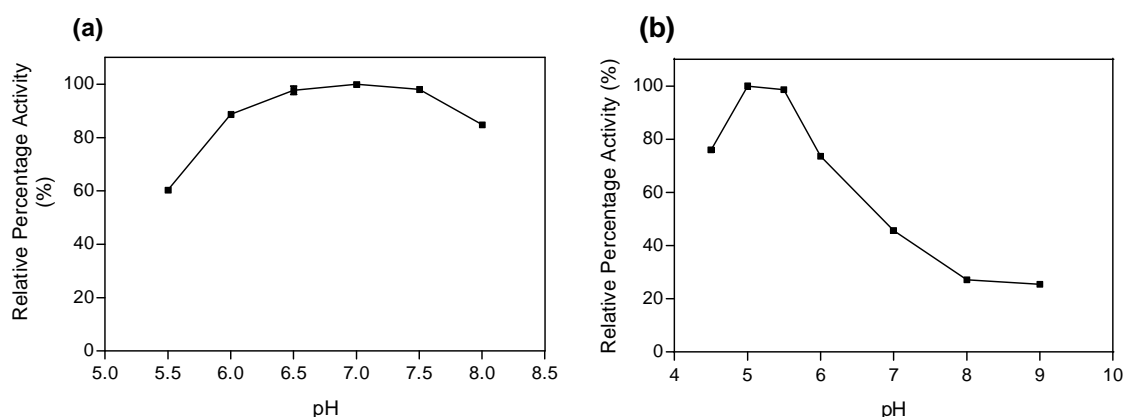
Figure 4.6:   Effect of pH on the activity of purified OYE from **a)** *C. metallidurans* CH34 and **b)** *S. solfataricus* P2.   Activity at pH 7.0 and pH 5.5 (optima, respectively) were taken as 100%. Error bars indicate standard deviation.

4.3.2.2.   Effect of temperature on enzyme activity

The effect of temperature on the activity of the two OYEs towards the ubiquitous OYE substrate 2-cyclohexenone was investigated by monitoring the oxidation of NADPH. Figure 4.7 depicts the effect of temperature on enzyme activity towards 2-cyclohexenone in MOPS buffer, with pH adjusted to be 7 at each temperature step. Relative activities are shown, with maximum activity taken as 100%. *Cm*OYE was more than 50% active between 20°C and 35°C. Optimum activity for the enzyme was determin ed at 25°C, slightly lower than the optimum growth temperature of *C. metallidurans* CH34 of 30°C. OYE from *S. solfataricus* P2 was more than 50% active between 50°C and 85°C.   Optimum temperature of the enzyme was determined at 80°C, slightly higher than the optimum growth temperature of the organism at 70°C (She *et al.*, 2001). These results were similar to that reported for OYEs from other thermophilic organisms. The thermophilic CrS was reported to exhibit optimum activity at 65°C and above 50% activity at 80°C, wh ich correlates well with the optimum growth temperature of the organism of origin, *Thermus scotoductus* SA-01 (Opperman *et al.*, 2008). While the complete temperature profile for the activity of the OYE from *Geobacillus kaustophilus* has not been reported, the enzyme has been shown to catalyse the conversion of ketones to the corresponding α/β-unsaturated compounds (see 4.3.2.6.) at temperatures above the 65°C growth optimum of the source organis m, although these reactions did not require nicotinamide cofactor (Schittmayer *et al.*, 2011).
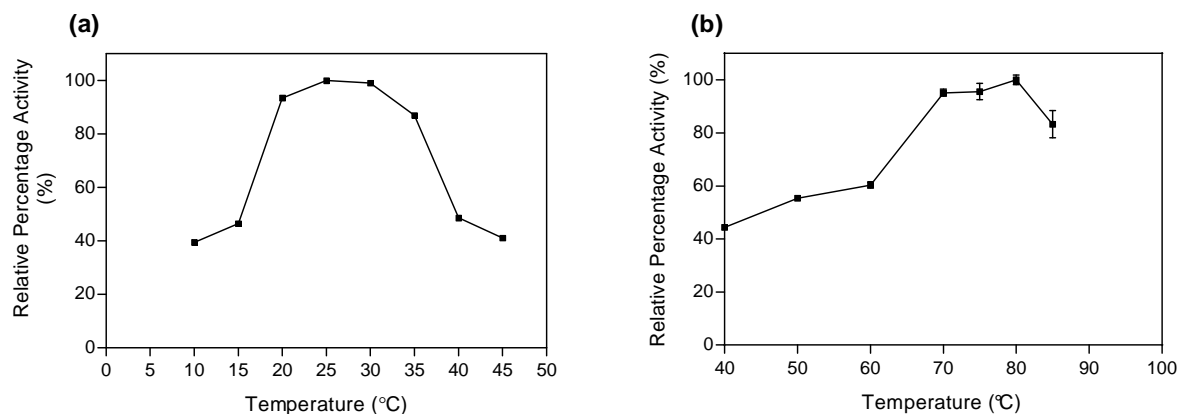
Figure 4.7: Effect of temperature on the activity of purified OYE from **a)** *C. metallidurans* CH34 and **b)** *S. solfataricus* P2. Activity at 25°C and 55°C (optima, respectively) were taken as 100%. Error bars indicate standard deviation.

4.3.2.3. *Steady state kinetics*

Kinetic constants for the reduction of 2-cyclohexenone by the two purified OYEs were determined by measurement of initial velocities at various concentrations of 2-cyclohexenone. Reactions were performed at the optimum pH and temperature conditions as determined for each enzyme and were monitored by measuring NADPH oxidation. Reactions were performed aerobically and blank rates were measured, representing the oxidation of NADPH by molecular oxygen in the absence of 2-cyclohexenone. Figure 4.8 shows the hyperbolic dependence of 2-cyclohexenone concentration to specific activity determined for the two purified OYEs. Both purified OYEs exhibited Michaelis-Menten kinetics during the reduction of 2-cyclohexenone, with reaction rates reaching maximum at high substrate concentrations. Table 4.1 shows the catalytic parameters determined for the two purified OYEs, along with kinetic parameters for the "classical" OYEs OYE1 (*Saccharomyces pastorianus)* and morphinone reductase (*Pseudomonas putida)* and for the "thermophilic-like" OYEs, CrS (*Thermus scotoductus* SA-01), xenobiotic reductase A (*Pseudomonas putida*) and YqjM (*Bacilis subtilis*).

OYE from *C. metallidurans* CH34 was determined to have a V_{max} of 5.71 µmol.min^{-1}.mg^{-1}, approximately double the V_{max} determined for OYE from *S. solfataricus* P2. For both OYEs, V_{max} values are significantly lower than that determined for the thermophilic homologue CrS from *T. scotoductus* SA-01 (Opperman *et al.*, 2010). K_m for reduction of 2-cyclohexenone by *Cm*OYE was determined at 5.05 mM, comparable to K_m determined for CrS (Opperman *et*

al., 2010) and morphinone reductase (Barna *et al.*, 2002). For both *Cm*OYE and *Ss*OYE, significant differences in the affinities for 2-cyclohexenone were apparent when compared to the homologues in the "thermophilic-like" group of OYEs, XenA and YqjM. Closer inspection revealed that, for both *Cm*OYE and *Ss*OYE, affinity towards 2-cyclohexenone is more comparable with that of morphinone reductase than to XenA and YqjM. The greatest difference in affinities was observed for both *Cm*OYE and *Ss*OYE when compared to the "classical" OYE, OYE1 from *S. pastorianus* (Brown *et al.*, 2002).

Substrate turnover (signified by K_{cat}), is highly comparable among OYE from *C. metallidurans* CH34, "classical" OYE1 (Brown *et al.*, 1998) and thermophilic-like YqjM (Fitzpatrick *et al.*, 2003), while turnover determined for OYE from *S. solfataricus* P2 was slightly lower. These results were clear despite *Cm*OYE and *Ss*OYE bearing much more sequence similarity to YqjM than to OYE1. Both purified OYEs exhibited superior turnover numbers compared to morphinone reductase (Barna *et al.*, 2002), but significantly lower compared to the closer related (in terms of sequence identity) CrS (Opperman *et al.*, 2010).

Catalytic efficiency (signified by K_{cat}/K_m) of OYE from *S. solfataricus* is lower than that observed for *Cm*OYE, yet higher than (but still comparable to) the catalytic efficiency of morphinone reductase (Barna *et al.*, 2002). However, catalytic efficiencies of both thermophilic-like CrS (Opperman *et al.*, 2010) and YqjM (Fitzpatrick *et al.*, 2003) are approximately 20-fold higher than that determined for *Cm*OYE and *Ss*OYE. Catalytic efficiencies of "classical" OYE1 (Brown *et al.*, 1998) and the two purified OYEs' closer relative XenA, are significantly higher than that determined for *Cm*OYE and *Ss*OYE. Low catalytic efficiency of the *Cm*OYE and *Ss*OYE appears to be mostly due to high K_m values.
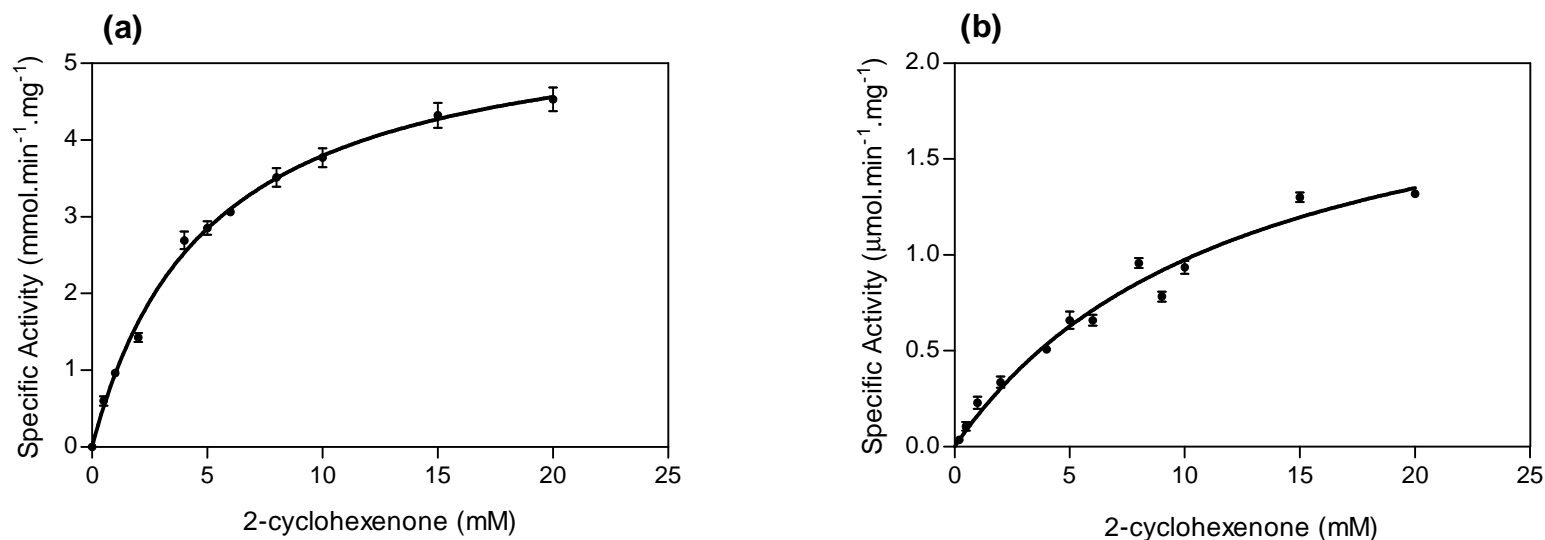
Figure 4.8: Steady-state kinetics of the purified OYE from **a)** *C. metallidurans* CH34 and **b)** *S. solfataricus* P2 illustrating the dependence of initial velocities against substrate concentrations. Error bars indicate standard deviation.

Table 4.1: Catalytic parameters of OYE homologues towards 2-cyclohexenone with NADPH as electron donor

Enzyme	V_{max} [µmol.min^{-1}.mg^{-1}]	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [mM^{-1}.s^{-1}]	Reference
OYE1 (*S. pastorianus*)	-	<0.01	4.2	>420	Brown *et al,.* 1998
MR (*P. putida*)	-	4.2	0.8	0.19	Barna *et al.*, 2002
CrS *(T. scotoductus* SA-01)	170.7 ± 3.9	3.68 ± 0.24	110.1	29	Opperman *et al.*, 2010
XenA (*P. putida* 86)		0.033	8	242	Griese, *et al.*, 2006
YqjM (*B. subtilis)*	-	0.293	4.38	15	Fitzpatrick *et al.*, 2003
OYE (*C. metalllidurans* CH34)	5.71 ± 0.11	5.05 ± 0.26	4.03	0.80	This study
OYE (*S. solfataricus* P2)	2.18 ± 0.16	12.37 ± 0.69	2.29	0.19	This study

To explore the substrate scope and selectivity of the two OYE homologues, a range of potential substrates were investigated (table 4.2). Biotransformation reactions were performed using purified enzymes and NADH as cofactor. Percentage yields were calculated using the starting substrate concentration and the product concentration determined from chromatographic analysis.

Cyclic enones with methyl substitutions on the Cβ position were not reduced by either *Ss*OYE or *Cm*OYE, similar to YqjM (Hall *et al.*, 2008b) and CrS (Opperman *et al.*, 2010). These compounds have been shown to act as substrates for OYE1 (Hall *et al.*, 2008a). Similarly, the only enal tested, citral, has a methyl group on the Cβ and no reactivity was detected for either *Cm*OYE or *Ss*OYE after a 3 h reaction. While no reactivity towards citral has been observed for CrS (Opperman *et al.*, 2010), citral has been shown to act as a substrate for YqjM (Hall *et al.*, 2008b) and XenA (Chaparro-Riggers *et al.*, 2007). 2-Methylcyclopentanone was only reduced by *Cm*OYE, albeit marginally, with 2.2% yield, while 2-methylcyclohexenone was not reduced by either *Cm*OYE or *Ss*OYE. OYE1 (Hall *et al.*, 2008a) and YqjM (Hall *et al.*, 2008b) have been shown to act on the cyclic substrates with methyl substitutions on the Cα position, 2-methylcyclopentenone and 2-methylcyclohexenone, with a preference for the cyclohexenone substrate observed for OYE1 (Hall *et al.*, 2008a).
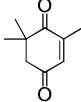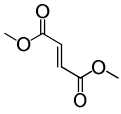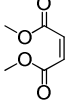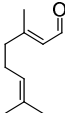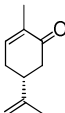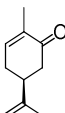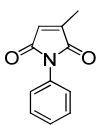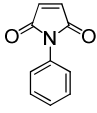
The dimethyl α,β-unsaturated ester, dimethylfumarate, also served only as substrate for *Cm*OYE. The *cis* configuration dimethyl α,β-unsaturated ester (dimethylmaleate) appeared to be the preferred substrate, serving as substrate for both *Ss*OYE and *Cm*OYE (with higher yields observed for *Cm*OYE).

The structurally more demanding substrates, the isomers of carvone, proved to be converted by both *Ss*OYE and *Cm*OYE, although extremely low activities were found for *Ss*OYE. The (R)-(-)-carvone was slightly preferred by the *Cm*OYE, with more than 1.5 mM converted after 3 h incubation. This contrasts with results reported for XenA, which exhibited no activity towards (R)-(-)-carvone (Chaparro-Riggers *et al.*, 2007). Carvone has been reported to act as a substrate for CrS (Opperman *et al.*, 2010). For both *Ss*OYE and *Cm*OYE, conversion of (S)-(+)-carvone resulted in two clearly observable product peaks, corresponding to the two isomers of the product: (2R,5S)-5-isopropenyl-2-methylcyclohexanone and (2S,5S)- 5-isopropenyl-2-methylcyclohexanone. For *Cm*OYE, it was determined that the (2R,5S)-enantiomer of the product was produced in 96% enantiomeric excess. Conversions for

*Ss*OYE were too small to accurately determine *ee* values. For both *Cm*OYE and *Ss*OYE, however, conversion of (R)-(-)-carvone resulted in only one product peak.

The only substrates where higher conversion was observed for *Ss*OYE compared to *Cm*OYE were ketoisophorone and the N-phenyl-substituted maleimides. Although marginal, *Ss*OYE exhibited almost twice the conversion of ketoisophorone than *Cm*OYE. 2-methyl-N-phenylmaleimide was more readily reduced than its un-substituted counterpart N-phenylmaleimide, with both *Ss*OYE and *Cm*OYE reaching more than 50% conversion of the 1 mM substrate in 3 h. This preference for 2-methyl-N-phenylmaleimide was much more pronounced in *Cm*OYE (approximately twice the conversion) as compared to the minimal preference observed for *Ss*OYE. Ketoisophorone has been shown to act well as a substrate for OYE1 (Hall *et al.*, 2008a), YqjM (Hall *et al.*, 2008b), CrS (Opperman *et al.*, 2010) and XenA (Chaparro-Riggers *et al.*, 2007). Similarly, OYE1, XenA and YqjM have been shown to utilise both these maleimide substrates well, with no particular preference for either substrate (Hall *et al.*, 2008a; Hall *et al.*, 2008b).

Table 4.2. Substrate scope *Ss*OYE and *Cm*OYE

Substrate		Yield [%][a]	
		*Ss*OYE	*Cm*OYE
2-methylcyclopentenone		n.d.	2.2
3-methylcyclopentenone		n.d.	n.d.
3-methylcyclohexenone		n.d.	n.d.
Ketoisophorone		3.2	1.5
Dimethylfumarate		n.d.	7.6
Dimethylmaleate		7.2.	21.2
Citral		n.d.	n.d.
(R)-(-)-Carvone		1.3	53.4
(S)-(+)-Carvone		2.5	41.9
2-methyl-N-phenyl maleimide		68.9	54.4
N-phenylmaleimide		54.6	26.9

n.d. not detected after 3 h conversion

[a] All substrate concentrations were 3 mM, except for the maleimides (1 mM) and dimethylfumarate (1.5 mM)

4.3.2.5.   *Cofactor regeneration*

As discussed in 1.4.1., one of the major drawbacks in the application of OYEs are the requirement of expensive co-factors. Various co-factor regeneration systems are therefore being employed to overcome this limitation not only for OYEs, but redox enzymes in general (Fryszkowska *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008b; Hollmann *et al.*, 2010). The use of cofactor-regeneration systems such as glucose dehydrogenase, glucose-6-phosphate dehydrogenase and formate dehydrogenase systems have been shown to be effective when used in conjunction with OYE reactions (Hall *et al.*, 2007; Hall *et al.*, 2008a; Hall *et al.* 2008b). An alternative to these conventional enzymatic cofactor systems has been the use of a light-driven direct regeneration system that employs free FMN, EDTA and white light as a source of reducing equivalents (Grau *et al.*, 2009; Taglieber *et al.*, 2008). By directly regenerating the FMN of the OYE through exploitation of the photoexcitability of the flavin, the use of costly and unstable nicotinamide cofactors is not necessary. We therefore tested *Ss*OYE and *Cm*OYE for their applicability with the light-driven co-factor regeneration system (Taglieber *et al.*, 2008) and 2-cyclohexenone as model substrate (figure 4.9).
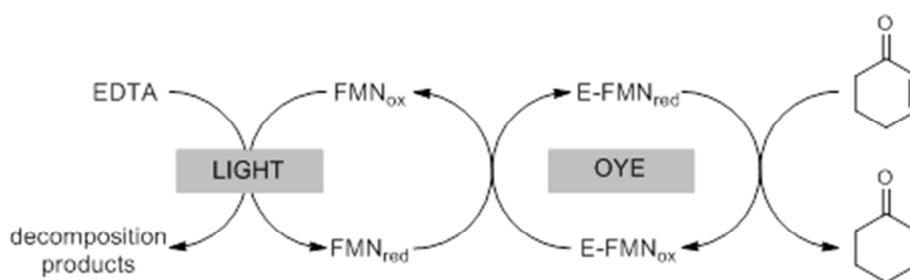
Figure 4.9:   Schematic representation of the light-driven cofactor regeneration pathway for the reduction of 2-cyclohexenone. E-FMN represents the enzyme-bound flavin group system (Grau *et al.,* 2009; Taglieber *et al.*, 2008)

Reactions were performed under the optimum pH and temperature conditions as determined for each enzyme, using substrates that fall within each enzyme's substrate scope, as previously determined. Reactions were performed under illumination (250 W bulb, 5 cm distance) and N_2 flux, as de-aeration of the reaction mixtures has been shown to greatly increase the conversion, compared to reactions under aerobic conditions (Grau *et al.*, 2009). Low conversions under aerobic conditions have been attributed to the formation of hydrogen peroxide by the reduced flavin in the presence of molecular oxygen, which has a detrimental effect on enzyme activity (Grau *et al*, 2009; Hollmann *et al.*, 2010). In addition, the presence

of hydrogen peroxide can result in the epoxidation of the alkene substrates (Hu *et al.*, 2002), rendering them inaccessible to the OYEs.

For OYE from *C. metallidurans* CH34, reactions with (R)-(-)-carvone, (S)-(+)-carvone and 2-cyclohexenone resulted in successful reduction of the substrates in conjunction with the light-driven cofactor regeneration approach. 2-cyclohexenone was successfully reduced by *Cm*OYE (figure 4.10), exhibiting 28% yield after 3 hours. *Cm*OYE exhibited >70% yield using (R)-(-)-carvone as substrate after 2 hours. Greater than 10% yield was observed for reactions with (S)-(+)-carvone, with the (2R,5S)-enantiomer of the product in excess. These results correspond with reduction of ketoisphorone and citral observed using YqjM when employing the light-driven system (Grau *et al.,* 2009; Taglieber *et al.,* 2008). In contrast, no reduction of the substrate dimethyl maleate was observed with *Ss*OYE.
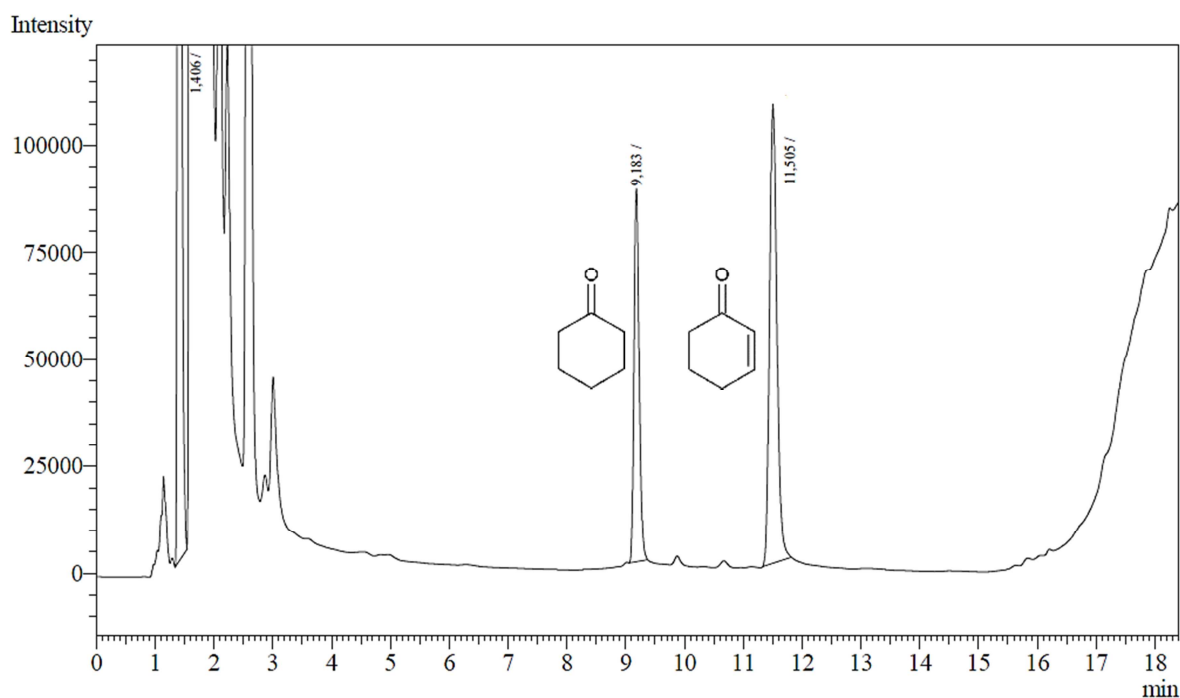
Figure 4.10:   Chromatogram obtained through GC analysis of reaction of *Cm*OYE with 2-cyclohexen-1-one   as   substrate   employing   the   light-driven   cofactor regeneration approach, showing peaks for 2-cyclohexenone and the product, cyclohexanone.

4.3.2.6.   *Reverse reactions*

In addition to the expected enone reduction, the ability of *Cm*OYE and *Ss*OYE to also catalyse the reverse (oxidation) reaction was investigated. Recent studies have shown the ability of the thermophilic OYE from *Geobacillus kaustophilus* to catalyse the reverse reaction (the desaturation of C-C bonds adjacent to a carbonyl) to give the corresponding α/β-unsaturated ketone (Schittmayer *et al.*, 2011).   This themophilic OYE homologue catalysed the conversion of numerous ketones to the corresponding α/β-unsaturated compound after 24 hours at 70°C.   These conversions included (amongst others) dihydrocarvone to carvone at 29% conversion, 2-methylcyclohexanone to 2-methylcyclohex-2-enone (56% conversion) and 2-methylcyclopentanone to 2-methylcyclopent-2-enone (56% conversion).   Reverse reactions were catalysed in the presence of molecular oxygen, without the presence of nicotinamide cofactor, but were attributed to the elevated reaction temperatures (Schittmayer *et al*, 2011).

The ability of *Cm*OYE and *Ss*OYE to catalyse the desaturation of C-C bonds adjacent to the carbonyl group at lower temperatures was investigated by using dimethyl succinate, (+)-dihydrocarvone, cyclohexanone and levodione (the product of the reduction of ketoisophorone by OYE) as substrates. These compounds were selected, as their corresponding α/β-unsaturated counterparts resulted in positive yields observed during substrate screening in 4.3.2.4. Reactions were performed at the optimum temperature and pH conditions as determined for each enzyme and incubated for 18 hours.

No product formation was observed for either *Cm*OYE or *Ss*OYE when utilizing dimethylsuccinate or levodione as substrates. Clear formation of product was observed for *Cm*OYE after 18 hours when utilizing (+)-dihydrocarvone (mixture of isomers) as substrate (figure 4.11). The formation of (S)-(-)-carvone was observed, with yields determined at 30% observed after 18 hours when utilizing (+)-dihydrocarvone as substrate. Due to the possible formation of hydrogen peroxide in the presence of molecular oxygen, the conversion of (+)-dihydrocarvone was repeated in the presence of catalase to determine if the removal of hydrogen peroxide will cause an improvement in the enzyme activity.   Although the formation of (S)-(-)-carvone was still observed (figure 4.12), no improvement in yield was observed, with yield determined at 19%.
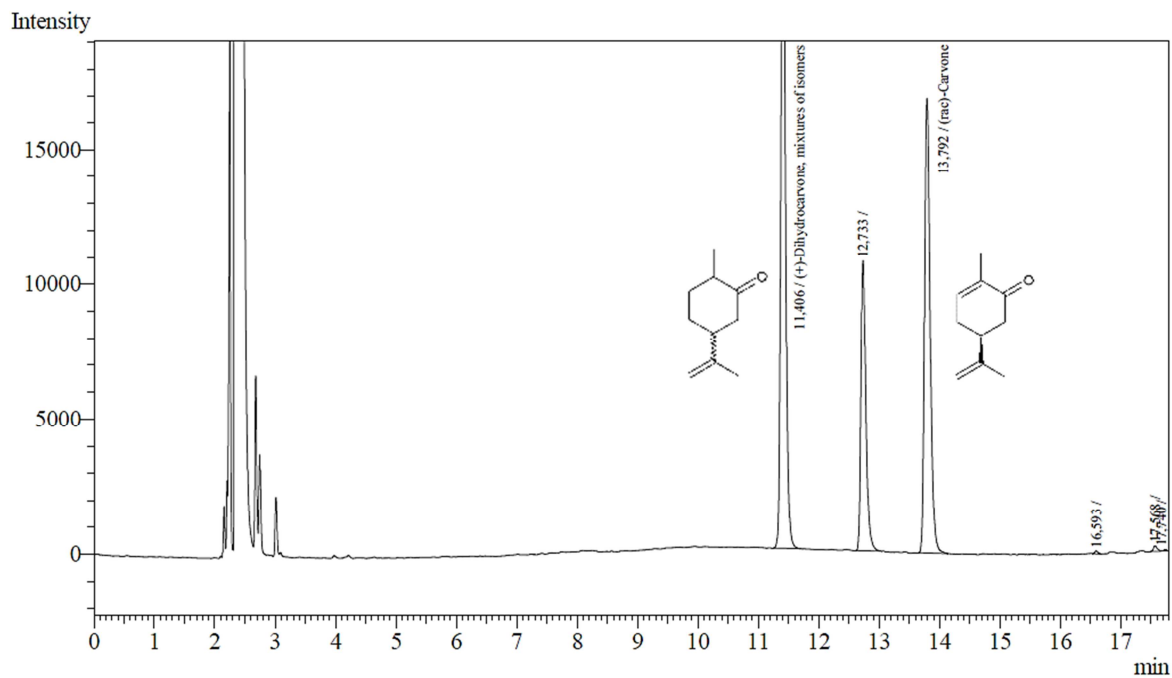
Figure 4.11: Chromatograms obtained through GC analysis of reverse reaction catalysed by *Cm*OYE with (+)-dihydrocarvone (mixture of isomers) as substrate, showing peaks for the isomers of (+)-dihydrocarvone and the product, (S)-(-)-carvone.
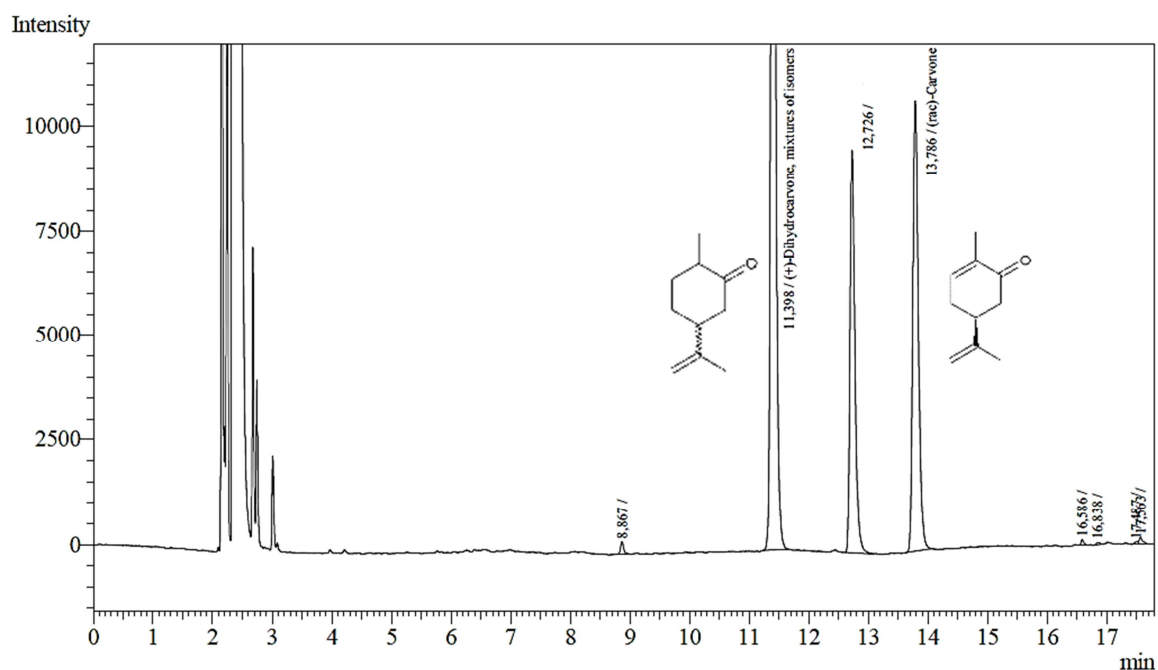
Figure 4.12: Chromatograms obtained through GC analysis of reverse reaction catalysed by *Cm*OYE with (+)-dihydrocarvone (mixture of isomers) as substrate in the presence of catalase, showing peaks for the isomers of (+)-dihydrocarvone and the product, (S)-(-)-carvone.

No product formation was observed for *Ss*OYE when utilizing cyclohexanone as substrate. However, for *Cm*OYE clear formation of product 2-cyclohexenone was observed after 18 hours when utilizing cyclohexanone as substrate (11% conversion). As with reactions utilizing (+)-dihydrocarvone, the conversion of cyclohexanone to 2-cyclohexenone was repeated in the presence of catalase to determine if the removal of hydrogen peroxide will cause an improvement in the enzyme activity. Although the formation of 2-cyclohexenone was still observed (figure 4.13), once again no improvement in yield was observed.
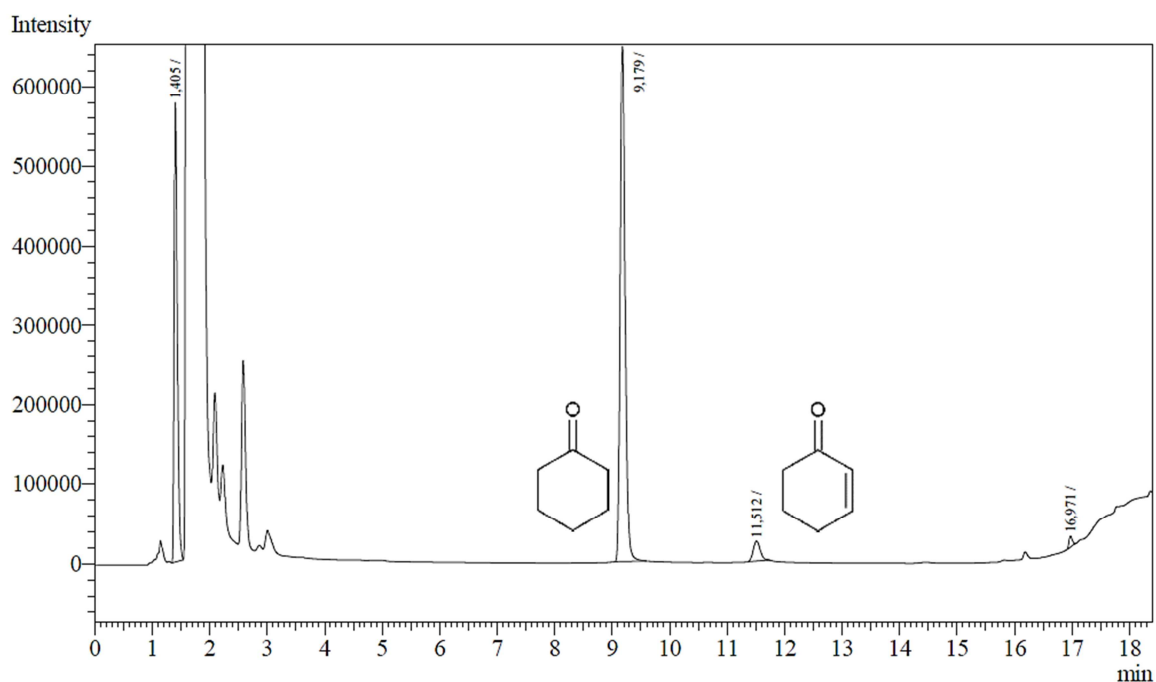
Figure 4.13: Chromatograms obtained through GC analysis of reverse reaction catalysed by *Cm*OYE with cyclohexanone as substrate in the presence of catalase, showing peaks for cyclohexanone and the product, 2-cyclohexenone.

4.3.3. Crystallisation of OYEs from *C. metallidurans* CH34 and *S. solfataricus*

Initial screening of conditions for the crystallization of OYEs from *C. metallidurans* CH34 *S. solfataricus* P2 was performed through the employment of commercially available screening kits (Hampton Research) for the screening of a vast number of potential crystallization conditions. These kits employ sparse matrix screening, a method that screens a combination of conditions (varying pH, buffer, additive and precipitant) with deliberate bias towards conditions that have previously been shown to generate crystals (Jancarik & Kim, 1991). Sparse matrix screening is the recommended initial screening step for proteins that have not yet been successfully crystallised.

For *Cm*OYE, initial screening revealed successful crystal growth of characteristically yellow crystals (typically coloured due to the oxidized flavin) within 2-5 days at numerous conditions, most notably with 0.2 M ammonium citrate containing 20% (w/v) polyethyleneglycol (PEG) 3 350 as precipitant, as well as with 0.1 M sodium malonate pH 6.0 containing 12% (w/v) PEG 3 350 as precipitant. However, these crystals were unfit for x-ray diffraction analysis as they were not single crystals with single nucleation sites (figure 4.14). These two conditions were thus further optimized by the use of two sets of grid screens and employing the hanging-drop vapour diffusion method. The first consisted of 0.2 M ammonium citrate with pH ranging from pH 4 to pH 6 in intervals of 0.5, combined with 10%, 12% and 15% (w/v) PEG 3 350. The second consisted of 0.1 M sodium malonate with pH ranging from pH 5 to pH 6.5 in intervals of 0.5, combined with 10%, 12% and 15% (w/v) PEG 3 350. In both cases, protein concentrations were also varied at 4, 6 and 8 mg.mL^{-1}.

Screening for the crystallisation of *Ss*OYE revealed the successful growth of characteristically yellow crystals (figure 4.14) within 2-5 at a limited number of conditions, including 0.15 M cesium chloride containing 15% PEG 3 350 as precipitant, 0.1 M Tris pH 8.5 containing 8% PEG 8 000, as well as 0.1 M citric acid pH 3.5 containing 5% 2-propanol and 6% PEG 20 000 as precipitant. As observed with the initial screening for *Cm*OYE crystal growth, multiple nucleation sites were observed in these crystal, thus rendering the crystals unsuitable for X-ray diffraction analysis. The above-mentioned conditions (with the exception of 0.1 M Tris pH 8.5 containing 8% PEG 8 000) were further optimized, using the hanging-drop vapour diffusion method. The first set of grid screens consisted of 0.12 M, 0.15M and 0.18 M cesium chloride, combined with 10%, 12%, 15% and 18% PEG 3 350 respectively. The second set of grid screens consisted of 6% PEG 20 000, combined with 5%, 8% and 10% 2-propanol, as well as 0.1 M citric acid at pH 3.5, 4, 4.5, 5 and 5.5. A third set of grid screens combined 5% 2-propanol with 6%, 10% and 15% PEG 20 000, as well as

0.1 M citric acid pH 3.5, 4, 4.5, 5 and 5.5.   In all cases, protein concentrations were also varied at 4, 6 and 8 mg.mL^{-1}.

After optimisation, single crystals were obtained for *Cm*OYE with 0.1 M sodium malonate pH 6 with 12% (w/v) PEG 3 350 and 6 mg.mL^{-1} protein. From the initial screening conditions, single crystals were also obtained at the following conditions: 0.1 M ammonium citrate, pH 5 with 12% (w/v) PEG 3 350 with 6 mg.mL^{-1} and 8 mg.mL^{-1} protein; 0.1 M sodium citrate tribasic dehydrate, pH 5.0 with 10% (w/v) PEG 6 000 at 8 mg.mL^{-1} protein; 0.2 M potassium sodium tartrate tetrahydrate, 0.1 M BIS-TRIS, pH 6.5 with 10% (w/v) PEG 10 000 with 8 mg.mL^{-1} protein; 4% (v/v) tacsimate, pH 6 with 12% (w/v) PEG 3 350 and 8 mg.mL^{-1} protein; 0.2 M succinic acid pH 7 with 20% PEG 3 350 and 6 mg.mL^{-1} protein.

Despite repeated attempts, optimisation of conditions failed to produced single crystals of *Ss*OYE suitable for X-ray diffraction analysis.   As a result, X-ray diffraction data was collected only for *Cm*OYE.
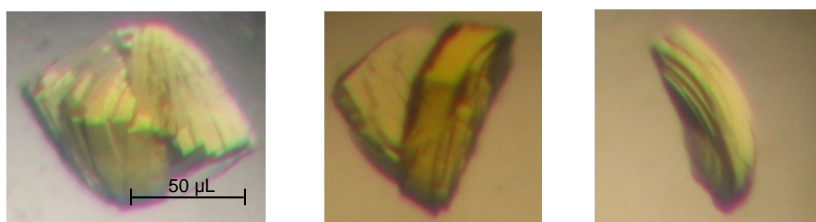
Figure 4.14:   Crystals of *Cm*OYE (with multiple nucleation sites) obtained through initial screening of crystallisation conditions.

4.3.4. Collection of X-ray diffraction data from *Cm*OYE crystals

Diffraction data from *Cm*OYE crystals were collected at a wavelength of 0.9173 Å. The data collection statistics for the crystal are given in table 4.3, with a typical diffraction pattern of an *Cm*OYE crystal shown in figure 4.15. Diffraction data was obtained with a resolution range of 57.99-1.93 Å. The crystal belonged to the space group $P2_12_12$ with unit cell parameters a=95.53, b=69.0, c=77.89 Å. Determination of the structure of *Cm*OYE by molecular replacement is currently being performed.
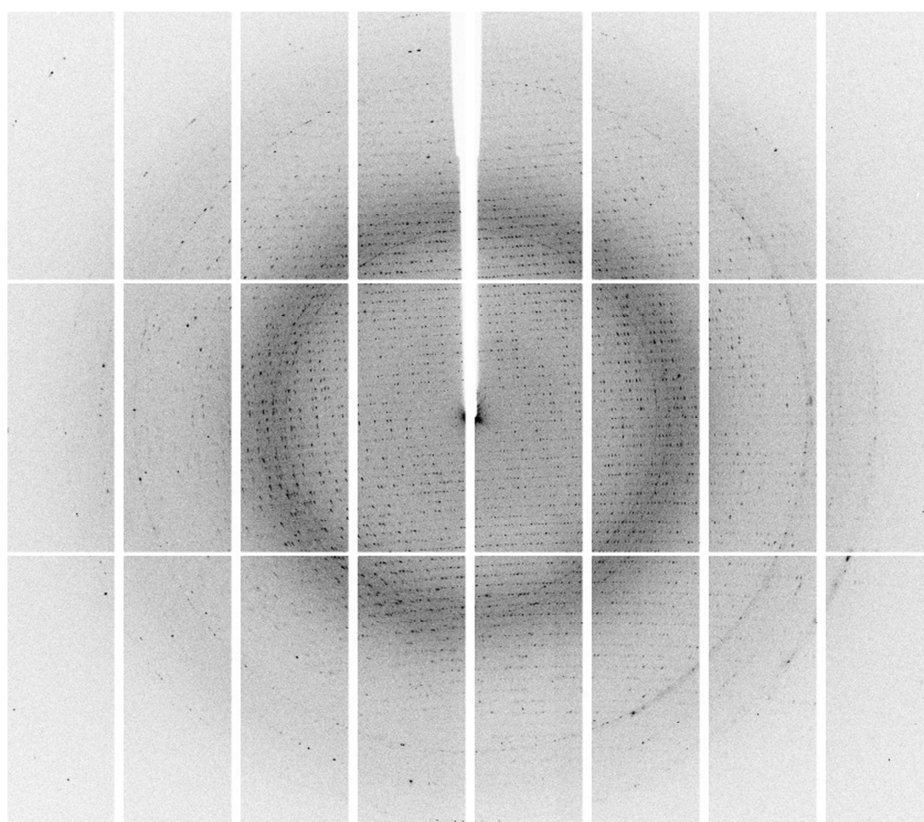
Figure 4.15:   Typical diffraction pattern of an *Cm*OYE crystal

Table 4.3: Crystal data and data-collection statistics for crystals of *Cm*OYE. Values in parentheses are for the highest resolution shell.

Space Group	$P2_1\,2_1\,2$
Unit-cell parameters (Å)	
a	95.53
b	208.75
c	77.89
Resolution range (Å)	57.99 - 1.93 (1.98 - 1.93)
Observed reflections	668291 (27757)
No. of unique reflections	110402 (7329)
Completeness (%)	93.6 (85.8)
Multiplicity	6.1 (3.8)
$\langle I/\sigma(I)\rangle$	10.4 (2.2)
$R_{merge}(\%)^{\#}$	0.119 (0.574)

[#]$R_{merge} = \Sigma_{hkl}\,\Sigma_i\,|I_i(hkl) - (I(hkl))| / \Sigma_{hkl}\,\Sigma_i\,I_i(hkl)$, where *$I_i(hkl)$* is the *i*th observation of reflection *hkl* and *<I(hkl)>* is the weighted average intensity for all observations of reflection *hkl*.

4.4. Conclusions

The two OYE homologues from *S. solfataricus* P2 and *C. metallidurans* CH3 were successfully heterologous expressed in *E. coli* and both enzymes were purified to homogeneity. Physicochemical characterization revealed that both enzymes function optimally near the optimal growth temperature and pH of their original bacterial sources. *Cm*OYE occurred as a monomer in solution, similar to XenA, with apparent monomer molecular mass of approximately 40 kDa, close to the predicted 42 kDa for the His_6-tag fused protein. Similarly, *Ss*OYE also occurred as a monomer in solution, with an apparent molecular mass of approximately 60 kDa, close to the predicted 65 kDa for the His-tag protein. The monomeric nature of *Ss*OYE is expected to be due to the 2,4,-dienoyl-CoA reductase-like second domain.

The catalytic efficiencies towards the substrate 2-cyclohexenone was higher for *Cm*OYE than for *Ss*OYE, but still significantly lower than that observed for other members of the "thermophilic-like" subclass of OYEs. Preliminary substrate scope and selectivity data revealed marked differences in the two OYEs catalytic preferences. Comparison of the two OYEs catalytic preferences to that of members of the "thermophilic-like" OYE group appeared complex: similarities were observed in the reduction of cyclic enones with methyl substitutions on the Cβ position, ketoisophorone and maleimide substrates, while differences were observed among the OYE homologues regarding the reduction of citral, carvone and cyclic enones with methyl substitutions on the Cα position.

*Cm*OYE exhibited similarity towards YqjM with successful reduction of substrates in conjunction with the light-driven cofactor regeneration approach. In contrast, reductions with *Ss*OYE employing the cofactor regeneration system were unsuccessful. In yet another difference between the two OYE homologues, *Cm*OYE successfully catalysed the reverse reaction as has been observed with the thermophilic OYE from *Geobacillus kaustophilus*, forming carvone from (+)-dihydrocarvone and 2-cyclohexenone from cyclohexanone. However, the study demonstrated the ability of the OYE homologue to catalyse the reverse reaction at temperatures lower than previously reported. *Ss*OYE failed to catalyse any of the tested reverse reactions.

Lastly, the successful crystallisation of *Cm*OYE and collection of X-ray diffraction data adds yet another OYE homologue for which structural data can be successfully determined.

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH

5.1. CONCLUSIONS

The old yellow enzyme (OYE) family, being a large and diverse group of flavoproteins with a broad substrate range, is a valuable resource in the search for potential new biocatalysts for the development of new, or the improvement of established, industrial processes (Toogood *et al.*, 2010; Williams & Bruce, 2002). Utilising a nicotinamide cofactor (NADH or NADPH) as reductant, OYE catalyses the reduction of activated alkenes by means of a two stage, bi-bi ping-pong mechanism. During the reductive half-reaction, the nicotinamide cofactor is oxidised by hydride transfer to the FMN cofactor of the OYE (Breithaupt *et al.*, 2001; Kohli & Massey, 1998). The oxidative half-reaction involves oxidation of the enzyme through hydride transfer from the reduced FMN to the activated C=C of the alkene substrate, followed by transfer of a solvent-derived proton to the Cα of the substrate. The oxidation half-reaction results in the net *trans* addition of [2H], occurring with absolute stereospecificity (Hall *et al.*, 2006; Karplus *et al.*, 1995; Kohli *et al.*, 1998).

From the catalytic capabilities and ever-broadening substrate range reported in literature, it can be concluded that the identification and study of novel members of the OYE family and the subsequent additions to the toolbox of available and characterised OYEs are vital in the search for new industrial biocatalysts. A vast variety of compounds have been shown to act as oxidative substrates for OYEs, all of which contain an electron-withdrawing (or activating) substituent (Stuermer *et al.*, 2007). Industrially-relevant, highly stereospecific reactions involving OYEs include (among others) the reduction of α-substituted cinnamaldehyde to L-phenylalaninol (an important intermediate in the synthesis of some HIV-1-protease inhibitors) by baker's yeast (Fardelone *et al.*, 2004; Fronza *et al.*, 2009), as well as the reduction of 2-cyclohexenone to cyclohexanone, a ketone produced to a large scale industrially (Adalbjörnsson *et al.,* 2010; Fryszkowska *et al.*, 2009), and the production of the precursors in carotanoid synthesis, levodione and actinol, through reduction of ketoisophorone (Hall *et al.,* 2008a; Hall *et al.*, 2008b; Wada *et al.*, 2003). Further strengthening the viability of OYEs as biocatalysts is the ability of the enzymes to be used in processes that avoid the need for costly nicotinamide cofactors in large-scale industial processes. These processes include the succesful use of strictly-controlled whole-cell biotransformations (Stuermer *et al.*, 2007; Toogood *et al.*, 2010),

as well as the utilization of purified OYE in conjunction with a cofactor regeneration system (Fryszkowska *et al.*, 2009; Grau *et al.*, 2009; Hall *et al.*, 2007; Hall *et al.*, 2008a; Taglieber *et al.*, 2008).

The wealth of sequence and structural data for members of the OYE family plays a pivotal role in understanding the catalytic mechanism and functional characteristics of these enzymes (Toogood *et al.*, 2010). Mutagenesis studies have revealed the importance of certain conserved active-site residues in the binding of substrates (Brown *et al.*, 1998) and in the catalysis of the oxidative half-reaction (Kohli & Massey, 1998), as well as in the binding and regulation of the flavin cofactor (Brown *et al.*, 2002; Spiegelhauer *et al.*, 2010; Xu *et al.*, 1999). In-depth knowledge of the structural and catalytic characteristics of OYEs is a valuable resource in the process of targeting and predicting the industrial viability of OYEs as biocatalysts (Toogood *et al.*, 2010). A recent study correlated the protein structures available for 12 unique OYEs with their respective functional characteristics and reported a strong association between the identity of two substrate-binding residues and the enantioselectivity of the enzymes (Oberdorfer *et al.*, 2011). The study subsequently divided OYEs into two subclasses – "classical" or OYE1-like OYEs and "thermophilic-like" or YqjM-like OYEs. While the identification of residues that can be used to effectively predict the functional characteristics of yet uncharacterised OYEs is highly desired, the technique is complicated by the fact that the targeted residues occur at a distance from each other (and from known important catalytic residues) in a linear amino acid sequence and were only identified as role-players due to the availability of the OYEs' crystal structures. When considering that a mere 12 unique OYEs have been structurally characterised, compared to the hundreds of gene sequences of yet uncharacterised OYEs available in databases, the need for an amino acid sequence fingerprint that can be used to effectively predict the functional characteristics of potential OYE targets becomes undeniable.

This study reported the possible existence of such a fingerprint, which targets two catalytic residues known to play a role in the binding of substrate (His191 and Asp194, OYE1 numbering; His164 and His167, YqjM numbering), in combination with a third residue which has been implicated in modulating the redox potential of the enzyme-bound flavin (Thr37, OYE1 numbering; Cys26, YqjM numbering). When grouping the large number of available OYE amino acid sequences among the Firmicutes, 3 classes of Proteobacteria and the Archaea according to the identity of the three above-mentioned residues, groupings were consistently highly comparable to that obtained through computationally intensive Maximum Likelihood phylogenetic algorithms. In addition to the two previously-mentioned subclasses reported in literature ("classical" and "thermophilic-like"), analysis of OYE sequences according to the 3-residue fingerprint identified two additional subclasses among OYEs of the Proteobacteria, as well as two new subclasses of OYE homologues among the

Firmicutes, and a further two new OYE subclasses among homologues from the Archaea. As members of the same OYE subclass have been reported to exhibit highly comparable enantioselective characteristics (Oberdorfer *et al.*, 2011), it is tempting to deduce that the OYEs belonging to the newly identified subclasses will also share functional characteristics and that the 3-residue motif identified in this study can be used as a possible sequence fingerprint to predict the functionality OYEs. An overview of the available literature revealed that all the OYE homologues for which functional and structural characterisation have been reported, belong to either the "classical" or the "thermophilic-like" subclasses. As it is clear that a correlation between the subclass grouping and functionality of OYE homologues can only be revealed with the availability of structural and functional data of yet uncharacterised OYEs, the remainder of the study was focused on the cloning, heterologous expression and characterisation of two new OYE homologues.

Two OYE homologues, one each from *C. metallidurans* CH34 and *S. solfataricus* P2 were targeted, amplified using PCR and cloned into the appropriate vectors for expression in a mesophilic host. The OYE from *C. metallidurans* CH34 was selected as it is a yet uncharacterised OYE homologue from a mesophilic organism, belonging to an OYE subclass that is dominated by OYEs from thermophilic organisms. Characterisation of the OYE provided vital data for comparison with other OYE homologues belonging to the same subclass. The second OYE, from *S. solfataricus*, was selected firstly as it belongs to an OYE subclass that currently contains no cloned or characterised OYEs. The homologue exhibits an unconventional OYE structure, bearing a strong resemblance to the 2,4-dienoyl-CoA reductase from *E. coli*. Lastly, cloning and characterisation of these two new OYE homologues served to provide two additional OYEs that can potentially be used as biocatalysts.

The successful expression and purification of active enzyme for both OYE homologues allowed for the characterisation of these two new members of the OYE family. Analysis of the purified enzymes revealed that both exist as monomers in solution and that the optimum reaction conditions for both OYEs corresponded well with the optimum growth conditions of their source organisms. Steady-state kinetics performed using the ubiquitous OYE substrate 2-cyclohexenone revealed significant differences in the kinetic constants of the two homologues, as well as compared to the kinetic constants reported for OYE homologues belonging to the same OYE subclasses. Investigation of the substrate range of the two OYEs further revealed significant differences in the functional characteristics not only between the two OYEs targeted in the study, but also in comparison to OYE homologues belonging to similar subclasses. Further functional characterisation of the two targeted OYE homologues revealed greater success in the utilization of the OYE from *C. metallidurans* CH34 for the reduction of substrates in conjunction with a light-driven cofactor regeneration approach, than was observed for the OYE homologue from *S. solfataricus* P2. In addition, the OYE from *C.*

metallidurans CH34 exhibited the ability to catalyse the reverse of the conventional OYE reaction at ambient temperatures – a phenomenon reported only for the OYE from *G. kaustophilus* and previously attributed to high reaction temperatures.

The successful crystallization of the two OYE homologues and the collection of diffraction data for the OYE from *C. metallidurans* CH34 confirmed that crystallographic techniques can be implemented to study the structural characteristics of the two OYEs beyond the construction of homology models.

In conclusion, this study served to add to the available toolbox of OYEs that serve as an important resource in the search for new biocatalysts. In particular, the OYE from *C. metallidurans* CH34 exhibits a reasonably wide substrate range, the ability to catalyse reductions in conjunction with a cost-effective, light-driven cofactor regeneration approach, as well as the ability to catalyse the reverse of the conventional OYE reduction reaction at ambient temperatures. In addition, crystallization of the OYE from *C. metallidurans* CH34 makes it an ideal candidate for further structural analysis and study, including possible directed evolution studies. Although exhibiting significantly lower catalytic efficiency and a narrower substrate range than other characterised OYE homologues, the OYE form *S. solfataricus* is a valuable biocatalytic asset due to its stability at high temperatures and non-physiological pH. In addition, the unconventional structure of the OYE homologue and the ability to crystallise, provides an opportunity to gain valuable insight into the mechanism of the enzyme and provides an ideal platform for future mutagenic studies.

Due to the great diversity observed among OYE homologues and the vast collection of OYE gene sequences available in databases, selection of a target OYE homologue for further analysis with the aim of providing industry with an efficient biocatalysts can be problematic. While a strong correlation between the functional characteristics (pertaining to enantioselectivity) and the sequence-based subclass an OYEs belong to has been reported (Oberdorfer *et al.*, 2011), this study indicated that care must be taken when drawing conclusions regarding an OYE's substrate specificity from its amino acid structure.

5.2. FUTURE RESEARCH

The cloning, functional and structural characterisation of many more OYEs – particularly from yet 'unexplored' subclasses – is vital in order to examine the validity of the 3-residue fingerprint identified in this study for use in the prediction of the functionality of new OYEs from sequence data. In addition, further characterisation of the OYE homologues from *C. metallidurans* CH34 and *S. solfataricus* P2 will be performed, including the screening of a wider range of substrates and the steady-state kinetic constants for more substrates.

The completion of the crystals structure of the OYE from *C. metallidurans* CH34 will provide valuable insight into the structure-function relationships of OYEs in the same subclass. The optimization of conditions for the crystallization of the OYE from *S. solfataricus* P2 and the subsequent collection of X-ray diffraction data and structure determination is imperative in order to gain understanding of the reaction mechanism of this unconventional OYE. Once structurally characterised, both OYE homologues will serve as valuable platforms for future directed evolution studies, particularly towards improving the catalytic efficiencies and substrate scope of the enzymes.

While the study provided basic evolutionary analysis of the vast number of available OYE sequences, the analysis was by no means exhaustive. Further in depth evolutionary analysis of OYE amino acid sequences may provide valuable information regarding the evolution of protein structure and function, the importance of conserved residues and the possible identification of future targets for mutagenic studies.

SUMMARY

The old yellow enzyme (OYE) family is a diverse group of flavoenzymes that catalyse the asymmetric reduction of activated C=C bonds of a wide variety of α/β-unsaturated carbonyl compounds. OYEs are attractive as biocatalysts due to the ability to perform *trans*-hydrogenation with high stereospecificity (Stuermer *et al.*, 2007). The number of functionally and structurally characterised OYEs has grown over the past decade, as have the enzyme family's substrate spectrum.

Vital in the search for new industrial biocatalysts among the OYE family is the structural and functional characterisation of new OYE homologues (Oberdorfer *et al.*, 2011; Toogood *et al.*, 2010; Williams and Bruce, 2002). This study investigated the sequence-based evolutionary relationship among the vast number of OYE homologues in the Proteobacteria, Firmicutes and Archaea, with particular attention paid to two substrate-binding residues in the catalytic site and a single residue implicated in the modulation of the redox potential of enzyme-bound FMN. A strong correlation was identified between grouping of OYE homologues through advanced maximum likelihood evolutionary analysis, and the grouping of OYEs into subgroups through the identity of the above-mentioned three target residues.

Two OYE homologues were selected for cloning, heterologous expression and subsequent characterisation. The first OYE (*Cm*OYE) was selected from the mesophile *Cupriavidus metallidurans* CH3 and belongs to the previously identified "thermophilic-like" subclass of OYEs (Toogood *et al.*, 2010), which includes mainly OYEs from thermophiles, but also OYEs YqjM from *B. subtilis* (Kitzing *et al.*, 2003) and XenA (*P. putida*; Griese *et al.*, 2006) from mesophiles. *Cm*OYE was regarded as an ideal target, as it adds to the short list of mesophilic OYEs from the subclass. The second OYE (*Ss*OYE) was selected form the hyperthermophile *Sulfolobus solfataricus* P2 and belongs to an as-yet uncharacterised subclass of OYEs. *Ss*OYE was regarded as an ideal target due to its potentially high thermal stability (an ideal characteristic in biocatalysts) and unconventional OYE structure (bearing high similarity to the two-domain enoyl-CoA reductase from *E. coli*, as revealed with the aid of homology modelling form the translated nucleotide sequence).

Both targeted OYEs were successfully cloned and heterologously expressed in *E. coli* as both unmodified and modified (with addition of N-terminal His6-tag). Due to the presence of rare codons in the gene sequence of *Ss*OYE, the protein was expressed in *E. coli* in the

presence of the plasmid pLySSRARE2. Purification of *Cm*OYE and *Ss*OYE through IMAC and size-exclusion chromatography provided homogenous protein solutions and revealed that *Cm*OYE and *Ss*OYE are present in solution as a monomers, with the monomeric nature of SsOYE possibly due to the presence of a second domain. Temperature and pH profiles of the two OYEs revealed an optimum temperature and pH for *Cm*OYE that corresponds well to the optimum growth conditions of the source organism. The optimum catalytic temperature of *Ss*OYE was identified to be significantly lower than that of the source organism, while the optimum catalytic pH was higher than the optimum growth pH of *S. solfataricus* P2.

Steady-state kinetics performed for both *Cm*OYE and *Ss*OYE with 2-cyclohexenone as substrate revealed that catalytic efficiency and affinity differed vastly between the two OYE homologues. While the K_m for *Cm*OYE was found to be comparable with the close homologue from *Thermus scotoductus* SA-01 (CrS; Opperman *et al.*, 2010), catalytic efficiency for both *Cm*OYE and *Ss*OYE was revealed to be significantly lower than that observed for the close homologues YqjM, XenA and CrS.

*Ss*OYE revealed a more limited substrate scope compared to *Cm*OYE. Maleimides were identified as good substrates for both, corresponding to activities reported for YqjM, XenA and CrS. No conversions were observed for cyclic enones with methyl substitutions on the Cβ position. However, certain compounds previously reported to act well as substrates for YqjM and XenA were not accepted as substrates by either *Cm*OYE or *Ss*OYE, or resulted in significantly lower conversion as reported. Neither *Cm*OYE nor *Ss*OYE exhibited activity towards citral, an enal with a methyl group on Cβ that has been reported to act as substrate for YqjM and XenA but not for CrS. *Cm*OYE exhibited marginal activity towards the Cα methyl-substituted 2-methylcyclopentanone and none towards 2-methylcyclohexenone, two compounds for which conversions have been reported for YqjM. Two isomers of carvone were successfully reduced by both *Cm*OYE and *Ss*OYE, an activity not exhibited by XenA. Ketoisophorone, a known substrate for YqjM, resulted in marginal conversion for both *Ss*OYE and *Cm*OYE, with *Ss*OYE exhibiting almost double the conversion observed for *Cm*OYE.

*Cm*OYE catalysed the conversion of carvones and 2-cyclohexenone in the absence of nicotinamide cofactor, but in conjunction with a light-driven cofactor regeneration approach. Utilisation of this cofactor-regeneration approach failed to produce any positive results in *Ss*OYE. Further functional characterisation involved investigating the enzymes' ability to catalyse the dehydrogenation of saturated ketones. This phenomenon has been reported

for the OYE from the thermophile *Geobacillus kaustophilus (*Schittmayer *et al.,* 2011) in the absence of nicotinamide cofactor and utilizing only molecular oxygen, but has been attributed to elevated reaction temperatures (70°C). Although not successful for *Ss*OYE, *Cm*OYE catalysed the conversion of cyclohexanone and (+)-dihydrocarvone to their corresponding unsaturated compounds at 25°C.

Lastly, crystallisation of *Cm*OYE was performed for the collection of X-ray crystallographic data for future structural characterisation. The enzyme was successfully crystallised and diffraction data collected at a resolution range of 57.99 - 1.93Å.

The study demonstrated that predicting functional characteristics from sequence data remains problematic for members of the OYE family.   Although the use of three catalytically important target residues as fingerprint is useful for elucidating the evolutionary relationship among the vast number of OYE homologues, these groupings do not necessarily result in the clustering of OYEs with similar functional characteristics. The need for more functional and structural data of OYE homologues (especially OYE homologues belonging to yet uncharacterised subgroups) remains if correlation between sequence similarity and functional similarity among OYE homologues is to be elucidated.

Keywords

Old yellow enzyme, enoate reductase, α,β-unsaturated carbonyl, cofactor regeneration, NAD(P)H-dependent oxidoreductase, flavin oxidoreductase, biocatalysis, maximum likelihood evolutionary analysis, extremozymes.

OPSOMMING

Die 'old yellow enzyme' (OYE) familie is 'n diverse groep flavo-ensieme wat die asimmetriese reduksie van geaktiveerde C=C verbindings van 'n wye verskeidenheid α/β-onversadigde karbonielverbindings, kataliseer. OYEs is gesogte biokataliste as gevolg van die vermoë om *trans*-hidrogenasie met 'n hoë stereoselektiwiteit uit te voer (Stuermer *et al.* 2007). Die aantal funksioneel- en struktureel gekarakteriseerde OYEs het oor die afgelope dekade gegroei, en so ook die ensiem-familie se substraat spektrum.

Die strukturele en funksionele karakterisering van nuwe OYE homoloë (Oberdorfer *et al*, 2011; Toogood *et al*, 2010; Williams en Bruce, 2002) is noodsaaklik in die soektog na nuwe industriële biokataliste onder die OYE familie is. Hierdie studie het die aminosuuropeenvolging-gebaseerde evolusionêre verband tussen die groot aantal OYE homoloë in die proteobakterieë, firmikiete en archaea, ondersoek. Besondere aandagis gegee aan twee substraat-bindende residu's in die katalitiese setel, asook 'n enkele residuwat in die regulering van die ensiem-gebonde FMN se redokspotensiaal, geïmpliseer is. 'n Sterk korrelasie tussen die groepering van OYE homoloë deur middel van gevorderde maksimum waarskynlikheid evolusionêre analise, en die groepering van OYEs in die subgroepe deur die identiteit van die bogenoemde drie teiken-residu's, is geïdentifiseer.

Twee OYE homoloë is gekies vir klonering, heteroloë uitdrukking en die daaropvolgende karakterisering. Die eerste OYE (*Cm*OYE) is gekies uit die mesofiel *Cupriavidus metallidurans* CH3 en behoort aan die voorheen geïdentifiseerde "termofiel-agtige" subklas van OYEs (Toogood *et al*, 2010), wat hoofsaaklik OYEs van termofiele, maar ook OYEs YqjM uit *B. subtilis* (Kitzing *et al.*, 2003) en XenA (*P. putida*; Griese, *et al.*, 2006) van mesofiele, insluit. C*m*OYE is beskou as 'n ideale teiken, omdat dit bydra tot die min voorbeelde van mesofiele OYEs in hierdie subklas. Die tweede OYE (*Ss*OYE) van die hipertermofiel *Sulfolobus solfataricus* P2 behoort aan 'n tans ongekarakteriseerde subklas van OYEs. *Ss*OYE is beskou as 'n ideale teiken as gevolg van die potensiële hoë termiese stabiliteit ('n ideale kenmerk in biokataliste), asook die onkonvensionele OYE struktuur (soortgelyk aan die twee-domein enoiel-KoA reduktase van *E. coli*, soos blootgelê met die hulp van homologie modellering vanaf die aminosuur opeenvolging).

Beide OYEs is suksesvol gekloneer en heteroloog uitgedruk in *E. coli* as beide ongemodifiseerde en gemodifiseerde (met die byvoeging van die N-terminale His_6-tag)

proteïen. As gevolg van die teenwoordigheid van seldsame kodons in die basispaaropeenvolging van *Ss*OYE, is die proteïen in *E. coli* uitgedruk in die teenwoordigheid van die plasmied pLySSRARE2. Homogene proteïen-oplossings is verkry deur suiwering van *Cm*OYE en *Ss*OYE deur IMAC en uitsluitingskromatografie. Die laasgenoemde suiweringstegniek het ook aangedui dat *Cm*OYE as 'n monomeer bestaan. *Ss*OYE as ook 'n monomeer in oplossing teenwoordig is, moontlik as gevolg van die teenwoordigheid van die tweede domein. Temperatuur en pH profiele van die twee OYEs se aktiwiteit het aangedui dat *Cm*OYE 'n optimale temperatuur en pH het wat ooreenstem met die optimum groeitoestande van die bron organisme. Die optimum katalitiese temperatuur van *Ss*OYE blyk aansienlik laer as dié van die bron organisme te wees, terwyl die optimum katalitiese pH hoër is as die optimum groei pH van *S. solfataricus* P2.

Gestadigte-toestand kinetika is uitgevoer vir beide *Cm*OYE en *Ss*OYE met 2-sikloheksenoon as substraat wat aangedui het dat die katalitiese doeltreffendheid en affiniteit aansienlik tussen die twee OYE homoloë verskil. Terwyl die K_m vir *Cm*OYE vergelykbaar met die van die OYE vanaf *Thermus scotoductus* SA-01 (CrS; Opperman *et al*, 2010) is, is die katalitiese doeltreffendheid vir beide *Cm*OYE en *Ss*OYE beduidend laer is as wat vir die homoloë YqjM, Xena en CRS waargeneem is

*Ss*OYE toon aktiwiteit teenoor 'n meer beperkte omvang substrate as *Cm*OYE. Maleimiedes is as goeie substrate vir beide geïdentifiseer, wat ooreenstem met die aktiwiteite gerapporteer vir YqjM, XenA en CrS. Geen omskakelings is waargeneem vir sikliese enone met metiel-vervangings op die Cβ posisie nie. Sekere verbindings wat voorheen as goeiesubstrate vir YqjM en XenA beskou is, is egter nie as substrate deur *Cm*OYE óf *Ss*OYE geïdentifiseer nie, of het gelei tot beduidend laer aktiwiteite as wat vir homoloë gerapporteer is. Nie *Cm*OYE of *Ss*OYE toon aktiwiteit teenoor sitral nie, 'n enal met 'n metielgroep op Cβ wat wel as substraat vir YqjM en XenA optree, maar nie vir CrS nie. *Cm*OYE toon marginale aktiwiteit met die Cα metiel-gesubstitueerde 2-metielsiklopentanoon en geen aktiwiteit met 2-metielsikloheksanoon nie, twee substrate waarvoor aktiwiteit met YqjM gerapporteer is. Twee isomere van karvoon is suksesvol deur beide *Cm*OYE en *Ss*OYE omgeskakel, 'n aktiwiteit wat nie vir XenA gerapporteer is nie. Beide *Ss*OYE en *Cm*OYE het baie lae aktiwiteit met ketoisoforoon, 'n bekende substraat vir YqjM, getoon. Vir *Ss*OYE is byna dubbel die omskakeling as wat vir *Cm*OYE waargeneem is, met ketoisoforoon getoon.

*Cm*OYE het die omskakeling van karvoon en 2-sikloheksenoon in die afwesigheid van 'n nikotienamied kofaktor, maar in samewerking met 'n lig-gedrewe kofaktor

regenereringsmetode, gekataliseer. Gebruik van hierdie kofaktor-regenereringsbenadering het geen positiewe resultate met *Ss*OYE geproduseer nie. Verdere funksionele karakterisering is gedoen deur die ensieme se vermoë om die dehidrogenasie van versadigde ketone te kataliseer, te ondersoek. Hierdie verskynsel is vir die OYE van die termofiel *Geobacillus kaustophilus* (Schittmayer *et al,.* 2011) in die afwesigheid van nikotienamied ko-faktor en met die gebruik van slegs molekulêre suurstof, gerapporteer. Hierdie verskynsel is egtertoe geskryf aan verhoogde reaksie temperatuur (70˚C). Hoewel geen sukses met *Ss*OYE behaal is nie, het *Cm*OYE die omskakeling van sikloheksanoon en (+)-dihidrokarvoon na die ooreenstemmende onversadigde verbindings by 25˚C gekataliseer.

Laastens is die kristallisasie van *Cm*OYE vir die versameling van X-straal kristallografiese data vir toekomstige strukturele karakterisering, uitgevoer. Die ensiem was suksesvol gekristalliseer en diffraksie data is versamel by 'n resolusie van 57.99-1.93 Å.

Dié studie het getoon dat die voorspelling van funksionele eienskappe vanaf aminosuuropeenvolging steeds problematies vir lede van die OYE familie is. Hoewel die gebruik van die drie geïdentifiseerde teiken-residue as vingerafdruk nuttig is vir aanduidings van die evolusionêre verhouding tussen die groot aantal OYE homoloë, kom hierdie groeperinge nie noodwendig ooreen met die groepering van OYEs met soortgelyke funksionele eienskappe nie. Die behoefte vir meer funksionele en strukturele data van OYE homoloë (veral OYE homoloë wat aan tans ongekarakteriseerde subgroepe behoort) bly groot as 'n korrelasie tussen aminosuuropeenvolging en funksionaliteit tussen OYE homoloë gevind wil word.

REFERENCES

Adachi, J. and Hasegawa, M. (1996). Model of amino acid substitution in proteins encoded by mitochondrial DNA. *Journal of Molecular Evolution*, **42**:459-468.

Adachi, J., Waddell, P. J., Martin, W. and Hasegawa, M. (2000). Plastid genome phylogeny and a model of amino acid substitution for proteins encoded by chloroplast DNA. *Journal of Molecular Evolution*, **50**:348-358.

Adalbjörnsson, B. V., Toogood, H. S., Fryszkowska, A., Pudney, C. R., Jowitt, T. A., Leys, D. and Scrutton, N. S. (2010). Biocatalysis with thermostable enzymes: structure and properties of a thermophilic "ene"-reductase related to old yellow enzyme. *Chembiochem*, **11**:197-207.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, **25**:3389-3402.

Baker-Austin, C. and Dopson, M. (2007). Life in acid: pH homeostasis in acidophiles. *TRENDS in Microbiology*, **15**:165-171.

Barna, T. M., Khan, H., Bruce, N. C., Barsukov, I., Scrutton, N. S. and Moody, P. C. E. (2001). Crystal structure of pentaerythritol tetranitrate reductase: "flipped" binding geometries for steroid substrates in different redox states of the enzyme. *Journal of Molecular Biology*, **310:**433-447.

Barna, T. M., Messiha, H. L., Petosa C., Bruce, N. C., Scrutton, N. S. and Moody, P. C .E. (2002). Crystal structure of bacterial morphinone reductase and properties of the C191A mutant enzyme. *The Journal of Biological Chemistry,* **277**:30976-83.

Barth, H. G., Boyes, B. E. and Jackson, C. (1994). Size exclusion chromatography. *Analytical Chemistry*, **66**:595R-620R.

Bornhorst, J. A. and Falke, J. J. (2000). Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology*, **326**:245-254.

Bougioukou, D., Kille, S., Taglieber, A. and Reetz, M. (2009). Directed Evolution of an enantioselective enoate-reductase: Testing the utility of iterative saturation mutagenesis. *Advanced Synthesis and Catalysis*, **351**:3287-3305.

Breithaupt, C., Strassner, J., Breitinger, U., Huber, R., Macheroux, P., Schaller, A. and Clausen, T. (2001). X-ray structure of 12-oxophytodienoate reductase 1 provides structural insight into substrate binding and specificity within the family of OYE. *Structure,* **9**:419-429.

Breithaupt, C., Kurzbauer, R., Lilie, H., Schaller, A., Strassner, J., Huber, R., Macheroux, P. and Clausen, T. (2006). Crystal structure of 12-oxophytodienoate reductase 3 from tomato: self-inhibition by dimerization. *Proceedings of the National Academy of Sciences of the United States of America,* **103**:14337-14342.

Brigé, A., Van den Hemel, D., Carpentier, W., De Smet, L. and Van Beeumen, J. J. (2006). Comparative characterization and expression analysis of the four Old Yellow Enzyme homologues from *Shewanella oneidensis* indicate differences in physiological function. *The Biochemical Journal*, **394:**335-344.

Brown, B. J., Deng, Z., Karplus, P. A., and Massey, V. (1998). On the Active Site of Old Yellow Enzyme. Role of histidine 191 and asparagine 194. *Journal of Biological Chemistry*, **273**:32753-32762.

Brown, B. J., Hyun, J. -W., Duvvuri, S., Karplus, P. A. and Massey, V. (2002). The role of glutamine 114 in old yellow enzyme. *The Journal of Biological Chemistry*, **277**:2138-2145.

Buque-Taboada, E. M., Straathof, A. J. J., Heijnen, J. J. and Van der Wielen, L. A. M. (2005). Microbial Reduction and *in situ* product crystallization coupled with biocatalyst cultivation during the synthesis of 6R-dihydrooxoisophorone. *Advanced Synthesis and Catalysis*, **347**:1147-1154.

Chaparro-Riggers, J. F., Rogers, T. A., Vazquez-Figueroa, E., Polizzi, K. M. and Bommarius, A. S. (2007). Comparison of three enoate reductases and their potential use for biotransformations. *Advanced Synthesis and Catalysis*, **349**:1521-1531.

Craig, D. H., Moody, P. C., Bruce, N. C. and Scrutton, N. S. (1998). Reductive and oxidative half-reactions of morphinone reductase from *Pseudomonas putida* M10: a kinetic and thermodynamic analysis. *Biochemical Journal*, **37**:7598-7607.

Dauphin, G., Gourcy, J. -G. and Vecshambre, H. (1992). Microbiological synthesis of optically active (2*R*,3*S*)-2,3-deuteriocyclohexan-1-ones and (2*R*,3*S*)-2-methyl-3-deuteriocyclohexan-1-one. Enantiospecific *anti*-addition of hydrogen to the double bond of cyclohex-2-en-1-ones. *Tetrahedron: Asymmetry* **3**: 595-598.

Demirjian, D. C., Morís-Varas, F. and Cassidy, C. S. (2001). Enzymes from extremophiles. *Current Opinion in Chemical Biology*, **5**:144-151.

Eck, R. and Simon, H. (1994). Preparation of (*S*)-2-substituted succinates by stereospecific reductions of fumarate and derivatives with resting cells of *Clostridium formicoaceticum*. *Tetrahedron* **50**:13631-13640.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**:1792-1792.

Edwards. T. E. and Staker, B. L. To be published.

Esterbauer, H. (1993). Cytotoxicity and genotoxicity of lipid-oxidation products. *American Journal of Clinical Nutrition*, **57**(Suppl):779S-786S.

Fardelone, L. C., Augusto, J., Rodrigues, R. and Moran, P. J. S. (2004). Baker's yeast mediated asymmetric reduction of cinnamaldehyde derivatives. *Journal of Molecular Catalysis B: Enzymatic*, **29**:41-45.

Fairbanks, G., Steck, T. L. and Wallach, F. H. (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry*, **10**:2606-2617.

Fitch, W. M. (1986). An estimation of the number of invariable sites is necessary for the accurate estimation of the number of nucleotide substitutions since a common ancestor. *Progress in Clinical and Biological Research*,

Fitch, W. M. and Margoliash, E. (1967). A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome c as a model case. *Biochemical Genetics*, **1**:65-71.

Fitzpatrick, T. B., Amrhein, N. and Macheroux, P. (2003). Characterization of YqjM. an old yellow enzyme homologue from *Bacillus subtilis* involved in the oxidative stress response. *Journal of Biological Chemistry*, **278**:19891-19897.

Fitzpatrick, T. B., Auweter, S., Kitzing, K., Clausen, T., Amrhein, N. and Macheroux, P. (2004). Structural and functional impairment of an old yellow enzyme homologue upon affinity tag incorporation. *Protein Expression and Purification*, **36**:280-291.

Fox, K. M. and Karplus, P. A. (1994). Old yellow enzyme at 2 Å resolution: overall structure, ligand binding, and comparison with related flavoproteins. *Structure* **2**:1089-1105.

Fronza, G., Fuganti, C. and Serra, S. (2009). Stereochemical course of baker's yeast mediated reduction of the tri- and tetrasubstituted double bonds of substituted cinnamaldehydes. *European Journal of Organic Chemistry*, **2009**: 6160-6171.

Fryszkowska, A., Toogood, H., Sakuma, M., Gardiner, J. M., Stephens, G. M. and Scrutton, N. S. (2009). Asymmetric reduction of activated alkenes by pentaerythritol tetranitrate reductase: specificity and control of stereochemical outcome by reaction optimisation. *Advanced Synthesis and Catalysis*, **351**: 2976-2990.

Godiska, R., Reuter, M., Schoenfeld, T., Sheets, L., Derr, A. and Mead, D. (2001). The CloneSmart Advantage: Easy cloning of "difficult" DNA targets. *eLucidations*, November, Issue 1.

Goris, J., De Vos, P., Coenye, T., Hoste, B., Janssens, D., Brim, H., Diels, L., Mergeay, M., Kersters, K. and Vandamme, P. (2001). Classification of metal-resistant bacteria form industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis* Steinle *et al*. 1998 emend. *International Journal of Systematic and Evolutionary Microbiology*, **51**:1773-1782.

Goretti, M., Ponzoni, C., Caselli, E., Marchigiani, E., Cramarossa, M. R, Turchetti, B., Buzzini, P. and Forti, L. (2009). Biotransformation of electron-poor alkenes by yeasts: asymmetric reduction of (*4S*)-(+)-carvone by yeast enoate reductases. *Enzyme and Microbial Technology*, **45**: 463-468.

Grau, M. M., van der Toorn, J. C., Otten, L. G., Macheroux, P., Taglieber, A., Zilly, F. E., Arends, W. C. E. and Hollmann, F. (2009). Photoenzymatic reduction of C=C double bonds. *Advanced Synthesis and Catalysis*, **351:**3279-3286.

Griese, J. J., Jakob, R. P., Schwarzinger, S. and Dobbek, H. (2006). Xenobiotic reductase A in the degradation of quinoline by *Pseudomonas putida* 86: physiological function,

structure and mechanism of 8-hydroxycoumarin reduction. *Journal of molecular Biology*, **361**:140-152.

Grisshammer, R. and Tucker, J. (1997). Quantitative evaluation of neurotensin receptor purification by immobilized metal affinity chromatography. *Protein Expression and Purification*, **11**:53-60.

Haas, E. (1938). Isolierung eines neuen gelben ferments. *Biochemische Zeitschrif*, **298**:378-390.

Hall, M., Hauer, B., Stuermer, R., Kroutil, W. and Faber, K. (2006). Asymmetric whole-cell bioreduction of an α,β-unsaturated aldehyde (citral): competing *prim*-alcohol dehydrogenase and C-C lyase activities. *Tetrahedron: Asymmetry*, **17**:3058-3062.

Hall, M., Stueckler, C., Kroutil, W., Macheroux, P and Faber, K. (2007). 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. *Angewandte Chemie*, **119**:4008-4011.

Hall, M., Stueckler, C., Hauer, B., Stuermer, R., Friedrich, T., Breuer, M., Kroutil, W. and Faber, K. (2008a). Asymmetric bioreduction of activated C=C bonds using *Zymomonas mobilis* NCR enoate reductase and old yellow enzymes OYE 1–3 from yeasts. *European Journal of Organic Chemistry*, **9**:1511-1516.

Hall, M., Stueckler, C., Ehammer, H., Pointner, E., Oberdorfer, G., Gruber, K., Hauer, B., Stuermer, R., Kroutil, W., Macheroux, P. and Faber, K. (2008b). Asymmetric bioreduction of C=C bonds using enoate reductases OPR1, OPR3 and YqjM: enzyme-based stereocontrol. *Advanced Synthesis and Catalysis*, **350**:411-418.

Hirata, T., Takarada, A., Matsushima, A., Kondo, Y. and Hamada, H. (2004). Asymmetric hydrogenation of N-substituted maleimides by cultured plant cells. *Tetrahedron: Asymmetry*, **15**:15-16.

Hollmann, F., Arends, I. W. C. E., Buehler, K. (2010). Biocatalytic redox reactions for organic synthesis: nonconventional regeneration methods. ChemCatChem **2**:762-782

Horikoshi, K. (2010). Extremophiles Handbook (Horikoshi, K., Bull, A. T., Antranikian, G., Robb, T. F. and Stetter, K. O., eds). pp. 267-328. Springer

Hu, Y. -Z., Tsukiji, S., Shinkai, S., Oishi, S., Hamachi, I. (2000). Construction of artificial photosynthetic reaction centres on a protein surface: vectorial, multistep, and proton-coupled electron transfer for long-lived charge separation. *Journal of the American Chemical Society*, **122**:241-253.

Hubbard, P. A., Liang, X., Schultz, H. and Kim, J. -J. P. (2003). The crystal structure and reaction mechanism of *Escherichia coli* 2,4-dienoyl-CoA reductase. *Journal of Biological Chemistry*., **278**:37553-37560.

Imhoff, J. M. (1995). Taxonomy and physiology of phototrophic purple bacteria and green sulphur bacteria. In: Anoxygenic photosynthetic bacteria (Blankenship, R. E., Madigan, M. T. and Bauer, C. E., eds), Chapter 1. Kluwer Academic Publishers, Dordrecht.

Jancarik, J. and Kim, S. -H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of Applied Crystallography*, **24**:409-411.

Jablonski, E. and DeLuca, M. (1977). Purification and properties of the NADH and NADPH specific FMN oxidoreductase from *Beneckea harveyi*. *Biochemistry*, **16**:2932-2936.

Jamieson, D. J. (1998). Oxidative stress responses of the yeast *Saccharomyces cerevisiae*. *Yeast*, **14**:1511-1527.

Jeffery, C. J. (2009). Moonlighting proteins - an update. *Molecular BioSystems*, **5**:345-350.

Karplus, P. A., Fox, K. M. and Massey, V. (1995). Structure-function relations for old yellow enzyme. *The FASEB Journal*, **9**:1518-1526.

Kataoka, M., Kotaka, A., Thiwthong, R., Wada, M., Nakamori, S. and Shimizu, S. (2004). Cloning and overexpression of the old yellow enzyme gene of *Candida macedoniensis* and; its application to the production of a chiral compound. *Journal of Biotechnology*, **114**:1-9.

Kawai, Y., Hayashi, M. and Tokitoh, N. (2001). Asymmetric synthesis of α-chiral ketones by the reduction of enones with baker's yeast. *Tetrahedron: Asymmetry*, **12:**3007-3013.

Khan, H., Barna, T., Bruce, N. C., Munro, A. W., Leys, D. and Scrutton, N. S. (2005). Proton transfer in the oxidative half-reaction of pentaerythritol tetranitrate reductase. Structure of the reduced enzyme-progesterone complex and the roles of residues Tyr186, His181, His184. *The FEBS Journal*, **272:**4660-4671.

Kitzing, K., Fitzpatrick, T. B., Wilken, C., Sawa, J., Bourenkov, G. P., Macherouz, P. and Clausen, T. (2005). The 1.3 A crystal structure of the flavoprotein YqjM reveals a novel class of Old Yellow Enzymes. *The Journal of biological chemistry*. **280**:27904-13.

Kohli, R. M., and Massey, V. (1998). The Oxidative half-reaction of old yellow enzyme. The role of tyrosine 196. *Journal of biological chemistry*, **273:**32763-32770.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modelling: Four approaches that performed well in CASP8. *Proteins*, **77**(Suppl):114-122.

Kurata, A., Kurihara, T., Kamachi, H. and Esaki, N. (2004). Asymmetric reduction of 2-chloroacrylic acid to (*S*)--2-chloropropionic acid by a novel reductase from *Burkholderia* sp. WS. *Tetrahedron: Asymmetry*¸**15**:2837-2839.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **277**:680-685.

Marcus, L., Hartnett, J. and Storts, D. (1996). The pGEM-T and pGEM-T Easy vector systems. *Promega Notes Magazine*, **58**:36-38.

Massey, V. and Schopfer, L. M. (1986). Reactivity of old yellow enzyme with α-NADPH and other pyridine nucleotide derivatives. *Journal of Biological Chemistry,* **261**:1215-1222.

Messiha, H. L., Munro, A. W., Bruce, N. C., Barsukov, I. and Scrutton, N. S. (2005). Reaction of morphinone reductase with 2-cyclohexen-1-one and 1-nitrocyclohexene: proton donation, ligand binding, and the role of residues Histidine 186 and Asparagine 189. *The Journal of Biological Chemistry*, **280:**10695-10709.

Müller, A., Hauer, B. and Rosche, B. (2007). Asymmetric alkene reduction by yeast old yellow enzyme and by a novel *Zymomonas mobilis* reductase. *Biotechnology and Bioengineering,* **98:**22-29.

Mueller, N. J., Stueckler, C., Hauer, B., Baudendistel, N., Housden, H., Bruce, N. C. and Faber, K. (2010). The substrate spectra of pentaerythritol tetranitrate reductase, morphinone reductase, *N*-ethylmaleimide reductase and estrogen-binding protein in the asymmetric bioreduction of activated alkenes. *Advanced Synthesis and Catalysis*, **352:**387-394.

Novy, R., Drott, D., Yaeger, K. and Mierendorf, R. (2001). Overcoming the codon bias of *E. coli* for enhanced protein expression. *inNovations*, **12**:1-3.

Oberdorfer, G., Steinkellner, G., Stueckler, C., Faber, K. and Gruber, K. (2011). Stereopreferences of old yellow enzymes: structure correlations and sequence patterns in enoate reductases. *ChemCatChem*, **3**:1562-1566.

Opperman, D. J., Piater, L. A., Van Heerden E. (2008). A novel chromate reductase from *Thermus scotoductus* SA-01 related to old yellow enzyme. *Journal of Bacteriology,* **190**:3076-3082.

Opperman, D. J., Sewell, B. T., Litthauer, D., Isupov, M. N., Littlechild, J. A. and Van Heerden, E. (2010). Crystal structure of a thermostable old yellow enzyme from *Thermus scotoductus* SA-01. *Biochemical and Biophysical Research Communications*, **393**:426-431.

Padhi, S. K., Bougioukou, D. J. and Stewart, J. D. (2009). Site-saturation mutagenesis of tryptophan 116 of *Saccharomyces pastorianus* old yellow enzyme uncovers stereocomplementary variants. *Journal of the American Chemical Society*, **131:**3271-8320.

Schittmayer, M., Glieder, A., Uhl, M. K., Winkler, A., Zach, S., Schrittweise, J. H., Kroutil, P., Macheroux, P., Gruber, K., Kambourakis, S., Rozzell, D. J. and Winkler, M. (2011). Old yellow enzyme-catalysed dehydrogenation of saturated ketones. *Advanced Synthesis and Catalysis*, **353**: 268-274.

Schmitt, J., Hess, H. and Stunnenberg, H. G. (1993). Affinity purification of histidine -tagged proteins. *Molecular Biology Reports*, **18**:223-230.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**:461-464.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I.G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon., P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlenois, R. L. Doolittle, W. F., Duquet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sense, C. W. and Van der Oost, J. (2001). The complete genome of the crenarchaeon

Sulfolobus solfataricus P2. *Proceedings of the National Academy of Sciences of the United States of America.*, **98**:7835-7840.

Shenoy, S.R. and Jayaram, B. (2010). Proteins: sequence to structure and function - current status. *Current Protein and Peptide Science*, **11**:498-514.

Shimoda, K. and Kubota, N. (2004). Asymmetric reduction of 2-substituted 2-butenolides with reductase from *Marchantia polymorpha*. *Tetrahedron: Asymmetry*, **15**:3827-3829.

Shimoda, K, Kubota, N., Hamada, H., Kaji, M. and Hirata, T. (2004). Asymmetric reduction of enones with *Synechococcus* sp. PCC 7942. *Tetrahedron: Asymmetry*, **15**:1677-1679.

Shoemaker, J. S. and Fitch, W. M. (1989). Evidence from nuclear sequences that invariable sites should be considered when sequence divergence is calculated. *Molecular Biology and Evolution*, **6**:270-289.

Simon, H., Bader, J., Günther, H., Neumann, S. and Thomas, J. (1985). Chiral compounds synthesized by biocatalytic reductions. *Angewandte Chemie International Edition in English*, **24**:539-553.

Smith, P. K., Krohn, R. I., Hermanson, G. T., Malia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, **150**:76-85.

Spiegelhauer, O., Werther, T., Mende, S., Dickert, F., Knauer, S.H., Ullman, G.M. and Dobbek, H. (2010). Cysteine as a modulator residue in the active site of xenobiotic reductase A: a structural, thermodynamical and kinetic study. *Journal of Molecular Biology*, **398**:66-82.

Stackebrandt, E., Murray, R.G. and Trüper, H. G. (1988). *Proteobacteria* classis nov., a name for the phylogenetic taxon including the "purple bacteria and their relatives". *International Journal of Systematic Bacteriology*, **38**:321-325.

Stott, K., Saito, K., Thiele, D.J. and Massey, V. (1993). Old yellow enzyme, the discovery of multiple isozymes and a family of related proteins. *Journal of Biological Chemistry*, **268**: 6097-6106.

Straßner, J., Fürholz, A., Macheroux, P., Amrhein, N. and Schaller, A. (1999). A homolog of old yellow enzyme in tomato. Spectral properties and substrate specificity of the recombinant protein. *The Journal of Biological Chemistry*, **274**:,35067-35073.

Studier, W. F. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology*, **219**:37-44.

Studier, W. F. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, **41**:207-234.

Stueckler, C., Hall, M., Ehammer, H., Pointner, E., Kroutil, W., Macheroux, P. and Faber, K. (2007). Stereocomplementary bioreduction of α,β-unsaturated dicarboxylic acids and dimethyl esters using enoate reductases: enzyme- and substrate-based stereocontrol. *Organic Letters*, **9**:5409-5411.

Stuermer, R., Hauer, B., Hall, M. and Faber, K. (2007). Asymmetric bioreduction of activated C=C bonds using enoate reductases from the old yellow enzyme family. *Current Opinion in Chemical Biology*, **11**:203-213.

Swiderska, M.A. and Stewart, J.D. (2006). Stereoselective enone reductions by saccharomyces carlsbergensis old yellow enzyme. *Journal of Molecular Catalysis B: Enzymatic*, **42**:52-54.

Taglieber, A., Schulz, F., Hollmann, F., Rusek, M. and Reetz, M. T. (2008). Light-driven biocatalytic oxidation and reduction reactions: scope and limitations. *Chembiochem*, **9**:565-572.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**:2731-2739.

Tang, Y. and Suga, T. (1992). Characterisation of two enone reductases from *Nicotiana tabacum* cultures. *Phytochemistry,* **31**:2599-2603.

Toogood, H .S., Fryszkowska, A., Hare, V., Fisher, K., Roujeinikova, A., Leys, D., Gardiner, J. M., Stephens, G .M. and Scrutton, N. S. (2008). Structure-Based Insight into the Asymmetric Bioreduction of the C=C Double Bond of α,β-unsaturated nitroalkenes by pentaerythritol tetranitrate reductase. *Advanced Synthesis and Catalysis*, **350:**2789-2803.

Toogood, H. S., Gardiner, J. M. and Scrutton, N. S. (2010). Biocatalytic reductions and chemical versatility of the old yellow enzyme family of flavoprotein oxidoreductases. *ChemCatChem*, **2**:892-914.

Trotter, E. W., Collinson, E. J., Dawes, I. W. and Grant, C. M. (2006). Old yellow enzymes protect against acrolein toxicity in the yeast *Saccharomyces cerevisiae. Applied and environmental microbiology*, **72**:4885-4892.

Uchida, K., Kanematsu, M., Morimitsu, Y., Osawa, T., Noguchi, N. and Niki, E. (1998). Acrolein is a product of lipid peroxidation reaction. *Journal of Biological Chemistry*, **273**:16058-16066.

Utaka, M., Konishi, S., Okubo, T., Tsuboi, S. and Takeda, A. (1987). A faciel synthesis of optically pure L-armentomycin and its D-isomer. Highly enantioselective reduction of the C=C double bond of methyl (*E*)- and (*Z*)-2,4,4-trichloro-2-butenoate by using baker's yeast. *Tetrahedron Letters*, **28**: 1447-1449.

Vandamme, P. and Coenye, T. (2004). Taxonomy of the genus Cupriavidus: a tale of lost and found. *Interantional Journal of Systematic and Evolutionary Microbiology*, **54**:2285-2289.

van den Hemel, D., Brige, A., Savvides, S. N. and Van Beeumen, J. (2006). Ligand-induced conformational changes in the capping subdomain of a bacterial old yellow enzyme homologue and conserved sequence fingerprints provide new insights into substrate binding. *Journal of Biological Chemistry*, **281**:28152-28161.

Vaz, A. D. N., Chakraborty, S. and Massey, V. (1995). Old yellow enzyme: aromatization of cyclic enones and the mechanism of a novel dismutation reaction. *Biochemistry*, **34:** 4246-4256.

Wada, M., Yoshizumi, A., Noda, Y., Kataoka, M., Shimizu, S., Takagi, H., Nakamori, S. (2003). Production of a doubly chiral compound, (4*R*,6*R*)-4-hydroxy-2,2,6-trimethylcyclohexanone, by two-step enzymatic asymmetric reduction. *Applied and Environmental Microbiology*, **69**:933-937.

Warburg, O. and Christian, W. (1932). Uber das neue Oxydationsferment. *Die Naturwissenschaften*, **20**:980-981.

Wasternack, C. and Kombrink, E. (2010). Jasmonites: structural requirements for lipid-derived signals active in plant stress responses and development. *Chemical biology*, **5**:63-77.

Wheeler, D. and Bhagwat, M. (2007). BLAST QuickStart: Example-driven web-based BLAST Tutorial. In: Comparative genomics: volumes 1 and 2 (Bergman, N.H., ed.), Chapter 9. Humana Press.

Whelan, S. and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution*, **18**:691-699.

Williams, R. E. and Bruce, N. C. (2002). "New uses for an old enzyme" – the old yellow enzyme family of flavoenzymes. *Microbiology*, **148**:1607-1614.

Williams, R. E., Rathbone, D. A., Scrutton, N. S. and Bruce, N. C. (2004). Biotransformation of explosives by the old yellow enzyme family of flavoproteins. *Applied and Environmental Microbiology*, **70**:3566-3574.

Winter, G. (2010). Xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography*, **43**:186-190.

Wu, S. and Zhang, Y. (2008). MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. *Proteins,* **72**:547-556.

Wu, J. T., Wu, L. H. and Knight, J. A. (1986). Stability of NADPH: Effect of various factors on the kinetics of degradation. *Clinical chemistry*, **32**:314-319.

Xu, D., Kohli, R. M. and Massey, V. (1999).The role of threonine 37 in flavin reactivity of the old yellow enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, **96**:3556-3561.

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, **39**:306-314.