Engineering of the PETNR active site to accommodate novel α/β substituted enone substrates.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Life Sciences

2010

Martyn Edward Hulley

Contents

Chapter 1: Introduction

Chapter 2: Materials & Methods

Chapter 3: Results 1

Chapter 4: Results 2

Chapter 5: Results 3

Chapter 6: Results 4

Chapter 7: Discussion

57,335 words in total

Abstract

Engineering of the PETNR active site to accommodate novel α/β substituted enone substrates.

Experiments facilitating the engineering of the PETNR active site to accommodate a range of non natural enone substrates with substituents localised on the α and β carbons of the unsaturated bond are described. In order to facilitate the high throughput purification of PETNR libraries poly histidine ($PETNR_{His}$) and biotin ($PETNR_{Bio}$) tagged PETNR variants were generated. High throughput protocols were developed for the automated generation, purification and screening of libraries in a 96 well format. Protocols were optimised and trialled using blocks consisting of $PETNR_{His}$ WT only and characterised in terms of intra block variation. A range of single site saturation mutagenic libraries were generated at positions in the active site consisting of T26, Y68, W102, H181, H184, Y186, Q241 and Y351. Sequencing results indicated randomised libraries with the occasional instance of bias evident. Expression and purification in a 96 well format was monitored by SDS PAGE and protein quantitation. Library activity was quantified and demonstrated to retain varying degrees of activity with the model substrate 2-cyclohexenone. Following this verification of the experimental protocol libraries were screened against a range of substrates analogous to substrates demonstrated to be active with $PETNR_{WT}$ but incorporating substituents at the α and β carbons. 'Hits' generated from these screening reactions were studied further by the determination of the specific activity and quantitation of substrate / product from biotransformation reactions. From these screening experiments totalling 3,600 individual reactions, 35 were identified as potential hits, of these 8 proved to be genuinely improved variants. Substituents at the β carbon were demonstrated to compromise the activity of the WT enzyme most severely. Positions 68, 102, and 351 were demonstrated to play an important role in the accommodation of substituents at the α carbon whilst residues 26 and 351 are important for the β carbon. The best variants demonstrated up to 9 fold improvements in poor substrates which represented rates in excess of those observed for model substrates.

The University of Manchester

Martyn Edward Hulley

Doctor of Philosophy

14 September 2010

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to thank my supervisor Professor Nigel Scrutton for affording me the opportunity to undertake this project, and Dr David Leys for acting as my advisor for the duration of this project.

I would also like to thank Dr Helen Toogood for instruction in all aspects of molecular biology and sound advice on all aspects of my project, Dr Anna Fryszkowska for guidance regarding the chemistry based aspects of my project, and all current and ex members of the Scrutton/Munro/Leys groups who have helped at various stages during the course of my studies.

On a more personal note in particular I would like to thank Bjorn Adalbjornsson, James Law and Mark Dunstan for making the experience more enjoyable.

An enormous amount of thanks go to my ever dependable parents Graham and Susan who been unstinting in their support and encouragement through all my endeavours for the last 27 years, and no doubt will continue to be into the future.

Finally I would like to thank my long suffering wife, Kelly, and our daughter, Megan, who have given me limitless encouragement, support and understanding.

Dedication

This thesis is dedicated firstly to my wonderful wife Kelly, who for the last 7 years has been a source of strength and encouragement. Also to my daughter, Megan, who has reminded me during the hard times of my motivation and helped me to maintain a sense of perspective.

Abbreviations

ADNT – Aminodinitrotoluene

ADP – Adenosine diphosphate

AIM – Auto induction media

AP – Antarctic phosphatase

CASTing – Combinatorial Active Site Saturation Testing

CH – 2-cyclohexen-1-one

CrS – Chromate reductase

DMSO – Dimethyl sulfoxide

DNP – 2,4-Dinitrophenol

dNTP – Deoxyribonucleotide triphosphate

DTT – Dithiothreitol

DYT – Double yeast tryptone

EDGN - Ethylene glycol dinitrate

EDTA – Ethylenediaminetetraacetic acid

EIMS – Electrospray Ionisation Mass Spectroscopy

EPPCR – Error Prone PCR

EtBr – Ethidium bromide

EWG – Electron withdrawing group

FACS – Fluorescence activated cell sorting

FID – Flame Ionization Detector

FMN – Flavin Mononucleotide

GC – Gas chromatography

GFP – Green fluorescent protein

GTN – Glycerol Trinitrate

GTNR – Glycerol Trinitrate Reductase

H^-TNT – Hydride Meisenheimer complex

HADNT – Hydroxylaminodinotritoluene

HBA - *p*-hydroxybenzaldehyde

HFPCR – high fidelity PCR

HPLC – High Performance Liquid Chromatography

ICASTing – Iterative combinatorial Active Site Saturation Testing

IPTG – Isopropyl β-D-1-thiogalactopyranoside

ISM – Iterative Saturation Mutagenesis

LB – Luria Bertani

MES – (n-morpholino) ethane sulfonic acid

MMR – Multiple mutation reaction

MR – Morphinone Reductase

NADH – reduced Nicotinamide Adenine Dinucleotide

NADPH – reduced Nicotinamide Adenine Dinucleotide Phosphate

NEMR – N-Ethylmaleimide reductase

Ni - NTA – Nickel-nitrilotriacetic acid

OFN – Oxygen Free Nitrogen

Onr – Organic nitrate reductase

OPDR – 12-Oxophytodienoate Reductase

OYE – Old Yellow Enzyme

PA – Picric Acid

PBS – Phosphate buffered saline

PCD – Programmed cell death

PCR – Polymerase chain reaction

PETN – Pentaerythritol Tetranitrate

PETNR – Pentaerythritol Tetranitrate Reductase

$PETNR_{Bio}$ – Biotin Tagged PETNR

$PETNR_{His}$ – His tagged PETNR

$PETNR_{WT}$ – Wild type PETNR

PNK – T4 polynucleotide kinase

ROS – Reactive oxygen species

SDS – Sodium dodecyl sulphate

SDS PAGE – Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SOC – Super optimal broth with catabolite repression

SSSM – Single site saturation mutagenesis

StEP – Staggered extension process

TAE – TRIS-acetate-EDTA

Taq – *Thermus aquaticus*

TB – Terrific broth

TFB – Transformation buffer

TOYE – Thermophillic OYE

TNT – 2,4,6 Trinitrotoluene

WT – Wild type

XR – Xenobiotic Reductase

1 Introduction

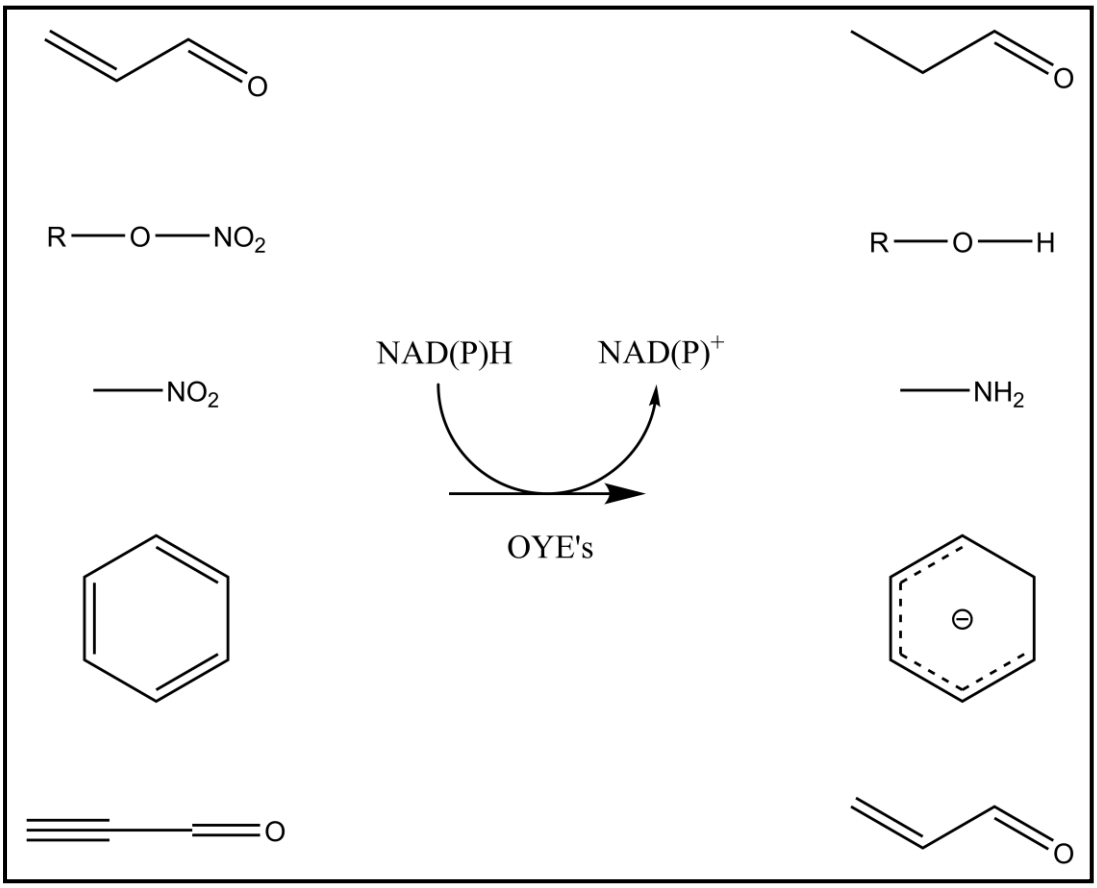
1.1 Overview

The aim of the work described in this thesis is to develop a robotic high throughput method for the generation of tailored biocatalysts, specifically to enable an extension in the substrate specificity of the flavoprotein oxidoreductase pentaerythritol tetranitrate reductase (PETNR).

This chapter will discuss the known properties of the Old Yellow Enzyme (OYE) family of oxidoreductases to which PETNR belongs. This discussion will focus on the; i) diversity of OYE family members, including a brief history of the family, ii) mechanistic similarities within the family, iii) properties of PETNR specifically, iv) reaction mechanisms of OYEs with a number of substrate classes, v) structure-function relationships within the family, to enable the identification of suitable targets for mutagenesis, vi) the experimental techniques used to generate and screen libraries to isolate improved biocatalysts.

1.2 OYE family overview

The OYE family of flavoproteins are NAD(P)H (reduced Nicotinamide Adenine Dinucleotide (Phosphate)) - dependent enzymes [1], a range of which are detailed briefly in Table 1.1. OYEs have been shown to reduce a variety of oxidative substrates including: α/β unsaturated carbonyls, e.g. 2-cyclohexen-1-one (CH) [2]; nitrate esters, e.g. pentaerythritol tetranitrate (PETN) [3]; nitro compounds, e.g. 2,4,6-trinitrotoluene (TNT) [4]; aromatic rings, e.g. also in TNT [5] and alkynes e.g. 4-phenyl-3-butyne-2-one [6] (Scheme 1.1). In terms of size, OYEs are capable of accommodating a wide range of substrates from acrolein at the bottom end of the spectrum to multi ring steroid systems at the other extreme [2].



Scheme 1.1. Schematic representation of the generalised reactions of OYE homologues. Compounds represent generalised substrates.

1.2.1 Historical perspective

The first OYE (OYE1) to be described was purified from *Saccharomyces carlsbergensis (S. carlsbergensis)* in 1932 by Warburg and Christian and initially described as 'gelbe ferment' [7]. The name derives from the yellow colouration present in the enzyme due to the non-covalently bound flavin mononucleotide (FMN) group [8]. The 'Old' in OYE was added to differentiate it from another yellow enzyme purified later [9].

OYE has played a significant role in the development of enzymology, providing the first example of a cofactor dependent enzyme directly involved in the catalysis of a redox reaction [1], work for which Hugo Theorell was awarded the 1955 Nobel prize for outstanding contribution in Physiology or Medicine [10].

1.2.2 Family members

While OYE1 is the best studied member of this family, a large number of homologous enzymes have been discovered and studied. The original old yellow enzyme (OYE1) from *Saccharomyces carlsbergensis*, was isolated from the native organism [11], which was found to consist of a mixture of 5 isozymes (termed OYE I-V). These enzymes all have a similar molecular mass (~50 kDa), indistinguishable by Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS PAGE), and identical spectral properties, however differ in their oligomeric state and contain subtle variations in the shape of the active site as demonstrated by ^{13}C NMR [12]. This OYE preparation was shown to be reduced by NAD(P)H [13]. The structure of recombinant OYE1 [14] was determined to 2 Å in oxidised and reduced forms, and with *p*-hydroxybenzaldehyde (HBA), β-estradiol and α-$O^{2'}$-6B-cyclo-1,4,5,6-tetrahydro-nicotinamide adenine dinucleotide phosphate (an NADPH analogue) as bound ligands [15]. Following the availability of a homogenous enzyme preparation the catalytic mechanism has been studied in great detail, particularly the oxidative half reaction, and will be discussed further in Section 1.4.

Two OYE enzymes from *Saccharomyces cerevisiae*, OYE2 – 3 respectively, were characterised demonstrating 92% and 80% identity to OYE1 [16, 17]. One focus of

research into these enzymes has concentrated on the identification of the physiological function of OYE homologues. Despite years of research a definitive answer to the conundrum of the natural function(s) and physiological substrate(s) of this class of enzyme remains elusive (Section 1.2.6).

Morphinone reductase (MR) is an NADH-dependent OYE homologue that catalyses the reduction of bulky α/β unsaturated carbonyl compounds, such as morphinone [18]. MR was initially isolated from a *Pseudomonas putida* strain due to its ability to utilise morphinone and codeinone as a sole carbon source [19]. The structure of MR has been solved [20], and this enzyme has become the focus of intense study to determine the mechanism for the OYE reductive half reaction.

Due to the rapid increase in available sequence data, attributable to the advent of rapid genome sequencing, a wealth of OYE like proteins have now been discovered. The majority are categorised as putative OYE based purely on sequence homology, but a significant number have been studied in rather more detail. These are listed and referenced in table 1.1.

Strain	Enzyme	Ref
S. carlsbergensis	Old yellow enzyme 1 (OYE 1)	[11]
S. cerevisiae	Old yellow enzyme 2 (OYE 2) Old yellow enzyme 3 (OYE 3)	[16] [17]
P. putida M10	Morphinone reductase (MR)	[18]
Enterobacter cloacae PB2	Pentaerythritol tetranitrate reductase (PETNR)	[21]
P. putida KT2440	Xenobiotic reductase (XR A – F)	[22]
Agrobacterium radiobacter	Glycerol trinitrate reductase (GTNR)	[23]
Arabidopsis thaliana	12-Oxophytodienoate reductase (OPR 1 – 2)	[24]
Lycopersicon esculentum	Oxophytodienoate reductase 1-3	[25]
P. putida II-B	Xenobiotic Reductase A (XR A)	[26]
Pseudomonas fluorescens I-C	Xenobiotic Reductase B (XR B)	[26]
Bacillus subtillis	YqjM	[27]
Thermoanaerobacter pseudethanolicus E39	Thermophillic OYE (TOYE)	[28]
Thermus scotoductus SA-01	Chromate reductase (CrS)	[29]
Escherichia coli	N-Ethylmaleimide reductase (NEMR)	[30]
Aspergillus fumigatus	OYE (AFOYE)	[31]
Candida macedoniensis	OYE (CMOYE)	[32]
Gluconobacter suboxydans	OYE (GSOYE)	[33]
Hansenula polymorpha	OYE (HPOYE)	[34]
Kluyveromyces lactis	OYE (KLOYE)	[35]
Trypanosoma cruzi	OYE (TCOYE)	[36]
Shewanella oneidensis	OYE (SOOYE)	[37]
P. putida	OYE (PPOYE)	[22]

Table 1.1 Summary of well characterised OYE homologues. A large number of additional putative OYEs have been identified from genome sequencing experiments, these are not included for reasons of clarity.

1.2.3 PETNR historical overview

PETNR is a NAD(P)H dependent OYE that catalyses the reduction of α/β unsaturated activated alkenes e.g. 2-cyclohexen-1-one (CH) [38], aromatic and aliphatic nitro-containing compounds e.g. PETN [3], and aromatic ring systems e.g. TNT [5].

PETNR was derived from a strain of *Enterobacter cloacae* strain PB2, originally isolated from explosives-contaminated soil, selected for its ability to utilize PETN as a sole nitrogen source [3]. The gene was originally named organic nitrate reductase (*onr*), though the enzyme is commonly known as PETNR [39]. Following initial identification and characterisation, structural and sequence homology between PETNR and the OYE (Old Yellow Enzyme) family of flavoenzymes enabled the assignment of PETNR enzyme to this family [7].

1.2.4 Generalised familial structure

Although there are structural differences between members of the OYE family of proteins, most markedly their variable oligomeric states, there are also common structural elements which appear in OYE homologues.

The overall monomeric structure of the OYE family is based on an 8 stranded α/β barrel fold [15, 40-43]. The substrate-binding channel is relatively large and can accommodate the binding of both small and bulky substrates such as acrolein and multi-ring steroid structures [2]. Substrate specificity is governed by the residues lining the active site channel, leading from the FMN to the surface of the protein. These residues are largely hydrophobic, with loop β3, loop β5 and loop β6 being highly variable [15]. An overview of the structure of OYE is shown in Figure 1.1.

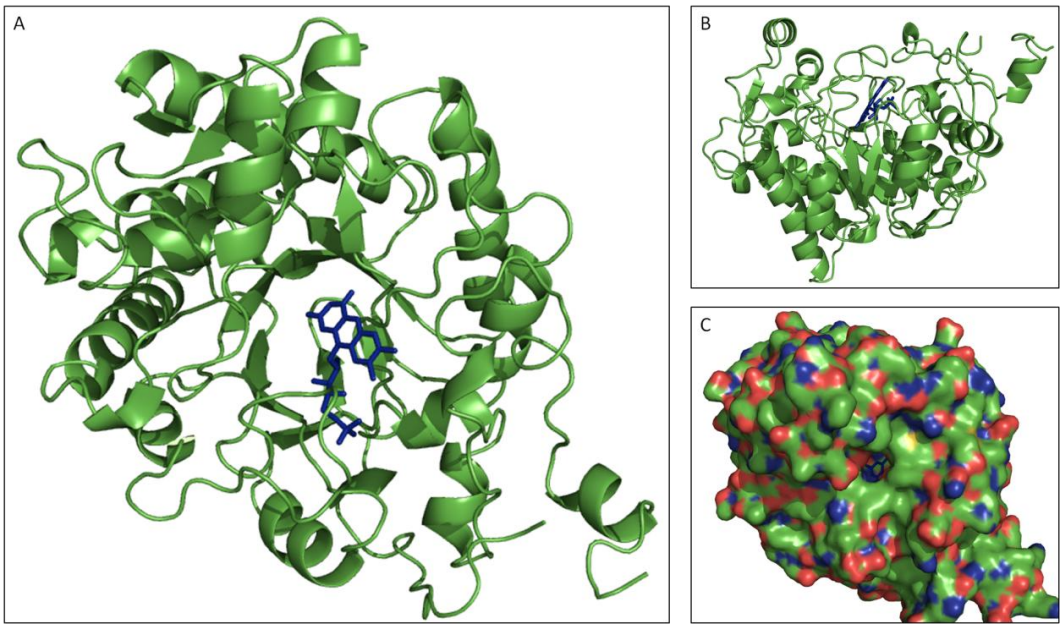


Figure 1.1. Overall structure of OYE; Panel A, the overall 8 stranded α/β barrel fold of OYE1 in cartoon format approximately along the axis of the barrel; Panel B, the same structure viewed from perpendicular to the barrel axis; Panel C, space fill model showing the substrate binding channel of OYE1 from 1OYA [15].

A number of OYEs exist and function as monomers such as PETNR [3] and glycerol trinitrate reductase (GTNR) [23]. There is, however, great variety throughout the family, for example OYE1 exists as a dimer [15] with a highly conserved dimer interface. This interface is highly conserved throughout the OYE family (even amongst non-dimeric proteins). In the case of MR which also exists as a dimer the dimer interface differs. In the case of 12-oxophytodienoate reductase 1 (OPDR) dimerisation has been shown to be of functional significance in self inhibition [42]. A more recent isolate YqjM from *B. subtilis* has been shown to exist as a tetramer with oligomer formation being important for function due to shared active site residues. In this unusual structure, R336 on a C terminal loop from one YqjM monomer contributes to the active site formation of the second monomer and is described as an 'arginine finger' [43]. This interaction along with the dimerisation interface can be seen in Figure 1.2.

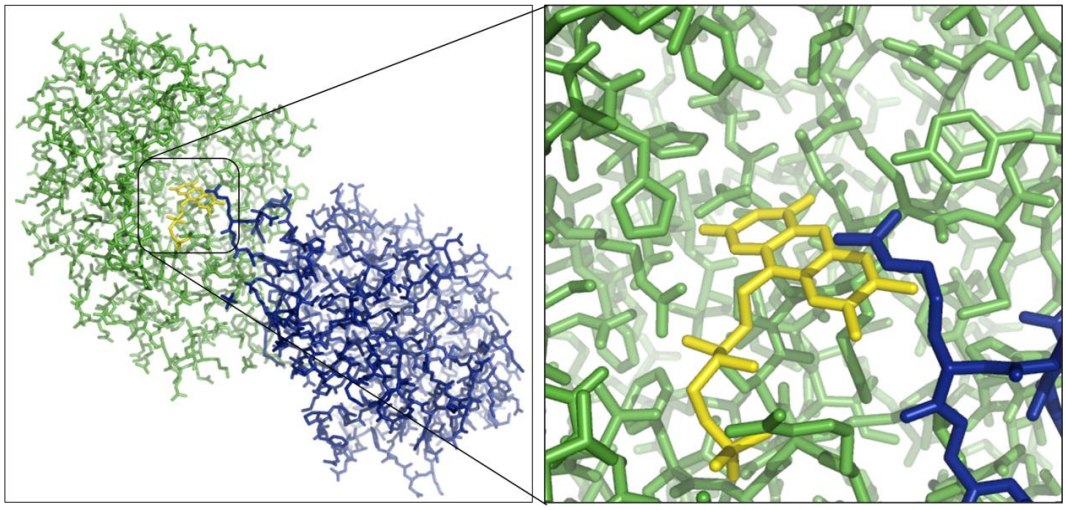


Figure 1.2. Dimerisation interface of YqjM. Inset shows the interaction of R336 with the active site of the opposite monomer. Green, monomer 1; blue, monomer 2; yellow, FMN. From 1Z41 [43].

Importantly, a distinction can be drawn here between two structurally variable subclasses. One subclass contains traditional OYE like proteins including various OYE's, PETNR, MR, etc. The other subclass is composed of largely thermophillic OYE homologues including YqjM, Thermophillic OYE (TOYE), Chromate Reductase (CrS), etc... This thermophile-like subclass is characterised by a highly conserved C terminal region incorporating an arginine finger motif which binds in the active site of the partner protein of the dimeric unit and is essential for function [28]. An alignment of both enzyme classes can be seen in Figure 1.3 highlighting key residues.

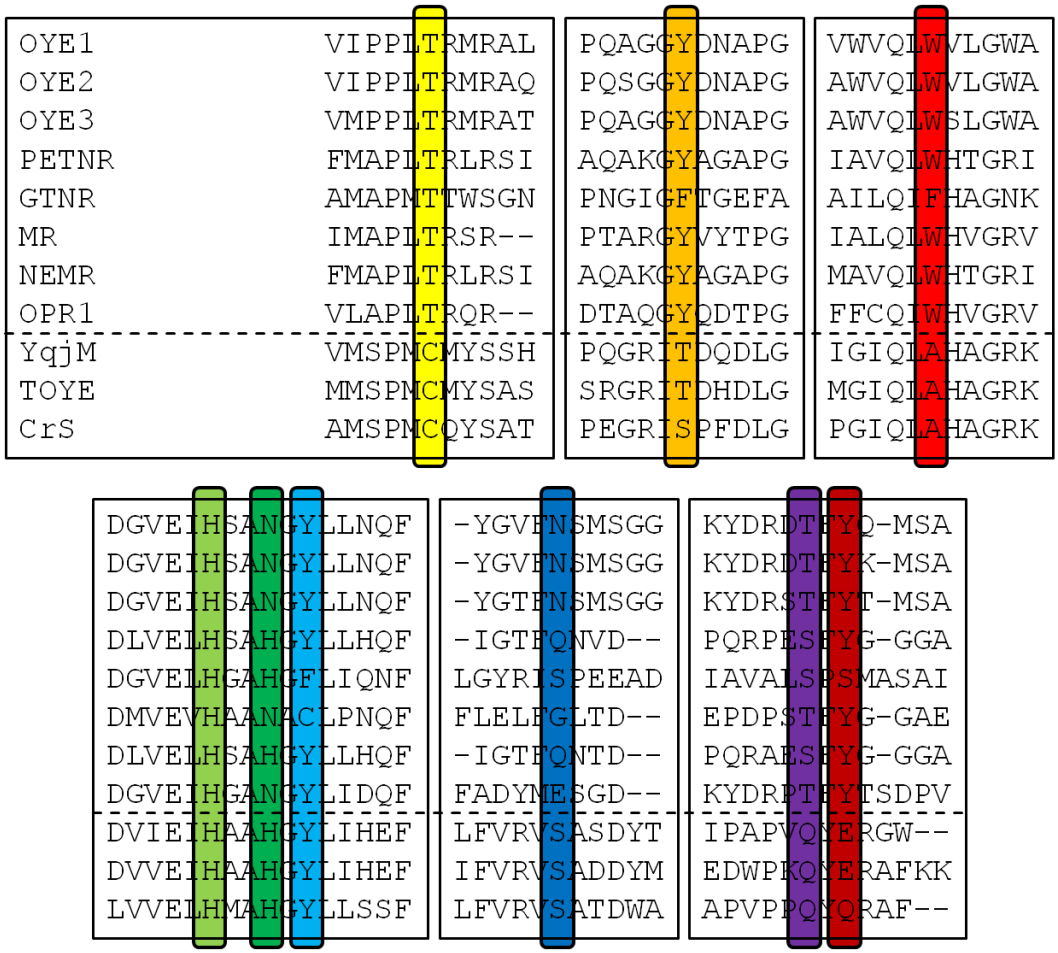


Figure 1.3 Alignment of a representative sample of OYE homologues highlighting key residues as discussed in the text below (PETNR numbering unless stated otherwise). Classical OYEs are above the dividing line thermophile like OYE are below. Yellow, residue 26; orange, residue 68; red, residue 102; light green, residue 181; dark green, residue 184; cyan, residue 186; dark blue, residue 241; purple, residue 333 (TOYE numbering); brown, residue 351.

1.2.5 General substrate profile

The OYE family of enzymes have a number of conserved activities throughout the family, including the reduction of unsaturated carbonyl compounds (enones) [2]. This reaction proceeds by the selective reduction of the unsaturated carbon double bond without alteration of the carbonyl group. A simple member of this substrate class, CH, has been recognised as a good general oxidative substrate for many members of this family [2].

Several of the OYE's, including PETNR, have been demonstrated to be able to perform the reductive denitration of nitrate esters and nitramines, particularly with

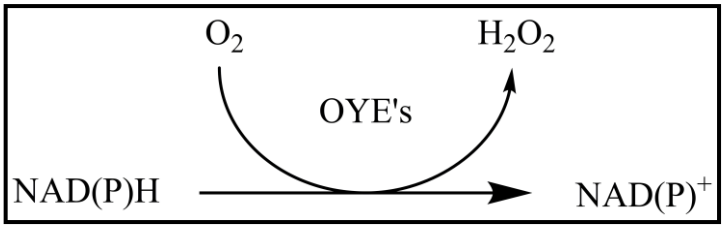
respect to the nitrate ester explosives, such as the reaction of PETNR with PETN and Glycerol Trinitrate (GTN), resulting in the respective alcohol product [44]. The direct reduction of nitro groups to give the respective amines proceeding through a hydroxyl amino intermediate has been observed in the reaction of PETNR with TNT [4]. OYE's have been demonstrated to also be capable of the reduction of aromaticised unsaturated bonds, even in the presence of nitro groups exemplified again by the reaction of PETNR with TNT [5]. An outline of all these reactions has been shown in Scheme 1.1.

1.2.6 Physiological role

Despite study of the old yellow enzyme family beginning almost 80 years ago the physiological role of the enzyme family is still a matter of contention. A number of theories as to what role the OYE homologues play in different organisms have been put forward including the recycling of cellular cofactors [33, 35], the biosynthesis of a range of compounds [16, 31, 45], the detoxification of toxic compounds [34, 36, 46], and the reaction to oxidative stress [27, 47, 48].

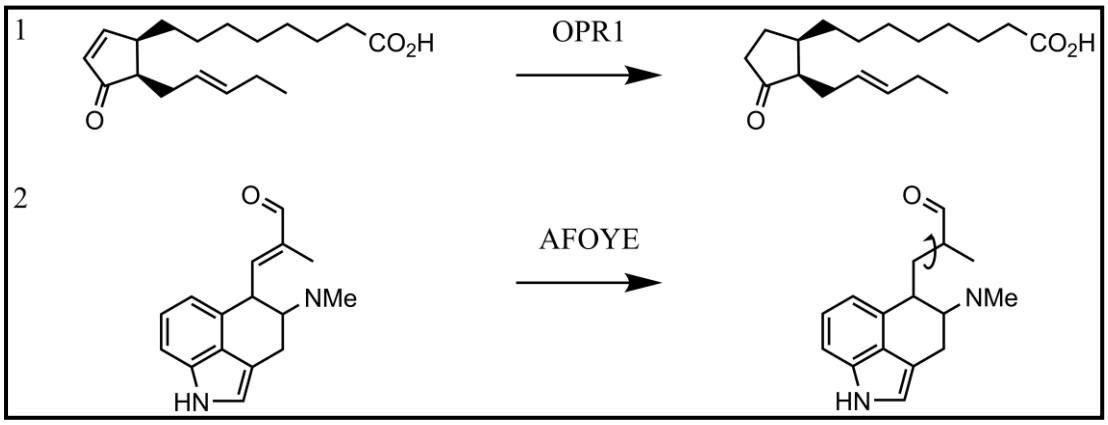
There is a high level of conservation throughout the OYE family, with OYE homologues being present in all kingdoms of life except higher eukaryotes, such as animals, and the presence of multiple homologues in a number of cases [12, 16, 22, 34, 37]. This indicates to the possibility of important natural function(s).

Early reports suggest that the ability to transfer electrons to molecular oxygen may be a facet of the natural role of OYE in NADPH recycling. It is suggested that in *G. suboxydans* $NAD(P)^+$ is recycled by OYE at the expense of oxygen allowing turnover of the pentose phosphate pathway to provide glycolytic metabolites (Scheme 1.2) [33]. A similar effect has also been observed in *K. lactis* deficient OYE where a decreased rate of NADPH oxidation was reported to be responsible for a substantial increase in division time of the organism [35]. Such a role in *S. carlsbergensis* is thought unlikely due to slow turnover of OYE1 with molecular oxygen [11].



Scheme 1.2. $NADP^+$ recycling Scheme showing the turnover of NAD(P)H by OYE using molecular oxygen as the oxidative substrate.

One well studied physiological role is in the biosynthesis of jasmonic acid, which is involved in plant signalling in the oxidative stress response [45]. The OYE homologue 12-oxophytodienoate reductase (OPR) has been demonstrated to reduce 9S,13S-12-oxophytodienoate to 1S,2S-3-oxo-2(29[Z]-pentenyl)-cyclo-pentane-1-octanoate (Scheme 1.3, panel 1) [45]. The most recent studies on *A. fumigatus* OYE suggest a role in the synthesis of ergot alkaloids in fungi [31]. The gene for the OYE homologue is located within the ergot alkaloid biosynthesis gene cluster and the purified protein has been demonstrated to catalyse the reaction of chanoclavine-1-aldehyde to dihydrochanoclavinealdehyde allowing free rotation prompting a ring closure reaction (Scheme 1.3, panel 2) [31]. Another biosynthetic pathway potentially relying on OYE 2 is the synthesis of the fungal membrane sterol ergosterol [16].



Scheme 1.3. Biosynthesis reactions of various OYE homologues. Panel 1, the reaction of OPR1; panel 2, reaction of AFOYE.

In *H. polymorpha*, an OYE has been demonstrated to have a role in detoxification of acrolein to propionaldehyde generated by alcohol oxidase metabolism of allyl alcohol. This observation lead to the proposal that OYE could be involved in the

detoxification of a range of toxic α/β unsaturated carbonyl species [34]. In *T. cruzi*, the causative agent of Chagas' disease, an OYE homologue has been demonstrated to metabolize a whole range of compounds used as drugs to treat the disease. In the case of naphthoquinone drugs, such as menadione, one electron reduction by TCOYE forms the naphthoquinone radical [36] which reacts with molecular oxygen to give the superoxide anion leading to cell death, which is thought to be the mode of action for this drug [49]. In the case of nitroheterocyclic drugs, such as nifurtimox, these have been demonstrated to undergo 2 electron reduction leading to detoxification [36]. A number of over-expressed proteins were identified from *A. thaliana* growing in TNT contaminated media. These proteins were classified as oxophytodienoate reductases (OPRs), a subclass of OYE's. A number of these proteins were observed to be over expressed in the cytosol and peroxisome of root cells in response to TNT exposure providing a detoxification function allowing the plant to grow [46].

OYE's have been shown to have a role on the progression of BAX-induced reactive oxygen species-mediated programmed cell death (PCD) in an artificial yeast system. OYE2 has been shown to have a protective effect against PCD. Conversely where OYE2 : OYE3 heterodimers are formed these serve to further sensitise the cell to the PCD mechanism [47]. Furthermore it is suggested that OYE3 is involved in the mechanism by which BAX induces reactive oxygen species (ROS) mediated PCD including the metabolism of lipid peroxidation products [48]. A more generalized role in the reaction to externally induced oxidative stress has been demonstrated for the OYE homologue YqjM [27].

It is unknown whether the repeatedly observed reactivity with α/β unsaturated activated alkenes is representative of the natural function or distinct from it, some of the reactions catalysed could represent opportunistic functions of the enzyme. This suggests that whilst maintaining a role in the organism's response to oxidative stress, many OYEs have been opportunistically utilised for a wide range of functions [37], e.g. the degradation of nitrate ester explosives by PETNR [3] and morphinone derivatives by MR [18] could represent recent examples of this phenomenon.

1.2.7 Biocatalytic functions and other potential uses

With the multitude of reaction mechanisms available from OYE homologues, it is not surprising that this family of enzymes have already attracted the attention of chemists seeking stereo and enantio selective biocatalysts.

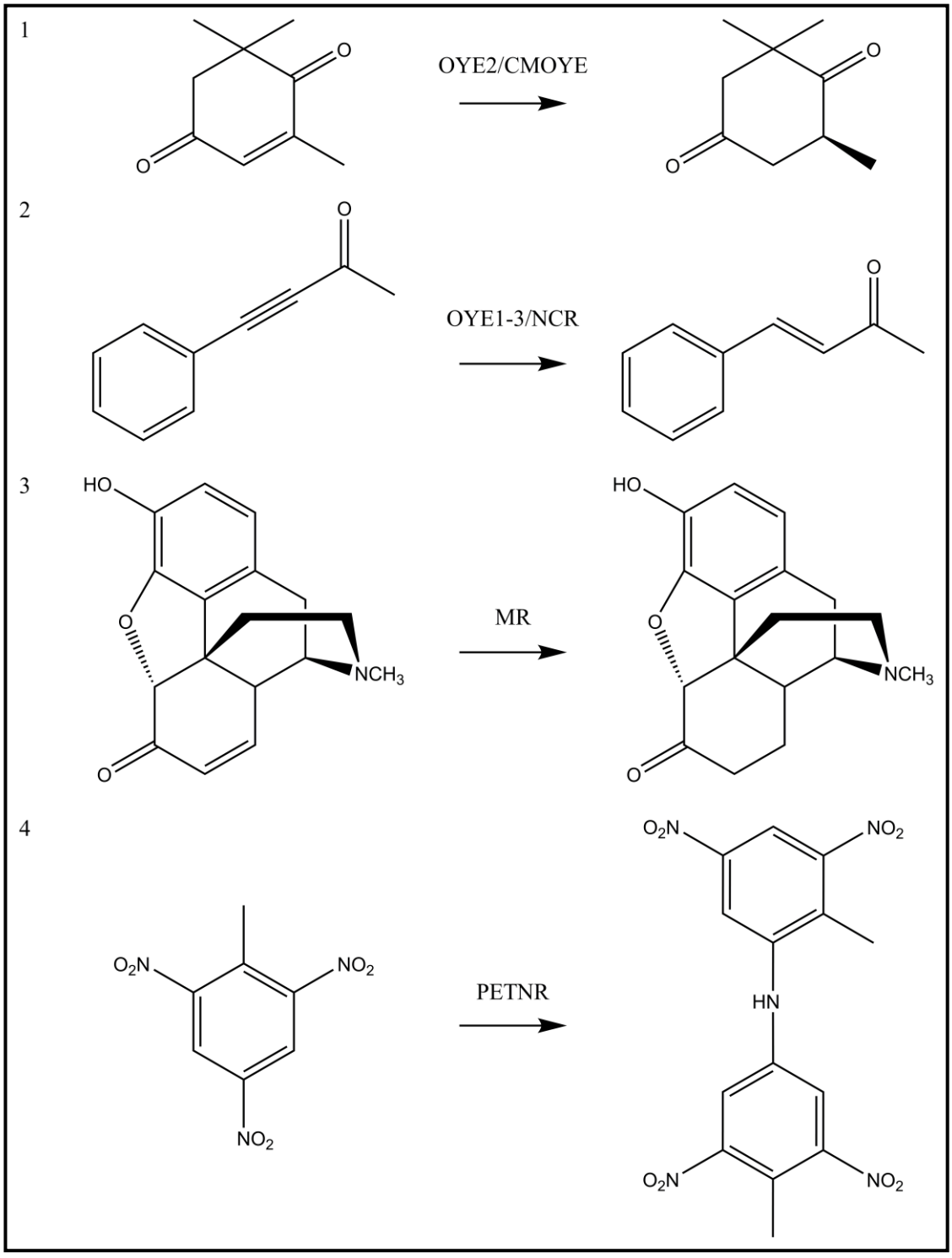
OYE homologues could potentially be used for the enantiospecific reduction of alkenes to give industrially useful products [50]. Reactions with alkenes conjugated to other electron withdrawing groups (EWG; required to induce δ+ charge on the β carbon, thus encouraging hydride attack) have also been studied including aldehyde [51], carboxylic acid, ester, nitro [52, 53], di-acids, di-esters [54].

An excellent example of an OYE being used in synthesis is the use of *S. cerevisiae* or *C. macedonis* OYE in the chiral synthesis of (6*R*)-levodione. *E. coli* cells expressing a cloned OYE construct were used to catalyse the whole cell stereoselective reduction of 2,6,6-trimethyl-2-cyclohexen-1,4-dione (ketoisophorone) to (6*R*)-2,2,6-trimethylcyclohexane-1,4-dione ((6*R*)-levodione). The reaction yielded 96.4 % product with a 95.5 % enantiomeric excess (e.e.) using OYE2 [55] while the *C. macedonis* enzyme yielded 98.7 % [32] (Scheme 1.4, panel 1).

A further reaction of OYE of particular interest to synthetic chemists is the recently identified regioselective conversion of alkyneones (ynones) to the corresponding enone, which could provide a new route for the synthesis of *trans*-enones e.g. the reaction of 4-phenyl-3-butyne-2-one (Scheme 1.4, panel 2). As is to be expected however the enone can also be further reduced to the corresponding saturated carbonyl, and synthetic utility will depend on the ability to prevent this further reduction [6].

The use of MR has been demonstrated for the synthesis of synthetic opiate compounds, including the reaction of morphinone for therapeutic use (Scheme 1.4, panel 3) [56]. Similar alkaloid compounds, including the reaction of chanoclavine - I aldehyde, have been derived from reactions with *A. fumigatus* OYE (Scheme 1.3, panel 2) [31].

The reaction of TNT with a small subset of OYE homologues, including PETNR, has been demonstrated to give diarylamine products, including secondary diarylamines (Scheme 1.4, panel 4) [57].



Scheme 1.4. Biocatalytic uses of OYE homologues

Other exploitations of OYE properties include the study of oxidative stress responses. Green fluorescent protein (GFP) fusions of OYE2-3 from *S. cerevisiae* were used to report the level of activity of the oxidative response in studies intended to demonstrate the degree of protection afforded by small molecule antioxidants [58]. The therapeutic activity of komaroviquinone against *T. cruzi* is reported to be attributable to the generation of ROS by OYE. It is also hypothesized that komaroviquinone can be used as a starting point for the discovery of novel therapeutics with a similar mode of action [49]. The use of PETNR for the bioremediation of explosive contaminated land was suggested early on in the study of this enzyme [5]. The over expression and bioremediation of solid media by oxophytodienoate reductase in *A. thaliana* has also been demonstrated [46].

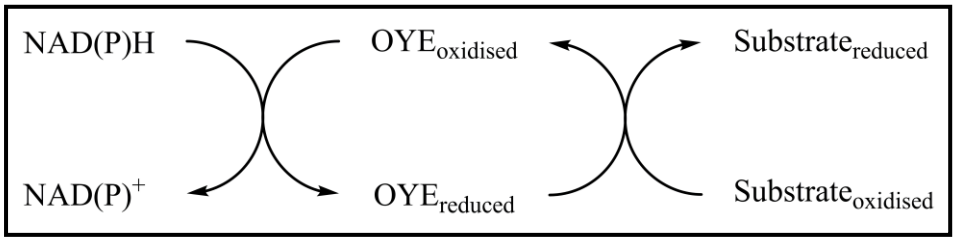
1.2.8 Interim summary

This section has described a number of important features of the OYE family of flavoproteins. Discussion of a number of OYE homologues and a brief historical account of the most prominent members helps provide context to the following more detailed chapters. Described above are the overall structural similarities with all OYEs being based an 8 stranded α/β barrel fold. However distinctions are drawn between classical OYEs and thermostable-like OYEs by the observation of an arginine finger motif.

A range of activities are common within the OYE family including the reduction of activated alkenes and alkynes, the reductive denitration of a range of compounds, the reduction of nitro groups and the reduction of aromatic rings. A number of OYEs have been demonstrated as useful for biocatalytic synthesis reactions and a wealth of other such reactions are potentially plausible. Despite decades of study, consensus regarding the physiological function of these enzymes remains elusive. However, a range of substrates with physiological relevance has been identified and studied.

1.3 Reactions and substrates of OYE homologues (specifically PETNR)

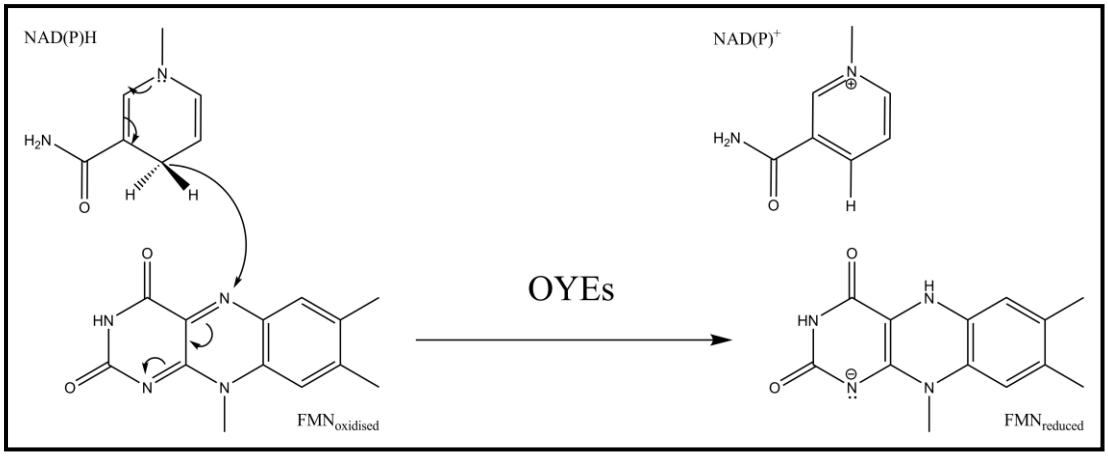
The catalytic cycles of oxidoreductase enzymes are generally divided up into two half-reactions. First, the reductive half-reaction is where oxidized enzyme reacts with an electron donor to generate the reduced, catalytically active form of enzyme. Second, in the oxidative half-reaction a second substrate is reduced and the enzyme is reoxidised. For convenience in describing the mechanistic aspects of the OYE homologues, each half-reaction will be discussed separately (Scheme 1.5).



Scheme 1.5. Generalised catalytic cycle of OYE homologues. Left half of the Figure the reduction of enzyme with NAD(P)H (reductive half reaction), Right half of the Scheme, regeneration of oxidised enzyme (oxidative half reaction).

1.3.1 Reductive half-reaction

In the reductive half-reaction, the reductant (NAD(P)H) donates two electrons and an associated proton (hydride; H^-) to the enzyme-bound FMN cofactor (Scheme 1.6). The reductive half-reactions of OYEs with nicotinamide cofactors have been shown to proceed via an enzyme-substrate charge transfer complex, which is characterised by an absorbance maxima at 560 nm [38]. Due to the fast reaction rates and complex nature of the transients, it is possible that there are additional unidentified intermediates in the reaction [38].

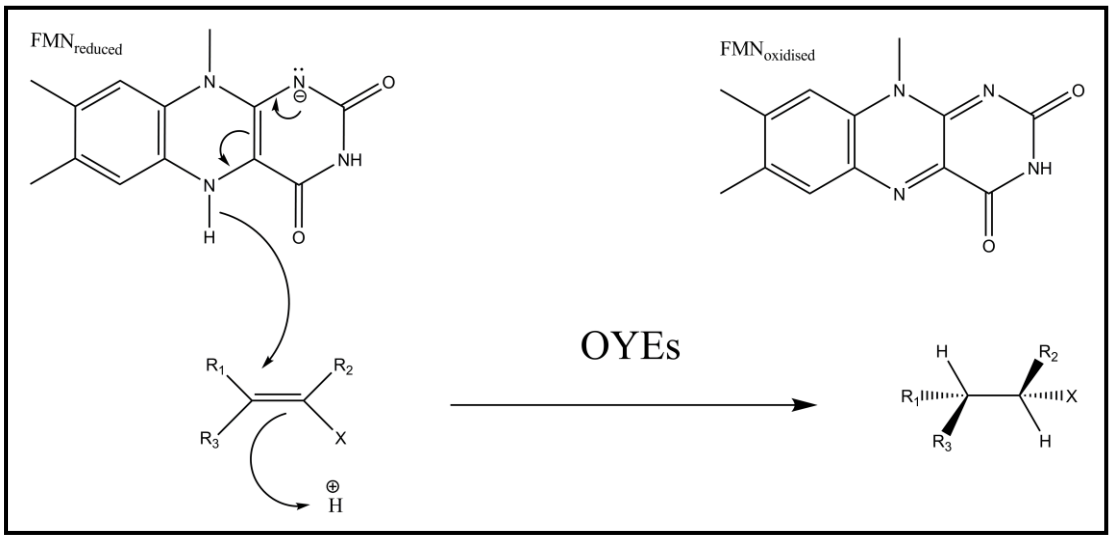


Scheme 1.6. Schematic representation of OYE reduction

1.3.2 Oxidative half-reaction

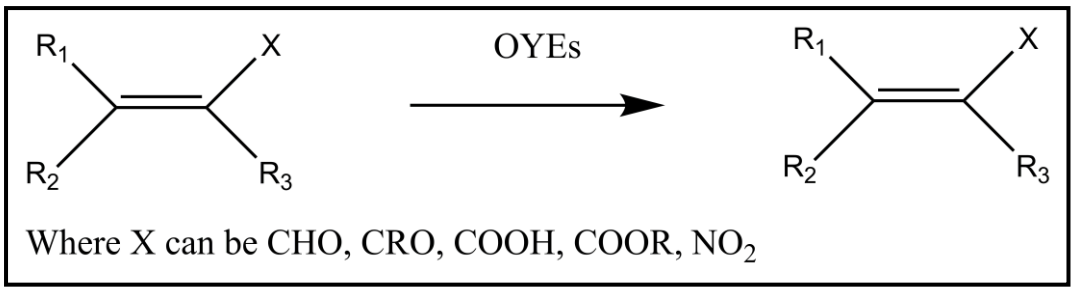
1.3.2.1 Reduction of α/β activated alkenes

The reaction is proposed to proceed via hydride attack at the β carbon, followed by protonation at the α carbon (Scheme 1.7) [6]. The main determinant of reactivity in this form of reaction is the activating group present, which is largely responsible for ligand/enzyme interaction [40].



Scheme 1.7. Schematic representation of the oxidative half reaction; X, activating group; H^+, derived from tyrosine residue or solvent, see Section 1.4.3.

OYE family activity has been detected with a wide variety of activating groups including aldehydes (enal), ketones (enone), carboxylic acids, esters, nitroalkenes as shown in Scheme 1.8



Scheme 1.8. Generalised substrate profile of activated alkene substrates with OYEs.

The most commonly described ene reduction of OYE's is the reduction of 'enone' compounds where the activating groups is a ketone. This reaction has been observed with a wide variety of substrates including cyclic and acyclic enones. An example of this type of reaction is the reduction of CH which has been recognized as a good general substrate for OYE homologues [1]. This activity has been demonstrated to extend to related compounds including maleimides, which contain two carbonyl groups [59].

One of the best characterised enone reductions is of the C_1 to C_2 double bond of the A ring of steroids such as prednisone. Steroids which lack this allelic bond, but like progesterone contain a C_4 to C_5 double bond, are effective inhibitors of PETNR [40]. Studies have also been performed using C_2 and C_3 substituted variants of cyclo pent/hex-enones, which have been shown to have a critical but differential effect on the substrate acceptance of OYEs [60]. Recent studies have widened the substrate acceptance of these substrates in YqjM by saturation mutagenesis studies [61].

A wide range of 'enal' compounds have been shown to be active with OYE homologues [2]. One example which has been observed in a wide range of OYE homologues is the reduction of cinnamaldehyde [21]. This class of substrates also includes a number of cyclic and acyclic terpenoids [59]. The substrate range for OYE1 has been demonstrated to include a range of short chain aliphatic, long chain aliphatic and branched aldehydes [2], with a similar profile for PETNR [59].

Additional to the reactions with enones and enals a range of other carbonyls have also been demonstrated to be substrates including; carboxylic acids and esters,

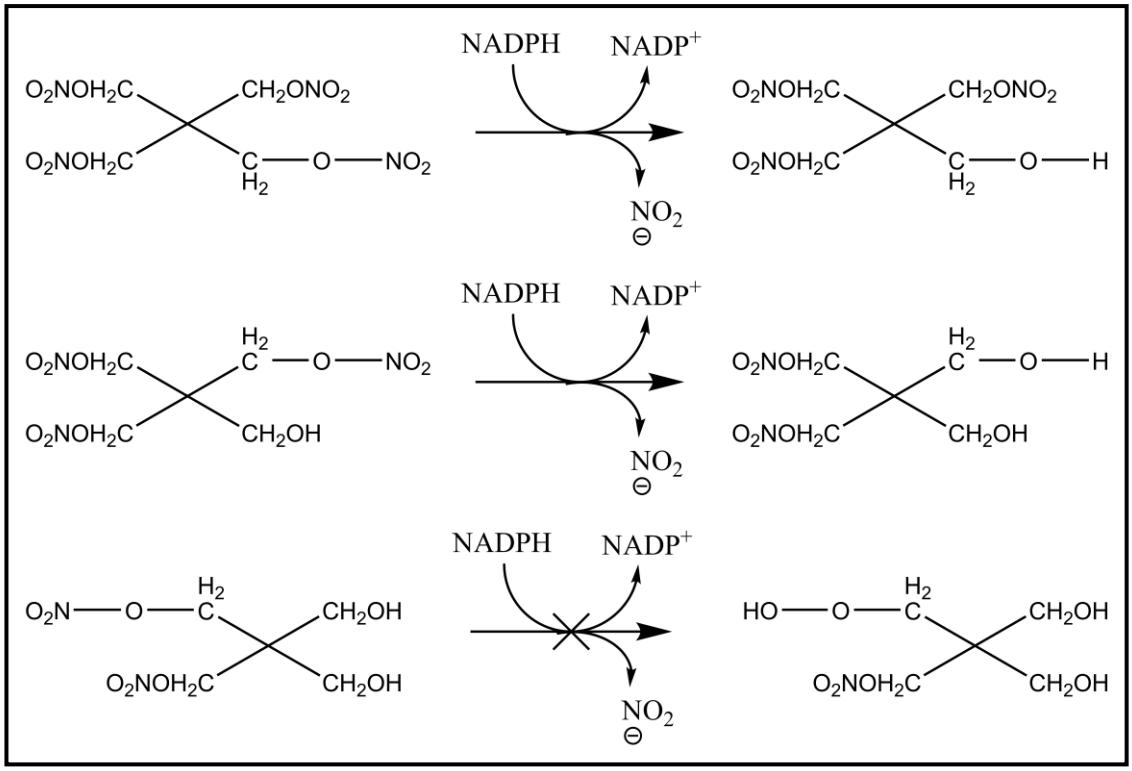
however, a strong preference for the di-acid species and the di methyl esters is observed [54, 62].

In addition to carbonyls nitro groups have also been shown capable of acting as activating groups for the reduction of unsaturated alkenes in OYEs, the simplest being nitrocyclohexene [63]. A range of α and β substituted nitro compounds have been studied with PETNR [63].

1.3.2.2 Reductive denitration of nitrate esters

Several OYE homologues have been isolated with the ability to remove the nitrate groups from nitrate esters, including PETNR, GTNR and XR. The ability of PETNR-containing bacteria to utilise nitrate explosives as a source of nitrogen stems from the ability of PETNR to reductively liberate nitrite from these compounds [3]. GTNR is capable of very similar oxidative chemistry to that of PETNR though with differences in the reductive chemistry [23]. Much of the mechanistic work performed on this pathway has been done on PETNR.

Identification of the metabolites of PETN reduction by PETNR were investigated by high performance liquid chromatography (HPLC) and electrospray ionisation mass spectroscopy (EIMS) [3]. NO_2^- groups were found to be sequentially removed, using one molecule of NADPH per NO_2^-, to give the trinitrate and dinitrate derivatives of PETN (Scheme 1.9). However, as mononitrate was not detected, the dinitrate is probably not a good substrate for the enzyme [3]. When similar experiments were performed with a whole cell system, the degradation products were more complex as enzymes other than PETNR can act on the di- and trinitrate metabolites to give the aldehyde derivatives (rather than the alcohol derivatives from the initial liberation of the nitrate group) of the intermediate metabolites [3].

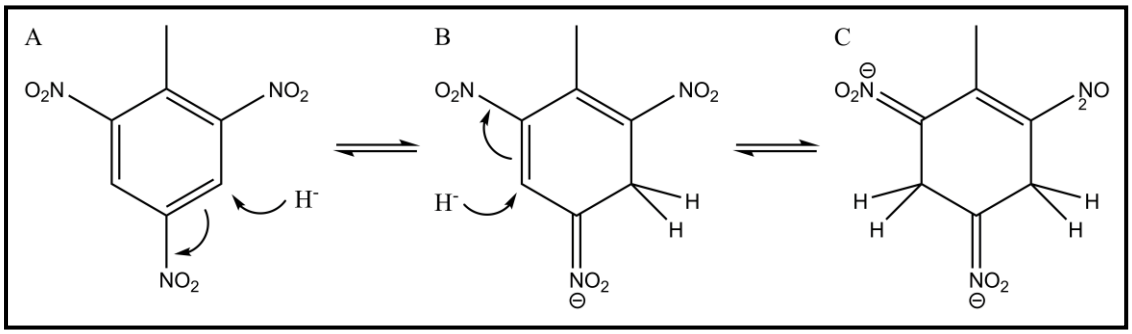


Scheme 1.9. Reductive denitration of PETN by PETNR. Panels A and B show sequential reductive denitration of PETN as catalysed by PETNR, panel C shows the expected reaction of the dinitrate substrate which is not observed experimentally.

PETNR has a similar activity toward other nitrate ester explosives such as GTN and ethylene glycol dinitrate (EDGN). Since the dinitrate is not a good substrate it can be inferred that other than the nitrate involved in the reaction a further two nitrate ester groups are required for good enzyme/substrate binding [3]. Since these initial experiments a whole range of nitrate esters were demonstrated to undergo the same reaction along with a number of nitramine species including erythritol tetranitrate, mannitol hexanitrate, xylitol pentanitrate, inositol hexanitrate, diglycerol tetranitrate, 1,2-propylene glycol dinitrate, 1,2-hexanediol dinitrate, 1,2-pentanediol dinitrate, ethylene glycol dinitrate, 1,2-butanediol dinitrate, N-nitro-diethanolamine dinitrate, butyleneglycol dinitrate, diethyleneglycol dinitrate, 1,3,5-Trinitroperhydro-1,3,5-triazine (RDX), nitroguanidine, ethylene-1,2-dinitramine, 1,3,5,7-tetranitro-1,3,5,7-tetrazocane (HMX) [44].

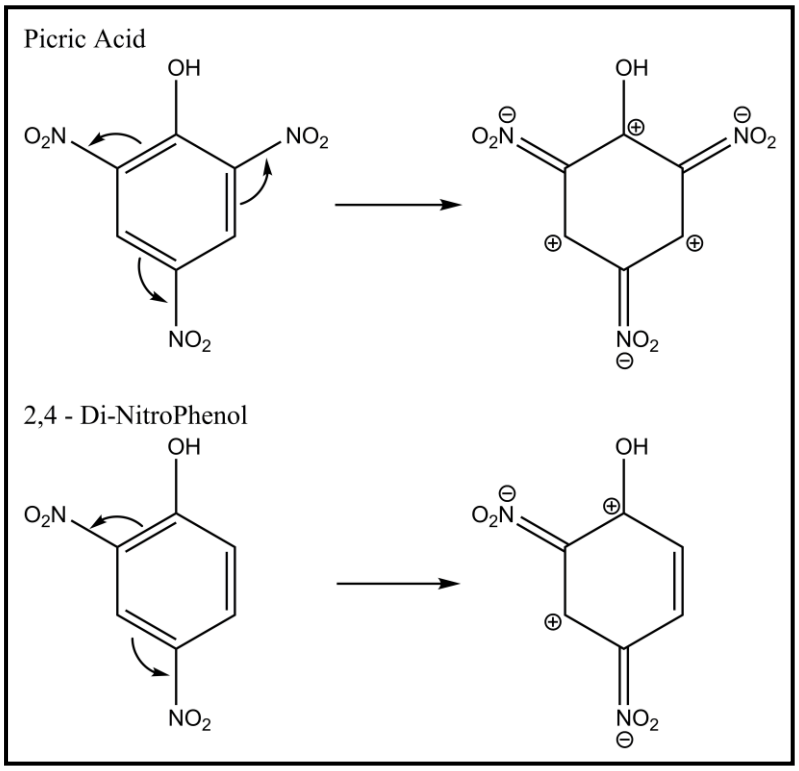
1.3.2.3 Reduction of the aromatic ring

At present the mechanism of dearomatisation has only been found in a few OYEs exhibiting high levels of homology with PETNR including PETNR, NEMR [4] and XRs [64]. Reduction of the aromatic ring was initially shown to proceed via the formation of a hydride Meisenheimer intermediate (H^-TNT; Scheme 1.10). This intermediate is generated by a hydride attack on the C_4 of the aromatic ring [5]. When following the reaction spectroscopically, orange/yellow products were detected from the reaction with TNT. These same orange/yellow products were also detected when PETNR was added to a chemically synthesised solution of hydride Meisenheimer complex, which has a purple/brown colouration. When chemically synthesised hydride Meisenheimer intermediate is left in an excess of borohydride, the orange colouration develops over an extended period of time indicating the orange/yellow products are a further reduction product of TNT [5]. The appearance of only the orange product in enzyme catalysed reactions indicates the hydride Meisenheimer complex exists only transiently and does not accumulate suggesting the overall rate limiting step is its formation [5]. The reaction in XR also proceeds via the addition of hydride to give di/hydride Meisenheimer intermediates and suggests a non-enzymatic mechanism for dimerisation and liberation of nitrite, which also involves hydroxylamino intermediates from the nitro reduction pathway indicating a balance between pathways rather than exclusion of one over another [57, 64].



Scheme 1.10. The ring reduction pathway including TNT substrate; Panel A, the hydride Meisenheimer intermediate; panel B, and the di-hydride Meisenheimer; panel C, generated during the reaction.

It has been suggested, by analysis of x-ray diffraction models of PETNR soaked with various substrates and inhibitors, that in order to facilitate hydride transfer the C_5 atom of the substrate needs to be localised above the N_5 of the flavin [38]. In substrates such as TNT and picric acid (PA), the C_3 and C_5 atoms are resonance-activated. In the case of the inhibitor 2,4-dinitrophenol (DNP), only C_3 has resonance activation due to the absence of a nitro-group on C_6 (Scheme 1.11). This is thought to be the mechanism of discrimination between substrate and competitive inhibitors as resonance activation at C_5 is a requirement for catalysis [38].

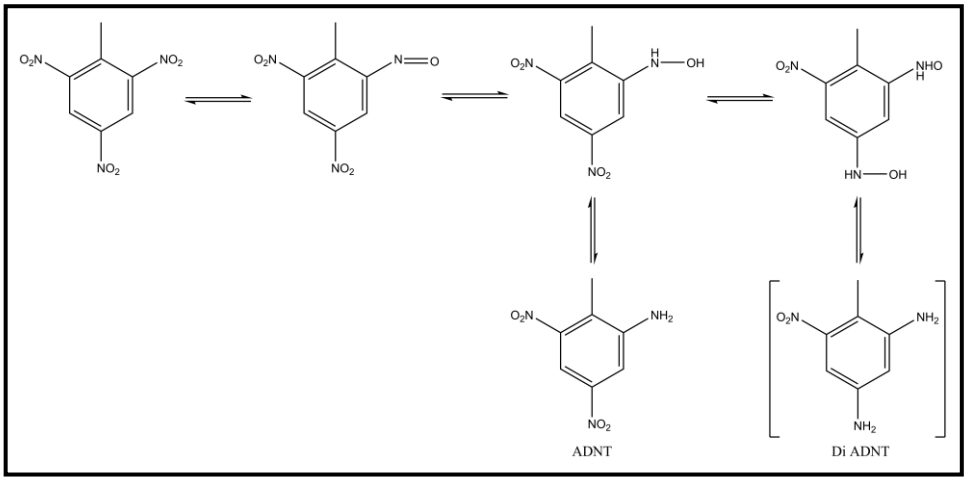


Scheme 1.11. Resonance activation of the substrate picric acid and inhibitor 2,4-Dinitrophenol (DNP) demonstrating possible mechanism for the delineation of substrate and inhibitor.

1.3.2.4 Nitro group reduction

The mechanism by which OYEs reduce nitro groups to amines proceeds via the formation of an initial hydroxylaminodinitrotoluene (HADNT) product. This is followed by either reduction of a second nitro-group of HADNT to dihydroxylamino-nitrotoluene, or reduction of the hydroxylamine group to form aminodinitrotoluene (ADNT; Scheme 1.12). While there seems to be potential for the dihydroxylamino-

nitrotoluene species to be further reduced to give a diaminonitrotoluene species, none was detected in the above study [4]. The nitro reductive pathway has also been shown to provide potential substrates for the non enzymatic condensation with intermediates of the ring reduction pathway [57, 64].

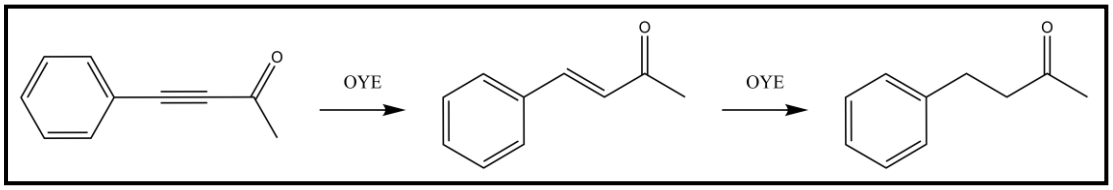


Scheme 1.12. Proposed reaction Scheme for the reduction of nitro groups in TNT. The undetected diaminotoluene species is shown in brackets.

As opposed to the ring reduction pathway, the nitro reduction pathway is consistently present in the OYEs tested [4]. In homologues where both pathways are present there is a delicate balance struck between the nitroreductive pathway and the direct reduction of the aromatic ring. The flux through these two pathways can be manipulated in PETNR by the substitution of a single amino acid [4].

1.3.2.5 Alkyne reduction

In a more recent development the reduction of the alkyne 4-phenyl-3-butyne-2-one (Scheme 1.13) has been observed using a range of OYEs, suggesting a previously unknown activity of OYEs which may prove to have potential as a biocatalyst [6].



Scheme 1.13. Alkyne reduction observed in OYEs, Initial reduction to the alkene followed by second undesirable reduction to the alkane product.

1.4 Structure-function relationships in OYE family

A number of structure-function relationships have been elucidated for the OYE family of enzymes. These are critical to understanding the enzymes function at a molecular level and as such will be critical in our attempt to manipulate the function of the enzyme. In the following section, the numbering of residues refers to PETNR residue identity throughout unless otherwise stated.

1.4.1 FMN binding

Each OYE monomer non-covalently binds a single FMN group, however, the nature of the interactions responsible for coordinating and modulating the function of the prosthetic group have only begun to become clear in the last 20 years. From the crystal structure of OYE1, the flavin prosthetic group was shown to be buried centrally within the 8 stranded barrel fold with access to the *si* face of the FMN from the solvent filled access channel (Figure 1.4) [15].

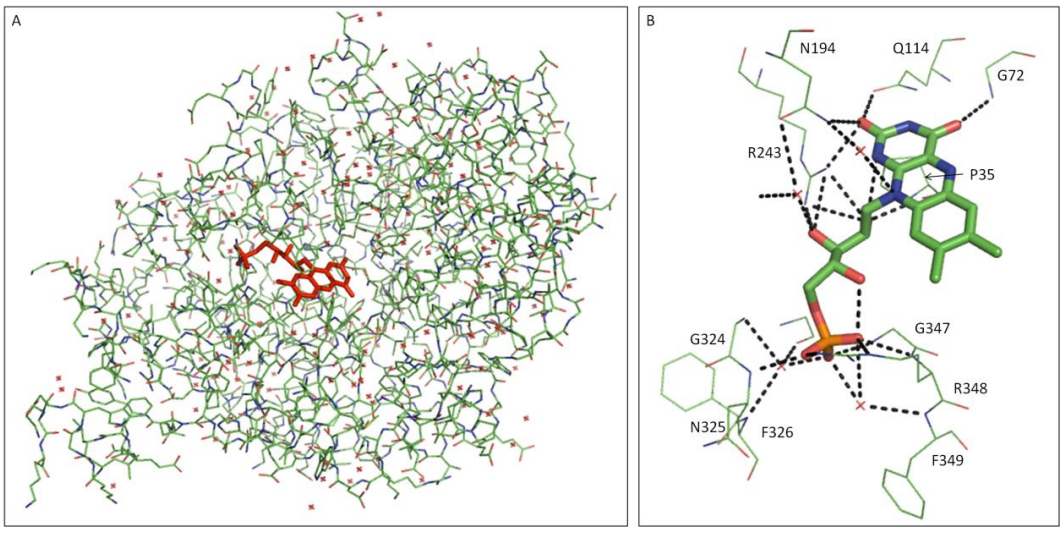


Figure 1.4. Binding of FMN to OYE1; panel A, localisation of FMN in the active site of OYE1; panel B, binding of FMN in the OYE1 active site.

The flavin is held in place by interactions from a wide number of protein interactions including side-chain hydrogen bonds, main-chain hydrogen bonds and water mediated hydrogen bonding (Table 1.2) [15, 65]; interactions similar in number and nature have been observed from structures of PETNR [40], MR [41] and OPR [45].

Interaction	Residue atoms	FMN atoms
Side Chain	T37	O4
	Q114	N3, O2
	R243	O2', O3', O2
	R348	PO
Main Chain	P35	O2'
	T37	N5
	G72	O4
	N325	PO
	G347	PO
Water Mediated	G324	PO
	F326	PO
	G345	PO

Table 1.2. Interactions responsible for the coordination and mediation of FMN in OYE

1.4.1.1 Role of residue 26 (OYE 37)

This residue sits adjacent to the isoalloxazine ring of the FMN and forms hydrogen bonds between the backbone N of T37, the N_5 of FMN and the side chain OH to the O_4 group of the FMN (Figure 1.5; [66]).

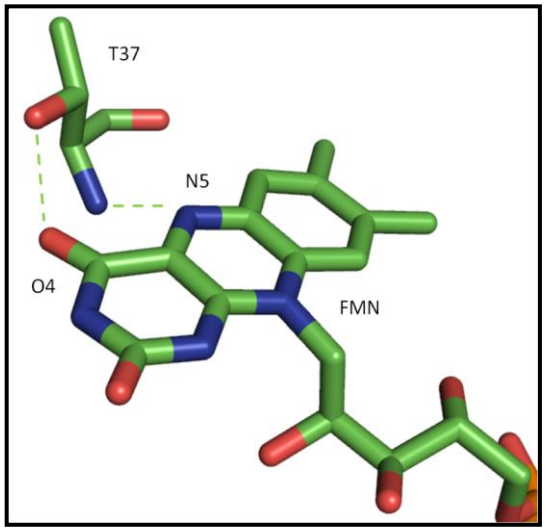


Figure 1.5 Hydrogen bonding between the Thr37 residue in OYE1 and its flavin group. From the PDB file 1OYA [15].

Mutagenesis of this residue to alanine in OYE1 caused the rate of the reductive half reaction to be significantly reduced, though the rate of the oxidative half-reaction with several substrates increased [66]. This suggests the mutation has destabilised the reduced form of the flavin leading to a faster transfer of electrons to the substrate [66]. Similar results were observed with the T32A mutant of MR [67]. In YqjM, this residue is not conserved and its role is thought to be partially fulfilled by Y375 [43].

1.4.1.2 Role of residue 100 (OYE 114)

Q114 from OYE1 is another highly conserved residue in the active site of the OYE family of enzymes. Its side chain atom O1 often hydrogen bonds to the N3 of the flavin isoalloxazine ring and the N1 interacts similarly with the O2 of the FMN. The crystal structure of the Q114N mutant of OYE1 showed the isoalloxazine ring of the flavin was shifted 12° away from its position in the wild type enzyme, causing the ring to form hydrogen bonds with the His/Asn pair. This could explain the lowered affinity of this mutant for the oxidative substrate [68] as this would require the breaking of hydrogen bonding of the FMN to enzyme [68].

1.4.1.3 Role of residue 181/184

These residues have been demonstrated to have a significant effect on the binding of FMN [69]. The OYE1 N194H mutant incorporating histidine at both positions bound FMN poorly, which may be due to steric hindrance in a smaller OYE active site [69]. Similar results have also been obtained by mutating the critical His residues in GTNR [70].

1.4.2 Nicotinamide coenzyme binding

The consensus sequence for binding nicotinamide coenzyme is well conserved among NADH-binding proteins and is represented by the sequence GXGXXG [71]. This sequence is usually well conserved in NADH binding proteins, but less so in proteins which bind NADPH. These residues form of a tight helix and provide space for the cofactor to bind [71]. There are often additional residues which stabilise parts of the bound cofactor. For example two arginine residues are usually found in

many NADPH-binding proteins downstream of the GXGXXG sequence responsible for coordinating the 2' phosphate group present in NADPH but not in NADH, suggesting a possible mechanism for distinction between cofactors [71]. Further arginine residues in loop 3 (R142/R130) are thought to play a role in binding the 2' phosphate group of NADPH [40].

Trends have been observed between the consensus sequence conservation of OYE homologues and the coenzyme preference, for example PETNR (GXGXXG – not conserved, Arg pair – well conserved) prefers NADPH while GTNR (GXGXXG – well conserved, Arg pair – not conserved) prefers NADH. However this relationship is not absolute as MR contains a slightly modified consensus sequence and has a conserved Arg pair, yet it prefers NADH. Overall OYEs exhibit a requirement for NADH but a preference for NADPH, so the method of binding the cofactor is likely to exclude NADPH rather than actively select for either [72].

Residues 181 and 184 have also been demonstrated in PETNR and MR to have a role in NADPH binding of the respective enzymes, which provide contrast between His/His and His/Asn pairs, and enzymes with different reductive cofactor preference. In the case of MR, the reductive half reaction is affected by altering the enzymes affinity for the cofactor enzyme charge transfer complex [73] with a similar role implied in PETNR [69]. In MR further research has shown that in the N189A mutant multiple conformations of NADH, all catalytically viable, were predicted and demonstrated to be kinetically distinguishable further supporting a role for this residue in binding of substrates [74].

1.4.3 Oxidative half reaction

1.4.3.1 Role of residue 26

T26 is conserved throughout all classical OYEs, in thermostable like OYE's it is conserved as Cys. In PETNR and OYE it is positioned above the isoalloxazine ring of the FMN cofactor. Studies have shown that T26 in PETNR prevents the reduction of the C4 to C5 bond in steroid inhibitors, such as progesterone, by restricting access to a catalytically active conformation [73]. In order for hydride transfer to occur

with oxidative substrates, the angle between the N10 and N5 atoms of the flavin and the substrate C_β atom must be in the region of 125-170°. However, crystal structures have shown that the N10 - N5 - C4 bond angle between PETNR FMN and progesterone is much too small [73]. The structure showed that progesterone is not bound in a catalytically optimal orientation (Figure 1.6) [73].

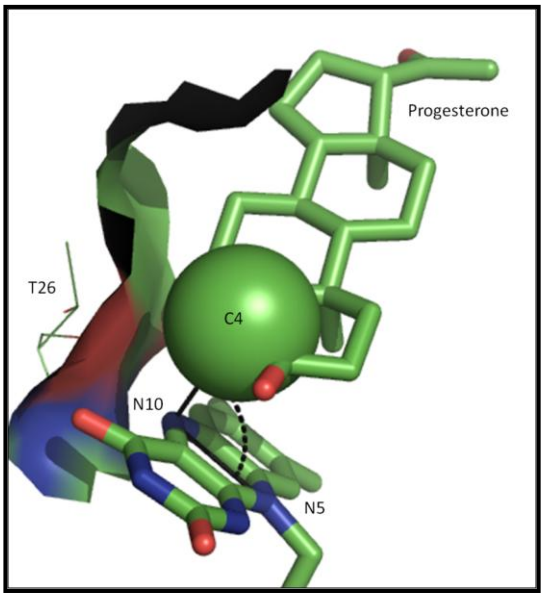
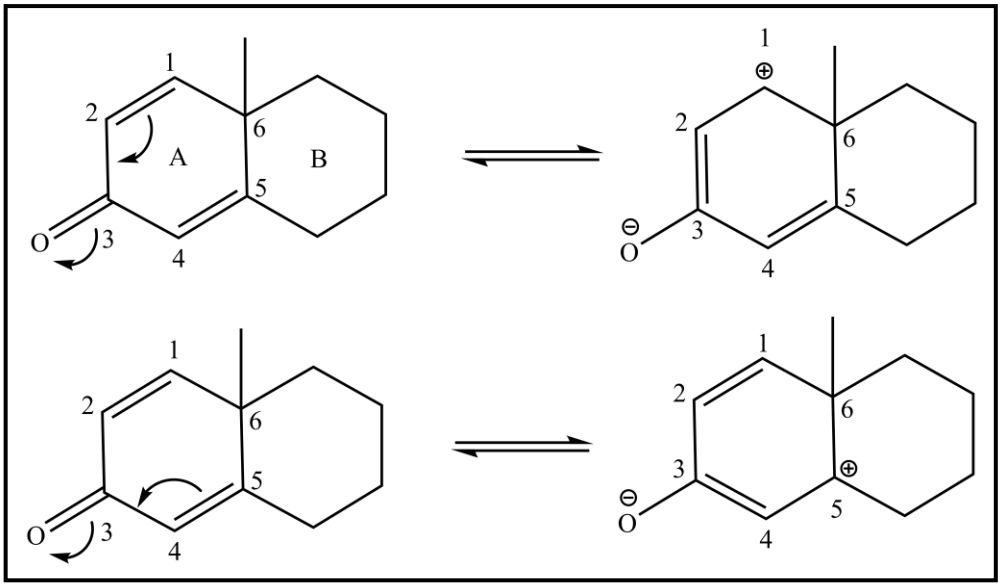


Figure 1.6. Structure of PETNR-bound progesterone showing the steric clash between progesterone and T26 side chain in the active site of PETNR. Progesterone and FMN, atom coloured sticks with green carbons; T26, shown as a surface; C4, shown as sphere. Taken from the PDB file 2ABA [73].

Site-directed mutagenesis of T26 to smaller residues may remove the steric hindrance with steroid substrates. This could open up the possibility of engineering mutants able to catalyse the addition to the C4 to C5 bond of these inhibitors. An analysis of the chemical environment of the double bonds suggests that both bonds may become partially activated by resonance forms, although the C5 is more sterically hindered (Scheme 1.14).



Scheme 1.14. Resonance forms of the A and B rings of generalised steroid backbone showing the partial charge activation on the C1 (top) and C5 (bottom).

1.4.3.2 Role of residue 68

The crystal structure of PETNR has shown that Y68 is one of a pair of spatially related tyrosines (along with Y351) which are positioned in the active site channel [40]. A role in oxidative binding is likely due to their positioning at the mouth of the active site channel (Figure 1.7). In MR a role as proton donor was suggested due to the absence of Y186, however such a function was disproved [67]. Interestingly mutants generated at this position proved to be catalytically competent, implying its identity is not strictly conserved [67] suggesting modification at this position could provide functionally improved mutants.

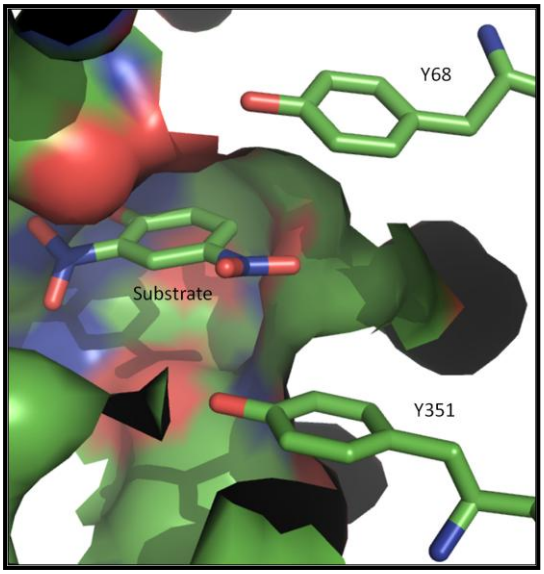


Figure 1.7 Location of the tyrosine residues 68 and 351, in PETNR, with relation to the active site tunnel; residues/substrate, atom coloured sticks; PETNR active site, surface. Taken from the PDB 1GVO [38].

1.4.3.3 Role of residue 102

W102 is highly conserved throughout the classical OYEs but not so in the thermostable-like OYEs. The crystal structure of picric acid bound PETNR showed an unusual electron density around W102 [38] which, following solution of a higher resolution data set, was interpreted as W102 occupying different conformations, in which the W102 residues overlapped, in the presence and absence of substrate [75]. This residue has little effect on either the reductive half reaction or on the binding of steroid substrates. However it has a profound effect on the flux between the ring reduction pathway and the nitroreductive pathway [75]. The crystal structures of W102Y and W102F mutants showed this residue adopted a single conformation, which removed the steric clash with the C6 nitro-group of TNT (Figure 1.8; [75]).

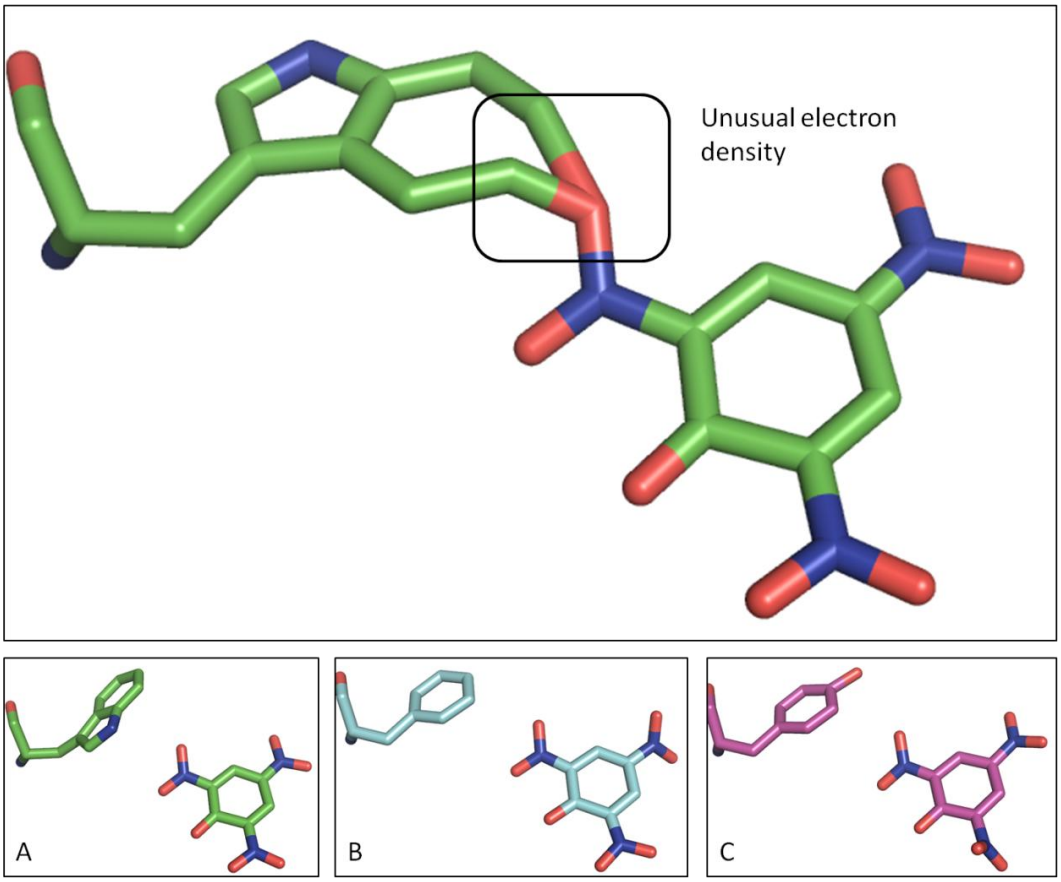
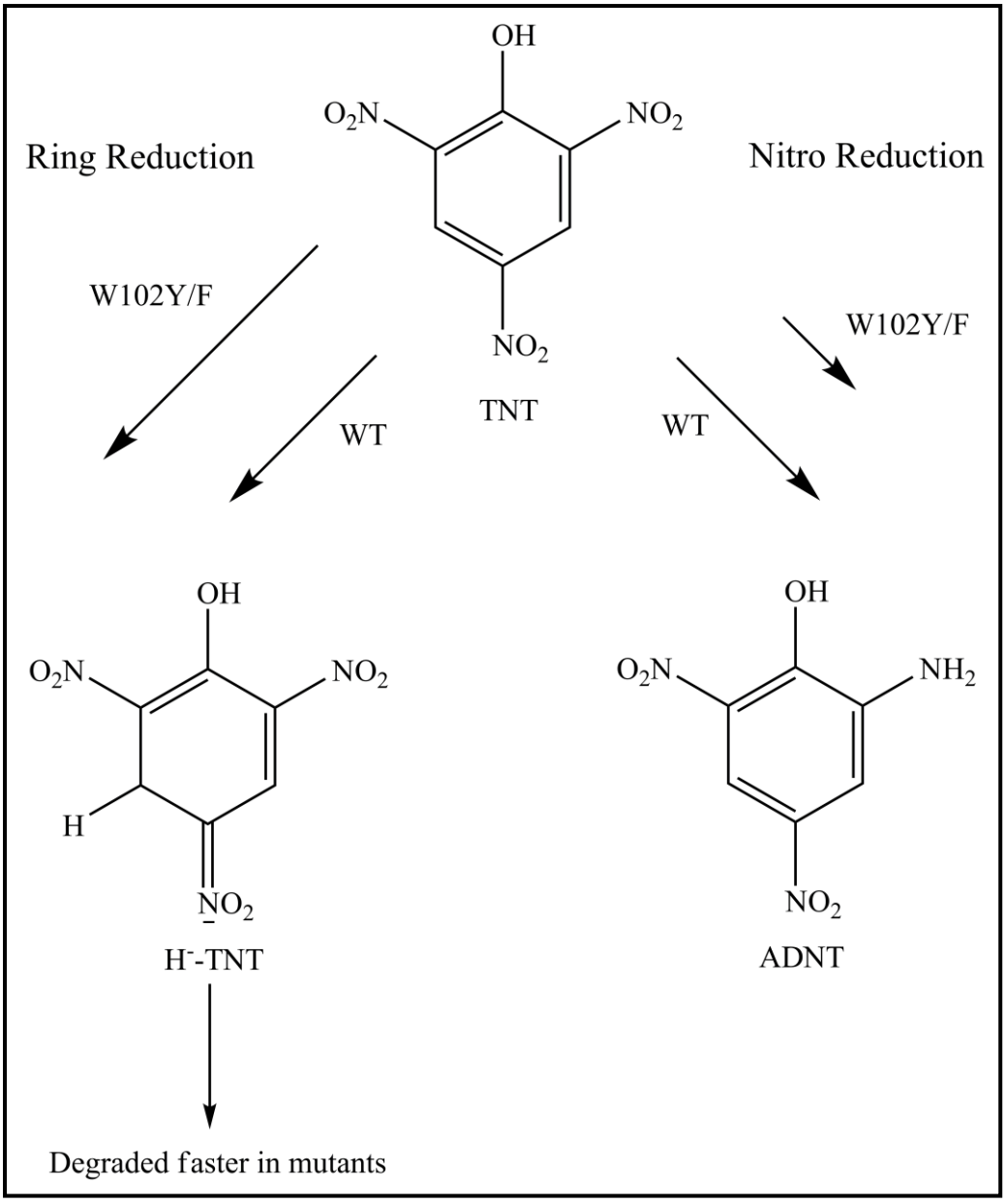


Figure 1.8. The role of Trp102 in substrate binding. Main panel, Unusual electron density surrounding Trp102 seen in the wild type PETNR structure shown as sticks with green carbons from PDB file 1VYR [75]: Panel A, alternate conformation of W102 shown as sticks with green carbons from 1VYR [75]; Panel B, W102F mutant shown as sticks with cyan carbons from PDB file 1VYP [75]; Panel C, W102Y mutant shown as sticks with magenta carbons from PDB 1VYS [75].

In these mutants, the hydride Meisenheimer intermediate from the ring reduction pathway was formed at the same rate, but was broken down at a faster rate than in wild type enzyme. This suggests that the mutant enzymes favour the ring reduction pathway over the nitroreductive pathway (Scheme 1.15; [75]). Therefore, W102 is an attractive target for PETNR mutagenesis due to the possibility of generating mutants favouring the ring reduction pathway, as well as turning on this activity in OYE homologues which do not possess this activity. The role of the respective residue in the reduction of α/β unsaturated activated alkene compounds was investigated in MR and shown to be non-essential for catalysis but to have a substantial influence on specificity [67].



Scheme 1.15 Flux down the two pathways available to some OYE family members. Bracketed arrows indicate the change in flux accompanying the mutations of W102Y and W102F (not to scale).

1.4.3.4 Role of residue 181/184

Residues 181/184 are highly conserved throughout the OYE family. These residues are located in the centre of the substrate binding pocket above the flavin (Figure 1.9), and play a central role in substrate recognition in the oxidative half reaction. Throughout the OYE family, H181 is invariant whilst the amino acid equivalent to position 184 varies between histidine and asparagine (Table 1.3). The presence of a

H181/H184 pair (as in PETNR and NEMR; [4]) is rare, with residue 184 usually an asparagine (Figure 1.10).

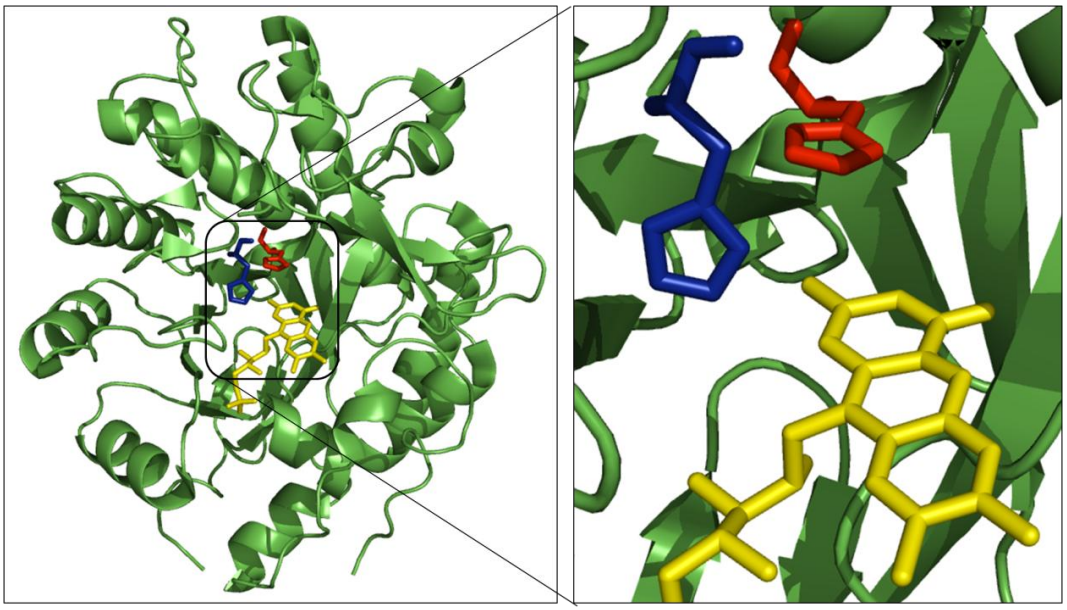


Figure 1.9 Overall structure of PETNR. His181 is shown as red sticks, 184 is shown as blue sticks, and FMN is shown as yellow sticks. Taken from the PDB file 1H50 [40].

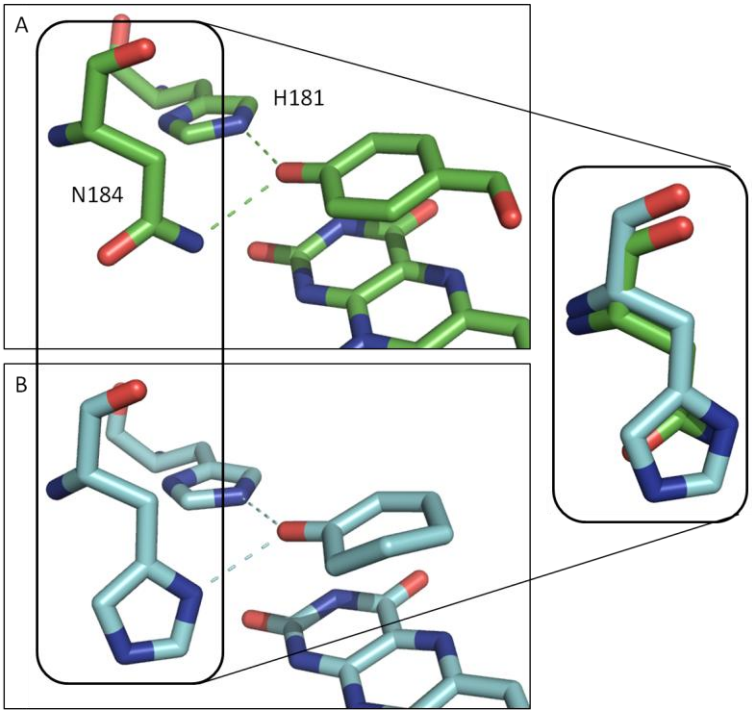


Figure 1.10 Interaction between the active site coordinating residues and enone substrate, comparison of the interaction of a His/His and His/Asn pair. Panel A, His/Asn, OYE1/HBA, green atom coloured sticks, 1OYB [15]; Panel B, His/His, PETNR/CH, cyan atom coloured sticks, 1GVQ [38]; Inset, Overlay of His 184 and the respective Asn showing colocalisation of coordinating nitrogen.

Protein	Residue 181	Residue 184
PETNR	His	His
NEMR	His	His
EBP	His	His
OPDR	His	His
XR	His	His
YqjM	His	His
TOYE	His	His
MR	His	Asn
GTNR	His	Asn
OYE1	His	Asn
OYE2	His	Asn
OYE3	His	Asn

Table 1.3 Selected examples of OYE family members and their configuration around the 181/184 positions.

The binding of ligands to the PETNR active site has been extensively studied by crystallographic methods. The recognition of the substrate by the enzyme active site is achieved through interactions between the crucial 181 and 184 residues (Figure 1.11). Crystallographic studies of OYE1 identified residues equivalent to 181 and 184 as potential hydrogen bond donors responsible for the coordination of the substrates *p*-hydroxybenzaldehyde and *p*-estradiol in a position where interaction between the substrate and prosthetic group can occur [15].

Similar studies with PETNR have demonstrated that the residues form hydrogen bonds with the carbonyl group of the α/β unsaturated substrate and a range of other functionalities (Figure 1.11) leading to the localisation of the resonance activated double bond above the N5 atom of the flavin, in the correct orientation for hydride transfer [38, 40]. Other examples include OPR1 [45], GTNR [70], MR [67].

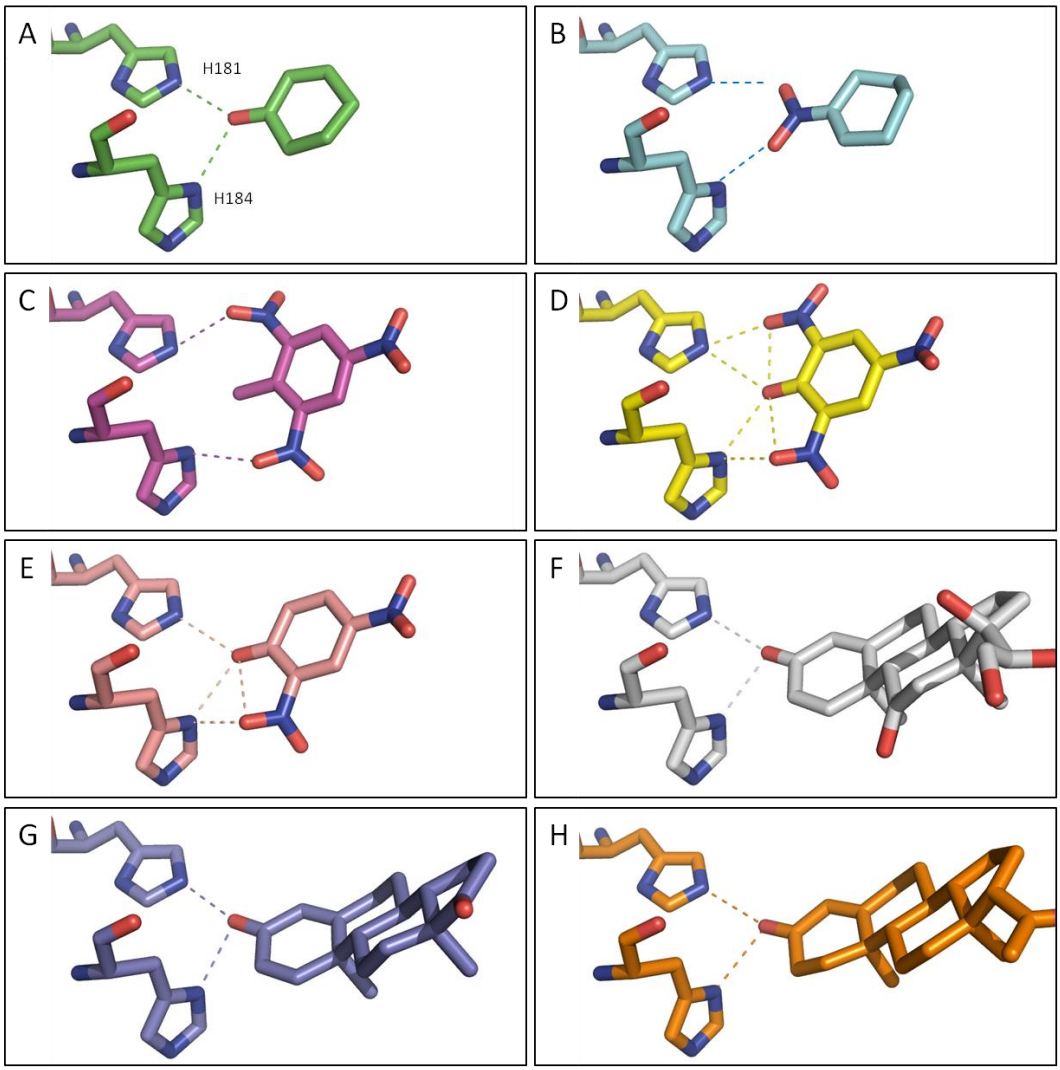


Figure 1.11. Binding of a number of PETNR substrates into the active site through interaction with H181 and H184. Panel A, 2-cyclohexenone, 1GVQ [38]; Panel B, 1-nitrocyclohexene, 3F03 [63]; Panel C, 2,4,6 trinitrotoluene, 1GVR [38]; Panel D, picric acid, 1GVS [38]; Panel E, 2,4 dinitrophenol, 1GVO [38]; Panel F, prednisone, 1H61 [40]; Panel G, progesterone, 2ABA [73]; Panel H, 1,4-androstadien-3,17-dione, 1H62 [40]. All residues and ligands are shown as atom coloured sticks.

While the interactions between the 181/184 residues and substrates differ amongst OYEs, the nature of the interactions are similar. For example, the hydrogen bond between N1 of H184 and O1 of CH in PETNR [40] is replaced by a hydrogen bond between the N2 of A194 and O4 of a similar substrate hydroxybenzaldehyde in OYE1 (Figure 1.10; [15]). OYE1 mutants in which H191N and N194H mutants were isolated demonstrated severely compromised substrate binding attributed to sub optimal enzyme substrate interactions [15, 69]. Interestingly problems with expression [15] and flavination of these mutants [69] have also been observed.

H191 has also been shown to be critical for the effective reduction of nitrate esters by PETNR [76] and GTNR [70], possibly by binding the terminal nitro ester. The above studies however also provide contrasting data with respect to K_d values for substrate-enzyme complexes. In the case of the reaction with CH poorer substrate binding was taken to account for a reduction in overall reaction rate [69]. In the case of GTN, better substrate binding was observed, allied to an inferior catalytic rate, explained by the enzyme binding substrate in an sub-optimal conformation for catalysis [76]. The hypothesis that oxidative substrate binding is responsible for the inferior catalytic efficiency is supported by the observation that reduced H181A and H184A mutants are still capable of being reoxidised [73]. Similar roles have also been observed for the respective residues in MR [67].

An interesting case study of the role of H181 and H184 in PETNR is the binding of steroid substrates and inhibitors. The crystal structure of PETNR is available with prednisone [40], progesterone [73] and 1,4-androstadien-3,17-dione [40], however despite this wealth of structural data the uncertainty surrounding the structural determinants of binding in catalytically active or inhibitory conformation remains [40].

There is also a correlation between the presence of a His at position 184 (or its equivalent) and the ability to catalyse the reduction of an aromatic ring [4]. When H184 in PETNR is mutated to Asn, the enzyme loses its ability to reduce the aromatic ring, as seen in naturally with OYE1. However, when provided with H^-TNT as a substrate, all enzymes (including the mutants) were still able to further reduce this intermediate and complete the reduction of the aromatic ring of this partially reduced compound. This suggests that H184 is most likely involved in the first hydride attack generating H^-TNT, and controls entry into the ring reduction pathway [4]. Mutation of this residue may allow us to deliberately affect the ability of the enzyme to enter the ring reduction pathway by the introduction of a single mutation.

Residues 181/184 are attractive targets for mutagenesis due to the importance of these residues in substrate binding. This may enable the accommodation of a wide

variety of non-natural substrates, including ones containing novel activating groups, e.g. nitrile, in place of the carbonyl which binds to residues 181/184. However it is also important to note that studies thus far have indicated catastrophic loss in activity for α/β unsaturated carbonyls, for example PETNR H181A and H184A with cyclohexenone [73]. However, this is thought to be due to binding issues rather than an inability to perform the correct chemistry [69]. It is also notable that mutations at these positions have led to mutants which could not be isolated or were incapable of interaction with FMN, for example N191H [69]. It has also been observed that generating mutants at these positions affects the reductive half reaction [69].

1.4.3.5 Role of residue 186

In OYE1, an active site Tyr (Y196) is known to act as the proton donor [77]. Mutagenesis of Y196 to Phe had no effect on the reductive half reaction, but the oxidative half reaction with CH was severely compromised [77]. Crystal structures of wild type OYE1 showed that Y196 is correctly positioned to account for the stereochemistry of the reduction of the double bond (Figure 1.12; [77]).

A similar role has also been hypothesised for Y186 in PETNR [38, 40] and OPR1 [45]. In GTNR the corresponding tyrosine is excluded from proton donation and thought to play a role in oxidative substrate binding [70]. However in the case of PETNR a role in proton donation was later ruled out in some substrates, by observation of the Y186F mutant, and solvent waters were implicated [73]

The corresponding residue in MR is C191 which has been shown not to be the proton donor by mutagenesis studies, nor to have any role in the oxidative chemistry [40]. In this case the proton donor is thought to be solvent water molecules [67]. Mutagenesis of this residue, in some OYE homologues, may open up the possibility of changing the stereochemical properties of the enzyme by altering the nature of the proton donation. In other reaction mechanisms of OYE, for example the reductive denitration pathway, residue 186 was demonstrated to play a role in substrate binding rather than for proton donation [76].

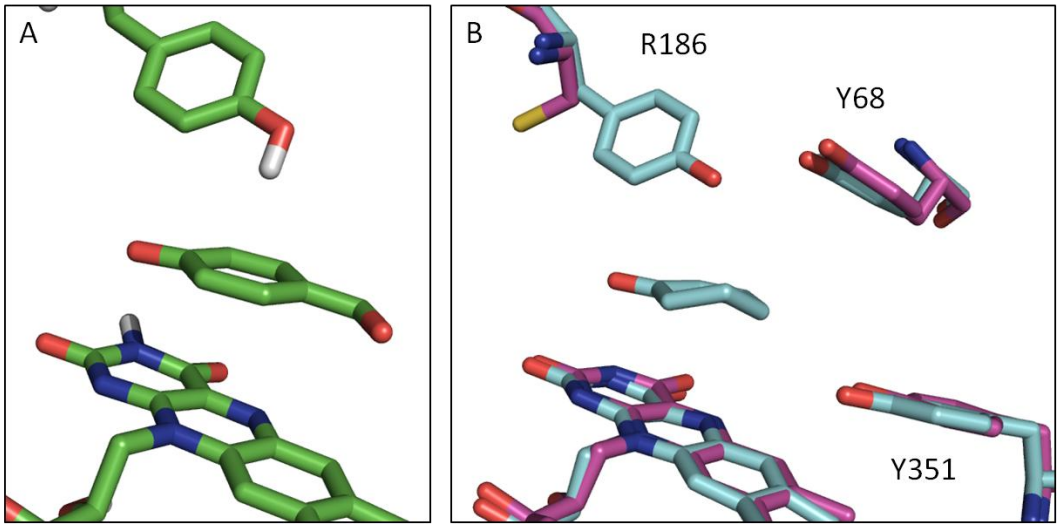


Figure 1.12. Tyr residues proposed as proton donors. Panel A, hydroxybenzaldehyde-bound OYE from 1OYB [15] showing the position of Y186 which has been demonstrated to act as the active site acid; Panel B, Overlay of residues posited to be active site acids in PETNR, Cyan, 1GVQ [38] and MR, Magenta, 1GWJ [41].

1.4.3.6 Role of residue 351

In classical OYE's residue 351 represents the other half of the tyrosine pair occupying a position in the mouth of the active site channel (Figure 1.13) [40]. The equivalent residue in OYE1 was initially highlighted as functionally important in the hydrogen bonding of the carbonyl oxygen of HBA in OYE1 crystal structures [15]. This residue was also demonstrated to move slightly to accommodate steroid substrates [15, 40]. This residue was also implicated as a potential hydrogen donor in MR, but was later demonstrated to have no such function [78]. Interestingly this residue is absent in the thermophilic class of enzymes [28].

1.4.3.7 Role of other residues

A wide range of other residues have been inferred from crystallographic studies [15, 40-42] to have mechanistic roles. Each of these residues is a potential site for library generation and hence influencing the substrate profile of PETNR.

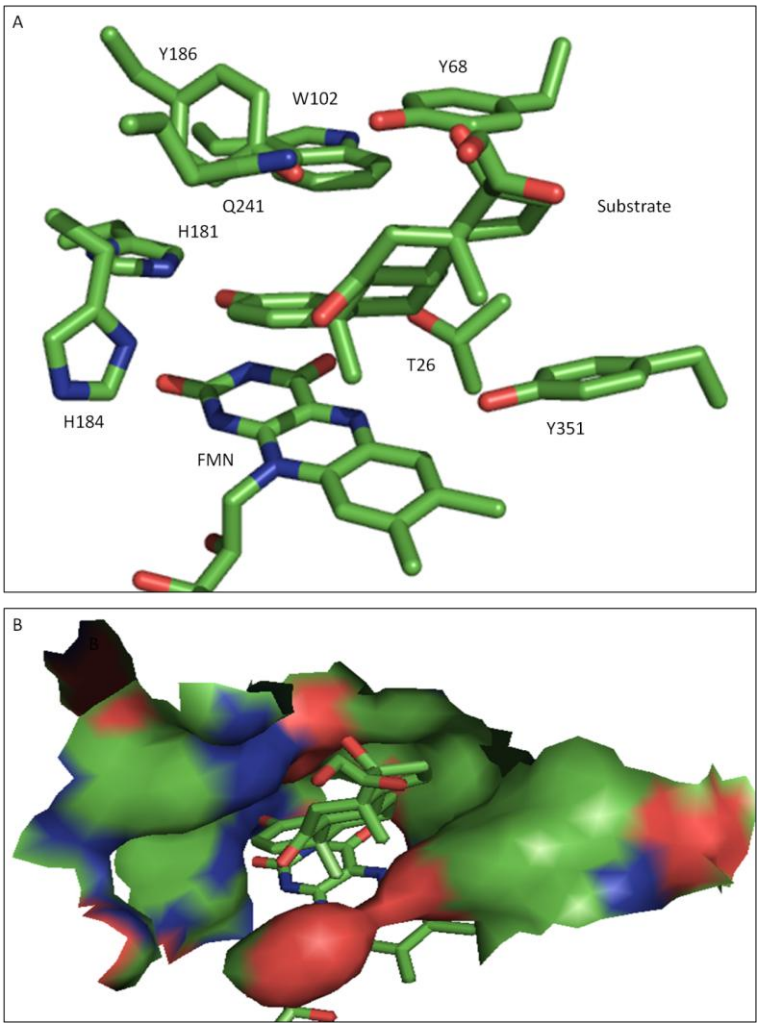


Figure 1.13. Figure showing the active site of PETNR occupied by the substrate progesterone, 1H61 [40]. Those residues identified in the text as being catalytically relevant have been highlighted, the inset gives a representation of the active site as a surface indicating where steric interactions may be present.

1.4.4 Oligomerisation

The oligomeric state of OYE homologues has been shown to vary greatly ranging from monomeric PETNR [3], through the dimeric OYE1 [15] and MR [41], up to the octomeric TOYE [28]. In the higher oligomeric states the enzymes are functional dimers due to the interactions of the arginine finger as discussed previously (Section 1.2.4).

In some cases information regarding the nature of the dimeric interface has been determined from crystallographic data (Figure 1.14). For example the analysis of crystal contacts in the model of OYE1 suggests a dimer interface incorporating a

number of direct and water mediated hydrogen bonds involving helices 4, 5 and 6 in both monomers [15], whilst the dimer interface in MR has been suggested to encompass helices 2 and 8 and β sheets A and B [41]. Whilst TOYE is considered as an octomer, the oligomeric organisation consists of 4 functionally active dimers rather than a traditional octomer, where dimerisation is achieved through yet another interface consisting of helices 1, E and F, oligomerisation of functional dimers occurs through the interaction of helices 6 and 7 [43].

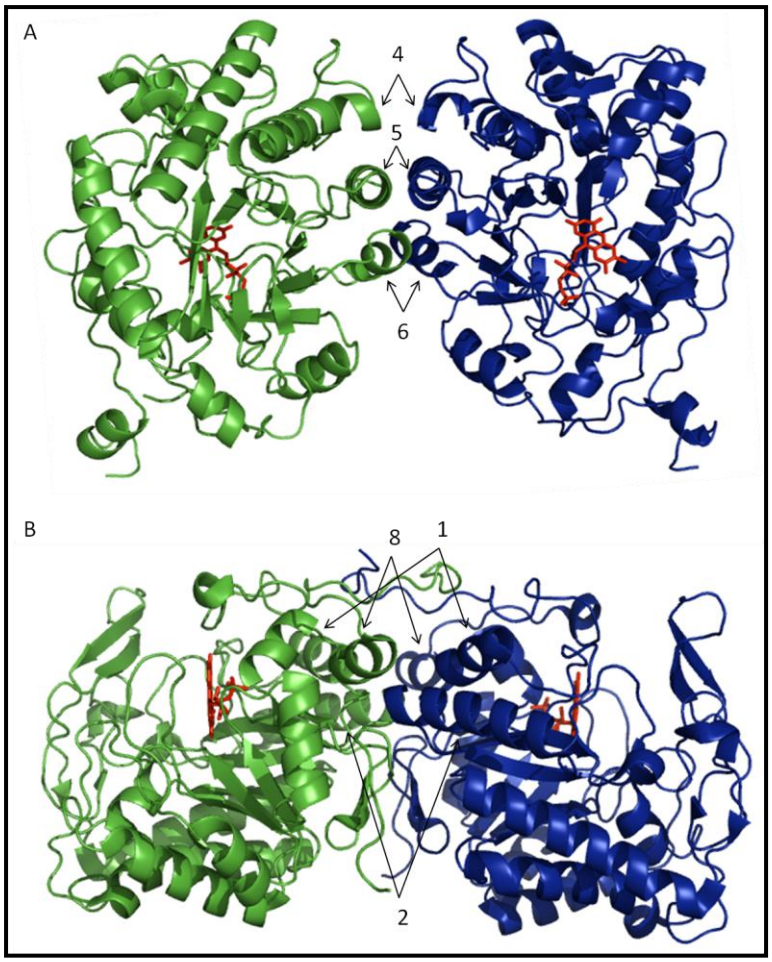


Figure 1.14. Dimerisation interfaces of OYE's. Panel A, dimerisation interface of OYE1 with interactions between helices 4, 5 and 6 from 1OYA [15]; panel B dimerisation interface of MR with interactions between helices 1, 2 and 8 from 1GWJ [41].

Interestingly a number of these dimeric interfaces have been demonstrated to be catalytically important. In the case of YqjM the formation of the dimer has been observed to result in the burying of R336 from one monomer into a second monomer to form a functional active site [43]. Interruption of this interaction, by

the addition of a poly histidine tag, lead to the loss of activity [79]. In the case of OPR dimerisation, the dimer formed by the binding of loop 6 in the active site was demonstrated to be functionally inhibited [42] highlighting another potential function of this phenomenon.

1.4.5 Summary

The information discussed in this Section is important in the consideration of two factors, first information regarding the structure-function relationships of the oxidative half reaction is used to inform the selection of residues at which to generate libraries. The review of the literature allied to analysis of the crystal structures of the OYE family has lead to the identification of T26, Y68, W102, H181, H184, Y186, Q241, and Y351 as a good initial selection of residues for library generation.

Additionally information regarding the structure-function relationships of OYE's with respect to a number of facets of enzyme activity, including oxidative half reaction, reductive half reaction, FMN binding and oligomerisation can be used to inform discussion of positive results and potential sources of error derived from the screening experiments.

1.5 Protein Engineering / Directed Evolution experiments

1.5.1 Introduction

There have been extensive studies of biotransformation reactions performed with OYE homologues both with isolated enzymes and whole cell preparations (Section 1.2.7). Inevitably, however, natural biocatalysts tend not to provide complete idealised substrate profiles. Improvements can be made by experiments which mimic the natural process of evolution, generation of diversity followed by application of selective pressure, on a micro scale over accelerated timescales in an effort to improve biotransformation efficiency and stereospecificity.

These experiments rely on the generation of genetic diversity followed by the application of selective/screening pressure to identify mutants with improved desired activity. Critical factors to the success of the experiment primarily involve the methods of library generation and screening [80].

The aim of this project is to apply these experimental techniques to PETNR, as an example of an OYE homologue, in order to evolve properties of the enzyme, leading to the generation of improved biocatalysts.

1.5.2 Library Generation

Critical to the success of these experiments is the method of library generation. A wide range of techniques for generating diversity have been developed ranging from single site techniques through whole genome mutagenesis [81]. Inherent in the discussion of diversity generation is the discussion and solution of 'the numbers problem' for the individual experiment which relates the screening effort to the library coverage which are both key determinants in the success of the experiment.

1.5.2.1 Gene Shuffling

The strategy behind gene shuffling experiments relies on the diversity present in a series of structurally related enzymes, positing the hypothesis that homologous enzymes have evolved separately to accommodate various functions and that by combining fragments of related enzymes, a number of functions may be conserved

or generated in hybrid clones [80]. This approach consists of fragmenting the pool of homologous genes, then randomly reassembling to give a large number of individual clones providing a library of DNA clones (Figure 1.15). This approach is described and was verified using *LacZ* as a template [82]. The major limitation to this kind of approach is often a lack of homologous enzymes [80], however in the case of PETNR a vast number of structurally similar enzymes have been identified, suggesting this may be an attractive approach for PETNR. However due to the wealth of structural data available for OYE homologues more targeted approaches are also available.

Methods have been developed for the generation of gene shuffled libraries to be completed by polymerase chain reaction (PCR), termed Staggered Extension Process (StEP), leading to the generation of gene shuffled libraries in a matter of hours. This procedure involves PCR with extremely short extension steps which allow only a short section of gene to be synthesised. The fragments are denatured and re-annealed, hopefully to a different parental gene [83].

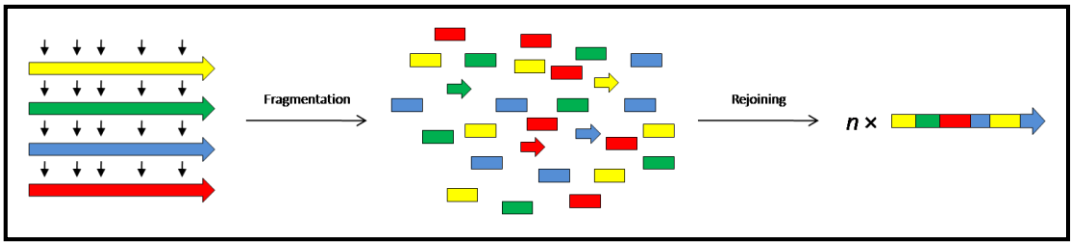


Figure 1.15. Schematic outlining the generation of mutagenic libraries by gene shuffling. Genes are initially fragmented before being pooled and rejoined at random generating a library of novel mutants.

1.5.2.2 Error Prone PCR (EPPCR)

Error Prone PCR (Figure 1.16) was originally studied using polymerases naturally lacking the 3' → 5' exonuclease activity required for genomic 'proofreading' such as *Thermus aquaticus* (*Taq*) DNA polymerase [84]. This leads to error rates of 0.001-0.02% errors per nucleotide per pass (resulting in 1 - 20 mutations per kb). The error rate achieved from *Taq* alone is rarely sufficient for the construction of highly randomised libraries and is often allied to the replacement of Mg^{2+} in the PCR

reaction with Mn^{2+} [85], the balance of Deoxyribonucleotide triphosphates (dNTPs) included in the PCR reaction for incorporation onto the growing chain [85], and additional nucleotide analogues which lead to various mis-incorporations [86]. This method requires very little prior knowledge of the structure-function relationship of the protein.

In vivo methods for the rapid generation of large random libraries using engineered mutator strains of *E. coli* (which lack DNA repair mechanisms) have been demonstrated to accumulate mutations at an enhanced rate [87]. Whilst not using PCR methods or mutagenic polymerases this method is random and attributable to errors in fidelity of replication.

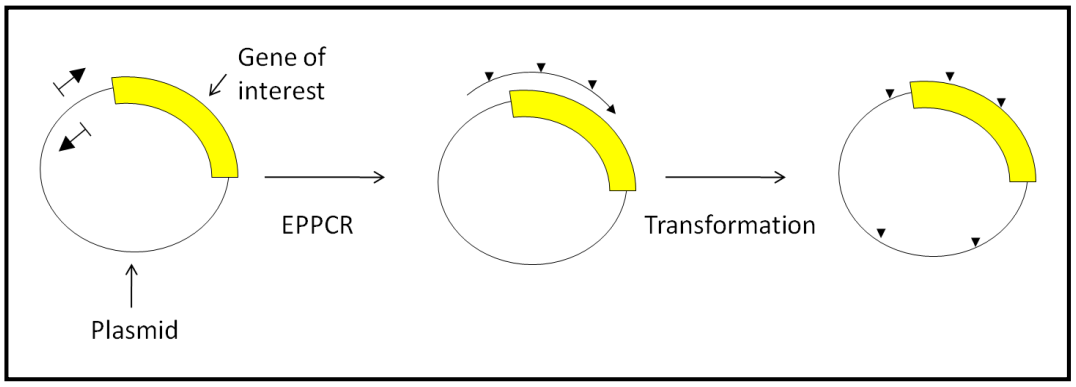


Figure 1.16. Scheme showing the principles behind error prone PCR methods. PCR of the plasmid or gene of interest under mutagenic conditions introduces mutations in the resulting product DNA. Small vertical triangles indicate the position of mutations in the EPPCR product.

The main disadvantage of these methods is that bias can be a significant problem. The fact that there are more residues distant from the active site dictates that a lot of mutants will have no effect on activity. Different polymerases vary in their tendency to mutate different nucleotides meaning some substitutions are represented more frequently than others. It is much more likely that some codons will be mutated to a selected subset of other codons, for example a single nucleotide change in a codon is more likely than all 3 changed therefore, AAA is more likely to be mutated to AAC than CCC. Mutations early in the PCR reaction are often present in a larger fraction of the products because they are templates for further cycles [81]. Despite the drawbacks and limitations of this approach even in

its simplest form EPPCR has led to improved biocatalysts, such as the improvement of subtilisin E in non-natural solvent [88].

Since the initial trials with EPPCR a number of improvements in the technique have to some extent overcome some of the limitations, such as the utilisation of multiple polymerases, each with distinctive mutational spectrum, in each reaction to offset mutational bias [89]. In order to localise the region where error prone PCR is carried out, thus limiting mutations to the gene of interest negating potential problems associated with whole plasmid methods, megaprimer methods (Figure 1.17) were developed where the mutagenic PCR is used to generate a mutagenic megaprimer which replicates the rest of the plasmid in conjunction with a regular high fidelity polymerase [90]. Where random libraries are utilised a number of possible mutations are incredibly rare by change (such as the mutation of 2 nucleotides in one codon), even in the absence of this type of bias the number of possible permutations containing all 20 amino acids in each position quickly become experimentally inaccessible, this means that only low levels of coverage are possible experimentally.

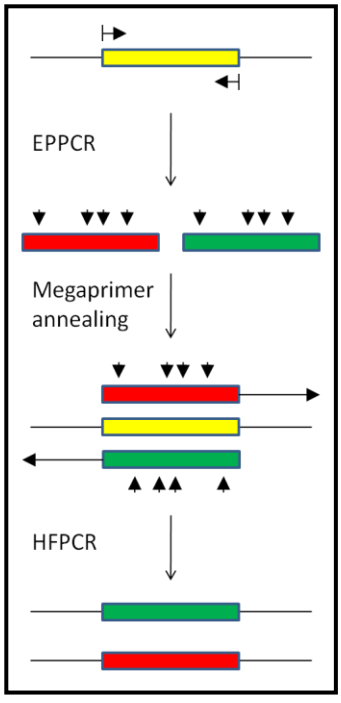


Figure 1.17. Generalised Scheme showing the principles behind megaprimer PCR library generation, The gene is amplified by PCR under mutagenic conditions, the resulting DNA sequence is used as a megaprimer for the high fidelity PCR of the rest of the plasmid.

An interesting extension of the idea of EPPCR is the idea of neutral drift, which while applicable to any form of mutagenesis seems most amenable to use with EPPCR methods. Neutral drift posits that by pre screening libraries and selecting clones that maintain wild type (WT) activity, you are selecting mutants which have demonstrated correct folding and catalytic competency. This leads to the generation of enhanced libraries which are more likely to yield functional mutants [91].

1.5.2.3 Site directed mutagenesis

Site directed mutagenesis is a PCR based method where mutagenic primers are incorporated into the gene at specific locations. This method is often used for the generation of rationally designed mutants using oligonucleotides of known sequence, but the use of degenerate oligonucleotide primers at specific positions can be used for the generation of mutagenic libraries [81].

In applying this range of techniques, a wide range of strategies have been employed to generate libraries both at single positions and multiple positions throughout the gene of interest (Figure 1.18). These approaches can require the screening of much smaller libraries and can approach complete coverage in certain cases. However the amount of sequence space sampled is extremely small compared to other approaches and to increase the likelihood of positive results, a substantial body of knowledge highlighting mechanistically or structurally important residues is needed to select appropriate residues to target.

The simplest application of this phenomenon to mutagenesis is the generation of single site saturation libraries [92]. These involve the substitution of all (or a number of) amino acids at a single position, the selection of which is generally guided by knowledge of structure function relationships. These single site libraries require smaller screening efforts to ensure a large degree of coverage, however also access corresponding lower sequence space [92].

An important consideration for making saturation type libraries is the nucleotide degeneracy used. By incorporation of a subset of different nucleotides at certain

positions one can limit the number of clones to be screened to give a high level of coverage. For example in NNN (where N represents any nucleotide) degeneracy, all 64 possible combinations of nucleotides are represented including all 20 amino acids, additionally NNK degeneracy (where K represents only Keto nucleotides i.e. T/G) also represents all 20 amino acids however only represents half the number of codons [93]. Use of the NNM degeneracy (when M represents aMino nucleotides) also represents half the number of possible codons, however, only represents 18 amino acids and includes 2 termination codons as opposed to 1 with NNK. This example represents the first level of minimisation with further reduced alphabets allowing higher degrees of coverage whilst simultaneously sacrificing sequence space sampled [94, 95].

Following initial site specific libraries being screened, hits can be subjected to multiple rounds of mutagenesis and screening at other positions, known as Iterative Saturation Mutagenesis (ISM) [96]. The screening of libraries in a different order leads to the possibility of several different pathways to the same mutant e.g. XAX/XBX could be accessed by screening library B with XAX as a template or screening library A with XBX as a template. It has been demonstrated that the relationship between all pathways leading to the same mutant are complex and all pathways do not necessarily lead to the same mutant, leading to a whole new layer of complexity added to the experiment [97].

Combinatorial active site saturation testing (CASTing) is a technique developed to maintain reasonably small libraries whilst evolving the active site of the enzyme to accept non wild type substrates [98]. In this method, two residues known to be important in catalysis and/or substrate binding are randomised to saturation. As combinations of positive mutations in more than one position may be additive, results from Combinatorial Active Site Saturation Testing (CASTing) experiments on multiple pairs of randomised residues could hopefully generate improved biocatalysts at a faster rate than screening single site saturation libraries [98]. The advantage of CASTing is that active site residues can be systematically screened whilst maintaining a more manageable library of only about 3,000 mutants (when compared to 160,000 mutants for 4 residues) per mutational round [98]. Since the

pairs of residues are located near each other, there is also the potential to find some improvements for which some degree of synergy between residues is required. This is unlikely in random mutagenesis or single site saturation mutagenesis. The repetition of library generation at novel positions using the leading candidates as templates has also been demonstrated and described as iterative casting [99].

Extending the idea of screening multiple positions at the same time leads to the multiple mutation reaction (MMR) method of generating libraries which uses multiple primers to introduce up to 9 distinct mutations on separate oligonucleotides and theoretically more mutations in a single reaction [100]. A further extension uses long degenerate oligonucleotides to randomise large cassettes within the genome whose positions are selected according to structure function information, and which has been demonstrated to be superior in this case to EPPCR methods [101].

The limitations of EPPCR with respect to saturation mutagenesis are demonstrated dramatically in the study of subtilisin, where the saturation mutagenesis of amino acids lead to superior increases in thermal stability [92]. It is important to note however that these amino acids were identified from EPPCR libraries demonstrating that both types of library generation can be accommodated in a successful strategy.

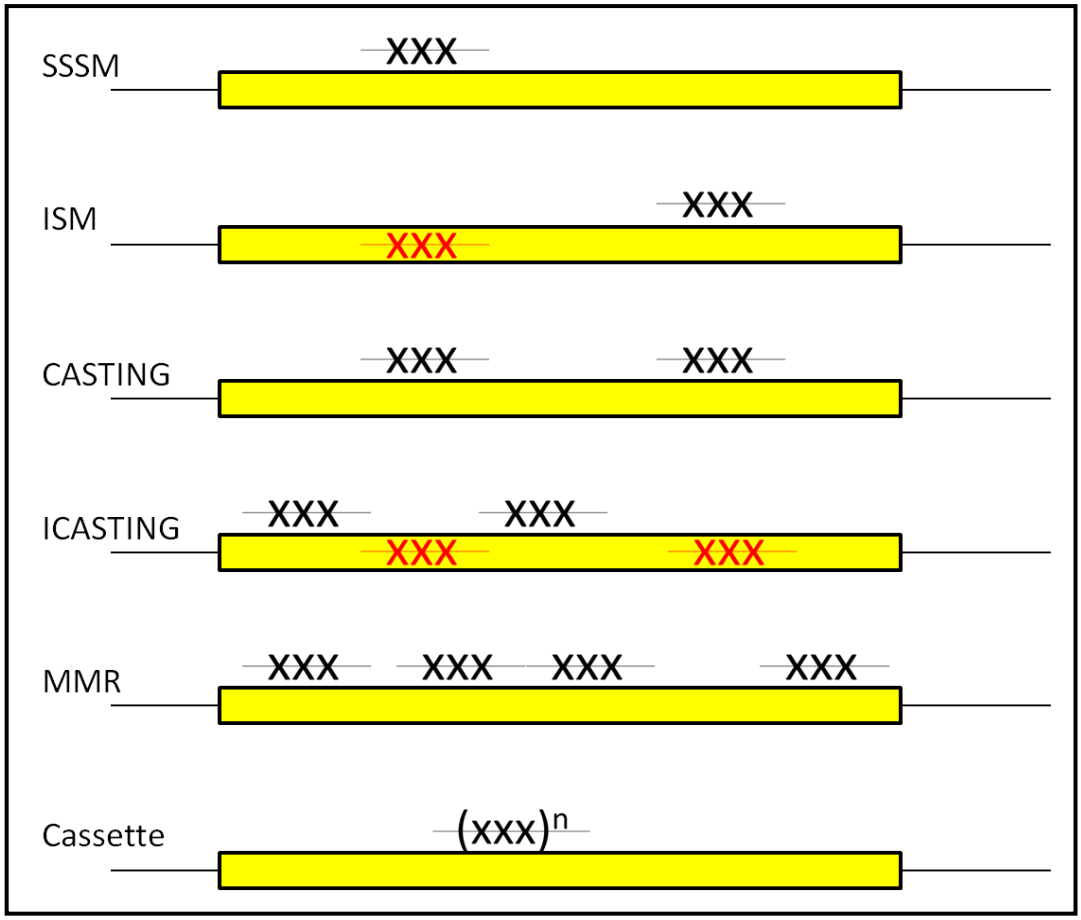


Figure 1.18. Generalised Scheme showing a variety of site directed mutagenesis methods of library generation. PCR of the gene of interest is performed with degenerate oligonucleotides using high fidelity PCR (HFPCR). SSSM, single site saturation mutagenesis; ISM, iterative saturation mutagenesis; CASTing, combinatorial active site saturation testing; ICASTing, iterative combinatorial active site saturation testing; MMR, multiple mutagenesis reaction.

1.5.2.4 Library generation statistics

The most important factor to consider when calculating the statistics describing mutagenic library construction is the equation outlined below (Equation 1.1, derived from first principles in [102] and [95])

Equation 1.1. Completeness of mutagenic libraries. P_c, the probability of a level of completeness; V, maximum degeneracy of the library.

This equation is compiled in the programme CASTER (an excel worksheet produced by the Reetz group) [103]. It is important to note however that all these

mathematical treatments make the critical assumption that all possible transformations are likely, and that no biological factors are taken into account. This simplification is required for the applicability of the treatments, however is not always accurate or representative of the true situation [102].

1.5.2.5 Summary

Throughout the planning of a protein engineering experiment it is important always to consider that you are balancing the coverage/completeness of the library you will be screening against the sequence space you are exploring. The extreme examples are demonstrated by SSSM libraries where experimentalists can get close to complete coverage of a library which represents a tiny proportion of the available sequence space. At the opposite extreme full gene random mutagenesis approaches allow you to sample a massive amount of sequence space, however a vast screening effort is required and it is impossible to approach complete coverage. Several factors which influence the decision of which library generation approach to use are summarised in Table 1.4. A consideration of when certain methods might be appropriate is given in [104]. Recent trends in the field show an increase in experiments which merge the random approaches with the focussed approaches and lead to the generation of smart focussed libraries.

Library generation method		Sequence Space Sampled	Screening Effort Required	Specificity
EPPCR		Very High	Very High	None
Site Directed	Single Site Saturation Mutagenesis (SSSM)	Very Low	Very Low	Very High
	Iterative Saturation Mutagenesis (ISM)	Low	Very Low	Very High
	CASTing	Low	Low	High
	Multiple Mutagenesis	Medium	Medium/High	Medium
	Cassette Mutagenesis	Medium/High	High	Low

Table 1.4. Summary of library generation methods including estimates of a number of factors to be considered when designing an experiment.

1.5.3 Library screening

Following library generation, it is necessary to identify those mutants which have acquired novel or improved functionalities. This step mimics the application of selective pressure present during natural macro-scale evolution, and can be achieved by selective or screening methods. Selective methods work on the principle that the desired property of the enzyme can provide an advantage/disadvantage to survival under certain conditions. This situation can occur naturally or under certain conditions be artificially engineered. Screening methods rely on monitoring a property which is proportional to the activity of the enzyme, due to the nature of the experiments a lot of experiments are performed with few improvements.

1.5.3.1 Selection

Selection has been demonstrated to be an excellent method for the isolation of clones grown under conditions lethal or growth inhibitory to WT constructs. Only clones which reduce the growth inhibitory phenotype are present. This enables very large libraries, with library sizes up to 10^{13} quoted [105], to be assessed easily

as only a small fraction of clones grow, growth being indicative of improved mutants.

In vivo selection techniques usually consist of auxotrophic host strains transformed with library DNA, in the case of a functional mutant the deficiency in the strain is complemented enabling survival. This method is generally applicable to modulating the specificity of enzymes and without modification is restricted to situations where substrates/products are essential/toxic, for example the rescue of the pyruvate auxotrophic *E. coli* PB25 using 2-Keto-3-deoxy-6-phosphogluconate aldolase mutant libraries [106]. Following careful consideration these approaches can be used in some cases to engineer other properties such as improvements in enantioselectivity, for example the linkage of the desired enantiomer to a metabolic precursor (positive selection) and the alternate enantiomer to a toxin (negative selection) [107]. This approach has been extended further by the development of three hybrid systems where the substrate is responsible for holding together a protein complex responsible for activation of transcription. Activity is then monitored using the expression of a reporter gene, for example the linkage between cephalosporinase activity and the expression of a *lacZ* reporter gene, allowing blue white screening [108, 109]. Variations in the transcription factor and reporter gene can lead to tailored selective techniques.

In vitro selection techniques have largely been exploited by the modification of enzyme substrate binding with the intention of changing the enzyme's affinity towards substrate, transition state, inhibitors and product, thus modifying catalysis. A number of display technologies, whereby folded proteins are expressed on the external surface of the organism and as such are tethered to its genetic material, have been demonstrated using phage [110], bacteria [111] and ribosomes [112]. Examples include the expression and immobilisation of membrane anchored Endoxylanase I mutants [113]. With modification this type of protocol can also be used to directly assess catalysis, by displaying both the enzyme and substrate on the cellular surface, catalysis can be performed in situ and variants can be selected using an antibody with affinity for the product [114].

1.5.3.2 Screening

Screening approaches are by far the most common method of applying selective pressure to mutant libraries in order to isolate improved variants [115]. They cover a multitude of approaches ranging from agar plate screening at the low end of the throughput spectrum, to approaches such as fluorescence activated cell sorting (FACS) which are at the limits of what can currently be achieved by screening approaches [115].

Agar plate screening involves the incubation of libraries on agar plates containing the substrate of choice. An important consideration is that the substrate/product/by-product/cofactor of the reaction is detectable. Common approaches include the use of colorimetric substrates, for example evolving β-fucosidase activity towards nitrophenyl-β-D-fucopyranoside substrate [116]; coloured secondary substrates, for example monitoring the uptake of safranin O in response to substrate turnover [117]; pH indicators, for example detection of the release of hydrofluoric acid during glycosynthase reactions using methyl red [118] or by degrading insoluble components of the media leading to clear zones, for example the hydrolysis of agar by agarase [119]. Additional external selective conditions can also be applied to agar plates during growth allowing the assay or multiple conditions such as temperature and pH [120].

Microtitre based methods are widely used in a number of different types of experiments but also have implications for the screening of mutagenic libraries. Individual library clones are grown in microplates followed by optional purification steps, before either being monitored spectrophotometerically, or by alternate methods such as gas chromatography (GC). A vast number of possible detection methods have been demonstrated for estimating the conversion of substrate to product including colour/fluorometric reaction participants, for example the characteristic absorbance of NADPH and indigo [121]; colour/fluorometric secondary metabolites, for example by the liberation of fluorophores from the secondary enzymatic reaction with endocellulase [122]; pH change, for example the use of 4-nitrophenol to screen for hydrolase activity [123]; temperature, for

example the monitoring using thermistors of temperature change associated with β-lactamase activity [124]; metal ion concentration, for example monitoring the Cu^{2+} ion concentration using a metal chelating compound with an orange absorbance, when the metal is chelated by the reaction product an orange absorbance is detected [125] and HPLC methods, for example monitoring turnover of the carbon bond forming reactions of transketolases [126].

A number of methods have also been described for the estimation of the enantiomeric excess of the product by including fluorometric methods, by calculating the rates of separate reactions both leading to the release of a fluorescent product [127] or mass spectrometry, for example using an isotopically labelled enantiomer [128]. Additionally, by relying on the principle of microscopic reversibility, often the reverse reaction can be screened with the chiral product used as substrate for the reverse reaction [129].

A vast quantity of assay methods have been published which are designed or adaptable for use in a microplate medium throughput format ranging in complexity and applicability (as discussed above). Microplate assays provide an almost ubiquitous option for the assay of mutagenic libraries. However, these provide far inferior throughput when compared to other techniques.

Several technologies have been exploited to increase the throughput of microplate screening experiments including the use of progressively smaller wells including 384 and 1,536 microplates and readers, all the way down to performing the reactions in 1,000,000 well plates [130]. Other groups have demonstrated that mutants can be pooled and hits can be sequentially plated and identified, multiplying the screening effort several fold [131].

FACS is the only screening procedure which can compete with selective methods in the throughput stakes with library sizes > 10^8 often quoted [132]. This approach requires the generation and localisation of a fluorescent reporter, for example membrane immobilisation [132]. Cells are then sorted on an individual basis by the presence or absence of fluorescence by specialised equipment allowing the identification of positive hits.

1.5.3.3 Summary

In general a selective approach is far superior to a screening approach as it allows the sorting of huge numbers of transformants to find the few positive mutants with minimal experimental effort. However at the higher throughput end of screening experiments, the number of transformants which can be screened is increasing dramatically. However the activity with many substrates does not lead to a selectable phenotype, meaning a selective approach is not possible. In some cases the selective phenotype can be engineered, however in a lot of cases a screening approach is the only one available.

1.5.4 Example protein properties engineered

A large number of protein properties have been shown to be evolvable, detailed in the following passage are a number of examples which have been reported previously. Whilst considering these examples an attempt has been made to provide examples of a variety of approaches with respect to library generation and assay method.

1.5.4.1 Thermostability

The generation of thermostable variants is a common theme for directed evolution. Enzymes evolved for an increase in the thermostability often demonstrate an increased stability in organic solvents [133] or extremes of pH [120], which is particularly of interest for biocatalytic purposes, allowing the utility of enzymes under a wider range of conditions.

Experiments associated with increasing thermal stability tend to adopt a whole genome approach to randomisation [119, 120, 133] since there is no 'active site' associated with thermal stability and residues affecting this property are located throughout the enzyme structure. However, in cases where crystallography data is available, experimentally derived B factors, which represent the degree of motion in the crystallographic model, can be used to guide targeted mutagenesis experiments. The programme B-fitter has been developed to identify potential hotspots for thermal stability [134].

Potential assay methods used include the application of heat to catalytic reactions, for example by application of heat to agar plates [120] and the direct monitoring of protein unfolding by a number of methods, for example circular dichroism [135].

1.5.4.2 Thermolability

Under certain circumstances, for example where the functional inactivation of an enzyme is required, it can be advantageous to engineer a decrease in protein stability into an enzyme. Whilst these examples are less frequent, previous experiments have included whole gene mutagenesis approaches, such as in the evolution of thermolabile transglutaminase [136], and more targeted approaches directed by structural knowledge, for example in the evolution of thermolabile lipases [137].

1.5.4.3 Solvent Stability

The engineering of increased stability of enzyme in a variety of solvents is intended to allow access to a greater range of conditions for biocatalysis with respect to solvent profile, since it is common for industrial reactions to use organic solvents [88]. Enzymes have largely evolved to function in an aqueous environment and a substantial subset of enzymes requires water for function. It is therefore not surprising that enzymes suffer from instability in the presence of organic solvents.

As with thermal stability there is no 'active site' for solvent stability, hence a large number of the experiments take a random approach to library generation, for example in the evolution of subtilisin E to function in dimethylformamide (DMF) [88, 133, 138, 139]. It has been noted that the majority of results from these experiments, predictably, provide results with mutants in solvent exposed areas. However, there is no immediately obvious mechanism for isolating these positions in library generation [139].

1.5.4.4 pH tolerance

Another condition limiting the utility of enzymes for industrial biotransformation is the often restrictive pH profiles [120]. The evolution of catalysts to function under

novel conditions of pH has also been described [120, 140, 141]. These experiments were largely carried out using random mutagenesis methods, for example using DNA shuffling methods to evolve xylanase B to function under more alkaline conditions [120, 140], however structure guided approaches were also used where appropriate, such as the use of position scores in the evolution of *Sphingomonas capsulata* prolyl endopeptidases (SCPEP) to function under gastric conditions [140, 141]. This property also provides an interesting aside giving an example of the evolution of a protein to increase its functionality as a therapeutic [141].

1.5.4.5 Increase native activity

Protein engineering approaches have been used frequently for the improvement of native activity i.e. the improvement of binding properties or the improvement of catalytic rate. These experiments are frequently performed using structural data to guide the construction of far more focussed libraries with reduced screening effort. Examples are available of random mutagenesis, for example in the improvement of the horse heart myoglobin peroxidase activity [142], a combination of random and saturation mutagenesis, for example in the improvement of benzoylformate decarboxylase carboligase activity [143] and more rational approaches, for example the improvement of pyranose-2-oxidase against a range of substrates involved in biofuel production [144]. In some cases this involves modifying the balance between reaction mechanisms to favour what is normally a by-product, as in the case of benzoylformate decarboxylase [143].

1.5.4.6 Increase activity towards poor substrate

One of the most common aims of protein engineering experiments is to widen the substrate acceptance of enzymes to accommodate novel substrates. A variety of methods have been used to generate libraries for changing substrate acceptance ranging from completely random approaches, for example in the evolution of adenosine diphosphate (ADP) glucose pyrophosphorylase to accommodate novel sugar substrates [145] to site directed approaches, for example in the evolution of transketolase to accommodate an aldehyde [146] with combinations of both methods also proving effective, for example in the evolution of P450 2A6 to

accommodate large indoles [147]. This represents by far the most popular use for enzyme evolution currently.

1.5.4.7 Variation in Stereo/Enantioselectivity

There have been many studies carried out on screening procedures monitoring the improvement of stereo and enantio selectivity of enzyme reactions. These methods require the detection of a chiral product or the monitoring of a chiral substrate. These experiments are often performed by generating diversity at specific positions identified as mechanistically important from structure/functional data, for example the modification of the enantiopreference of OYE1 [148]. Often less directed approaches are used to garner structural information which then targets positions to become the subject of further library generation, for example in the generation of stereocomplementary variants of N-acetylneuraminic acid lyase [149]. Sometimes more random approaches alone have been demonstrated to be efficient, for example in the reversal of enantiopreference in tagatose-1,6-bisphosphate aldolase [150].

One of the major problems encountered with this type of experiment is the development of suitable high throughput screens to distinguish between both substrate, product(s) and to determine the enantiopurity of product. When investigating reactions with alternate chiral substrates, initial efforts require the separation of enantiomers or reactions with alternate enantiomers, for example the monitoring of rates with chiral compounds in the evolution of *Pseudomonas aeruginosa* lipase [151], the screening for chiral products is discussed previously in Section 1.5.3.2.

A number of studies have suggested reasons for the enantio preference of enzyme systems both from experimental results in specific systems, for example the binding orientation of various enantiomers [152], and from more general principles, for example the principle of mirror image active site as derived from crystal structure analysis [153].

1.5.4.8 Variation in cofactor specificity

Rational design approaches have also led to the reorientation of coenzyme specificity, for example the realignment of glutathione reductase cofactor specificity [71], such results could potentially lead to the ability to replace more expensive coenzymes such as NADPH with less expensive alternatives such as NADH.

2 Materials & methods

2.1 Materials

2.1.1 Chemicals & reagents

Oxygen free nitrogen (OFN) was purchased from BOC. Deoxyribonucleotide triphosphates (dNTPs) were purchased from Fermentas. Bromophenol blue sodium salt, ethanol (ETOH), ethylenediaminetetraacetic acid (EDTA), glacial acetic acid, glycerol, glycine, isopropanol, D-lactose, magnesium sulfate ($MgSO_4$), methanol, potassium chloride (KCl), potassium dihydrogen phosphate (KH_2PO_4), potassium hydroxide (KOH), potassium phosphate (K_2HPO_4), sodium chloride (NaCl), sodium dihydrogen phosphate (NaH_2PO_4), sodium hydroxide (NaOH), sodium phosphate (Na_2HPO_4) were all purchased from Fisher chemicals. Agar, ampicillin, carbenicillin, isopropyl β-D-1-thiogalactopyranoside (IPTG), tris base were purchased from Formedium. Quant-it dsDNA broad range (BR) reagent, quant-it dsDNA high sensitivity (HS) reagent, qubit BR/HS standard #1 (TE buffer, 0 $ng\mu l^{-1}$ λ dsDNA), qubit BR standard #2 (TE buffer, 100 $ng\mu l^{-1}$ λ dsDNA), qubit HS standard #2 (TE buffer, 10 $ng\mu l^{-1}$ λ dsDNA) were purchased from Invitrogen. Dithiothreitol (DTT), magnesium chloride ($MgCl_2$), NADPH, sodium dodecyl sulphate (SDS), D-sorbitol, were purchased from Melford. Oligonucleotide primers (as described in Appendix 1) were purchased from MWG. Complete mini Ethylenediaminetetraacetic acid (EDTA) free protease inhibitor cocktail tablets were purchased from Roche. Ammonium sulphate ($(NH_4)_2SO_4$), calcium chloride (CaCl), cinnamonitrile, coomassie brilliant blue R250, cyclohexanecarbonitrile, cyclohexanecarboxylic acid, CH, cyclohex-1-enecarbonitrile, cyclohex-1-enecarboxylic acid, cyclohex-2-enone, cyclopentanone, cyclopent-2-enone, dimethyl sulfoxide (DMSO), ethidium bromide (EtBr), hexamine cobalt chloride, hydrochloric acid (HCl), imidazole, manganese chloride ($MnCl_2$), 2-mercaptoethanol, 3-methylcyclohexanone, 3-methylcyclohex-2-enone, 2-methylcyclopentanone, 3-methylcyclopentanone, 2-methylcyclopent-2-enone, 3-methylcyclopent-2-enone, 2-methyl-3-phenylpropanal, methyl 3-phenylpropanoate, α-methyl-*trans*-cinnamaldehyde, methyl-*trans*-cinnamate, mineral oil, nickel sulphate ($NiSO_4$), (n-morpholino) ethane sulfonic acid (MES), PCR

grade water, 3-phenylpropanal, 3-phenylpropanenitrile, 3-phenylpropanoic acid, potassium acetate, sodium acetate, sodium cacodylate, *trans*-cinnamaldehyde, *trans*-cinnamic acid, triton x-100 were purchased from Sigma Aldrich. All chemicals were of analytical grade or equivalent purity where available.

2.1.2 Media

For the production of selective media antibiotic was dissolved to the correct concentration, filter sterilised using a 0.2 µM syringe driven filter and added to media following autoclaving and cooling (in the case of agar media before hardening). For standard growth and purification of plasmid DNA during molecular biology protocols all strains were grown in Luria Broth (LB) consisting of 10 gl^{-1} tryptone, 5 gl^{-1} yeast extract, 10 gl^{-1} NaCl (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving. In order to isolate individual bacterial colonies during molecular biology protocols plates of LB agar containing 10 gl^{-1} tryptone, 5 gl^{-1} yeast extract, 10 gl^{-1} NaCl, 15gl^{-1} agar (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving, which following addition of antibiotic were poured into Petri dishes. Super optimal broth with catabolite repression (SOC) media used during transformation protocols was made by the addition of 20 gl^{-1} glucose, 20 gl^{-1} tryptone, 5 gl^{-1} yeast extract, 0.5 gl^{-1} NaCl, 2.44 gl^{-1} $MgSO_4$, 0.186 gl^{-1} KCl (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving. Initial purifications of PETNR were performed using double yeast tryptone (DYT) broth made by the addition of 16 gl^{-1} tryptone, 10 gl^{-1} yeast extract, 5 gl^{-1} NaCl (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving. Auto induction library growths were performed using a range of auto induction media (AIM) including; DYT AIM containing 16 gl^{-1} tryptone, 10 gl^{-1} yeast extract, 3.3 gl^{-1} $(NH_4)_2SO_4$, 6.8 gl^{-1} KH_2PO_4, 7.1 gl^{-1} Na_2HPO_4, 0.5 gl^{-1} glucose, 2 gl^{-1} Lactose, 0.15 gl^{-1} $MgSO_4$ (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving; LB AIM containing 10 gl^{-1} tryptone, 5 gl^{-1} yeast extract, 3.3 gl^{-1} $(NH_4)_2SO_4$, 6.8 gl^{-1} KH_2PO_4, 7.1 gl^{-1} Na_2HPO_4, 0.5 gl^{-1} glucose, 2 gl^{-1} Lactose, 0.15 gl^{-1} $MgSO_4$ (premixed solid components available from

Formedium) dissolved in the required volume of ddH_2O before autoclaving. Super broth (SB) AIM containing 35 gl^{-1} tryptone, 20 gl^{-1} yeast extract, 3.3 gl^{-1} $(NH_4)_2SO_4$, 6.8 gl^{-1} KH_2PO_4, 7.1 gl^{-1} Na_2HPO_4, 0.5 gl^{-1} glucose, 2 gl^{-1} Lactose, 0.15 gl^{-1} $MgSO_4$ (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving; Terrific broth (TB) AIM containing 12 gl^{-1} tryptone, 24 gl^{-1} yeast extract, 3.3 gl^{-1} $(NH_4)_2SO_4$, 6.8 gl^{-1} KH_2PO_4, 7.1 gl^{-1} Na_2HPO_4, 0.5 gl^{-1} glucose, 2 gl^{-1} Lactose, 0.15 gl^{-1} $MgSO_4$ (premixed solid components available from Formedium) dissolved in the required volume of ddH_2O before autoclaving;

2.1.3 Strains

Escherichia coli strain AVB100 (Avidity) genotype - MC1061, *ara*Δ139, Δ(*ara-leu*)7696, Δ(*lac*)*l*74, *galU, galK, hsdR2,* (r_{K-}, m_{K+}), *mcrB1, rpsL* (Str^r), *birA* gene stably integrated into the chromosome - was used for the expression and *in vivo* biotinylation of biotin tagged proteins. *Escherichia coli* strain JM109 (Promega) genotype - *recA1, endA1, gyrA96, thi-1, hsdR17,* (r_{K-}, m_{K+}), *supE44, relA1,* Δ(*lac-proAB*), [F´ *traD36, proAB, lacI*^{q}*ZΔM15*] - was used for growth/expression of recombinant PETNR and mutagenic libraries. *Escherichia coli* strain XL1 Blue (Stratagene) genotype - *recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac* [F´ *proAB, lacI*^{q}*ZΔM15, Tn10* (Tet^r)] - was trialed for the expression of mutant libraries. *Escherichia coli* strain XL10 Gold (Stratagene), genotype - Tet^r, Δ(*mcrA*)183*,* Δ(*mcrCB-hsdSMR-mrr*)173*, endA1, supE44, thi-1, recA1, gyrA96, relA1, lac,* Hte [F´ *proAB, lacI*^{q}*ZΔM15,* Tn*10* (Tet^r) Amy (Cam^r)] - was used for the replication and purification of plasmid DNA during molecular biology protocols.

2.1.4 Plasmids

The PETNR was originally cloned into pBluescript SK+ (Stratagene) plasmid. All the studies described in this thesis are performed using this construct.

2.1.5 Markers

For analytical agarose gel electrophoresis of plasmid DNA the Generuler 1 kbp ladder from Fermentas was used. This ladder consisted of DNA fragments from 250 bp to 1000 bp in 250 bp increments, 1000 bp to 4000 bp in 500 bp increments and 5

kbp, 6 kbp, 8 kbp, 10 kbp fragments. For analytical agarose gel electrophoresis of plasmid DNA the 1 kbp ladder from NEB was also used. This ladder consisted of DNA fragments from 0.5 kbp to 2 kbp in 0.5 kbp increments, from 2 kbp to 6 kbp in 1 kbp and fragments at 8 kbp and 10 kbp. For analytical agarose gel electrophoresis of smaller DNA fragments the Generuler 100 bp Ladder from Fermentas was used. This ladder consisted of DNA fragments at from 100 bp to 1000 bp in 100 bp increments. For analytical agarose gel electrophoresis of smaller DNA fragments the 100 bp Ladder from NEB was also used. This ladder consisted of DNA fragments at from 100 bp to 1000 bp in 100 bp increments and fragments at 1.2 kbp, 1.5 kbp, 2 kbp and 3 kbp. For SDS-PAGE analysis of proteins Pageruler plus prestained protein ladder purchased from Fermentas was used. This ladder consists of proteins with molecular weights of 10 kDa, 17 kDa, 28 kDa, 36 kDa, 55 kDa, 72 kDa, 95 kDa, 130 kDa, 250 kDa. For SDS-PAGE analysis of proteins low range prestained protein ladder purchased from bio-rad was used. This ladder consists of proteins with molecular weights of 21 kDa, 29 kDa, 34 kDa, 50 kDa, 77 kDa, 103 kDa.

2.1.6 Enzymes

Phusion DNA polymerase was purchased from Finnzymes. Antarctic phosphatase (AP), *BsmI*, T4 polynucleotide kinase (PNK), quick ligase and *Taq* DNA polymerase were purchased from NEB. Benzonase and rLysozyme were purchased from Novagen. DNAase and lysozyme were purchased from Sigma Aldrich.

2.1.7 Consumables

Biomek aluminium foil seals were purchased from Beckman Coulter. Econopac 10 DG column and ready gel tris-HCl precast gels (0.375 M Tris-HCl pH 8.8, 2.6 % bis-acrylimide, 4 % stacking gel, 12 % separating gel) were purchased from Bio-Rad. Parafilm and petri dishes were purchased from Fisher. Polypropylene 96 well microplates and UV star 96 well microplates were purchased from Greiner. High and medium volume robotic tips were purchased from Hamilton. Cloverleaf sitting drop crystallography trays and crystallography loops were purchased from Hampton. Qubit assay tubes were purchased from Invitrogen. D-Tube dialyser mega 6-8 kDa MW cut off and HisMag magnetic beads were purchased from Novagen.

QIAprep spin column and QIAquick spin column were purchased from Qiagen. Tubes (1 ml, 15 ml, and 50 ml) were purchased from Sarstedt. Air permeable 96 well format seal, thin walled PCR tubes, 96 well deep well blocks, 96 well PCR plate and 96 well silicone sealing mat were purchased from Starlabs. Vivaspin 20 spin concentrators 10, 30 kDa MW cut off were purchased from Vivascience.

2.1.8 Equipment

DB wax analytical GC column was purchased from Agilent. PCR workstation UV sterile hood was purchased from Alphalabs. SX18MV stopped flow was purchased from Applied Photophysics. Allegra X22R centrifuge, avanti J26-XP centrifuge and centrifuge tubes (1 L, 500 ml, 50 ml) were purchased from Beckman Coulter. Anaerobic glove boxes were purchased from Belle Technologies. Mini protean tetra cell electrophoresis apparatus and powerpac 300 power supply were purchased from Biorad. Gen 5 plate reader was purchased from Biotek. UV-50 bio UV–visible scanning spectrophotometer was purchased from Cary. JB1 water bath was purchased from Grant. Magnetic block and ML-STAR robotics platform were purchased from Hamilton. Titramax 1000 incubator was purchased from Heidolph. Quartz cuvettes (1 ml, 0.1 ml) were purchased from Hellma. HT incubator was purchased from Infors. Qubit fluorometer was purchased from Invitrogen. Accublock dry block heater was purchased from Labnet. StarPet E automatic multi-channel pipette was purchased from Starlabs. UV trans-illuminator was purchased from Syngene. TC 512 thermal cycler was purchased from Techne. French press, micro CL17 bench top centrifuge and owl B1 electrophoresis apparatus were purchased from Thermo Scientific. CP-3800 gas chromatograph with a combiPAL auto sampler was purchased from Varian.

2.1.9 Chromatography

Q-Sepharose resin (25 ml) was purchased from GE Healthcare. Immobilised monomeric avidin (10ml) was purchased from Pierce. Mimetic orange resin (25 ml) was purchased from Prometic Biosciences. Nickel-nitrilotriacetic acid (Ni - NTA) resin (25 ml) was purchased from Qiagen. Liquid chromatography column (18 ml), luer locked, non jacketed was purchased from Sigma Aldrich.

2.1.10 Buffers

2.1.10.1 Gel buffers

For the analysis of proteins by SDS PAGE; SDS loading buffer (2 ×; 62.5 mM TRIS-HCl, 2 % SDS, 25 % glycerol, 0.01 % bromophenol blue, 5 % 2-mercaptoethanol pH 6.8), SDS running buffer (25 mM TRIS-HCl, 192 mM glycine, 0.1 % SDS, pH 8.3 (unadjusted)), coomassie stain (0.25 % (w/v) Brilliant blue R250, 50 % methanol (v/v), 10 % glacial acetic acid and 40 % water), and destain (50 % methanol (v/v), 10 % glacial acetic acid and 40 % water) were used.

For the analysis of DNA by agarose gel electrophoresis DNA gel loading buffer (6 ×; 6 × TAE, 30 % glycerol (v/v), 0.25 % bromophenol blue (w/v)), tris acetate EDTA (TAE) buffer (40 mM tris acetate, 1 mM EDTA, pH 8) and DNA gel running buffer (1 × TAE, 0.001 % EtBr) were used.

2.1.10.2 DNA manipulation buffers

For the preparation of plasmid DNA buffers P1, P2, N3, PB, PE, EB were used from the miniprep kit as prepared by Qiagen. For the purification of DNA from an agarose gel or a PCR reaction buffers QG, PE, PBI were used as prepared by Qiagen. For PCR reactions 5 × Phusion GC buffer as prepared by Finnzymes or 10 × Standard Taq Buffer as prepared by NEB were used. For plasmid annealing STE buffer (10 mM TRIS-HCl, 50 mM NaCl, 1 mM EDTA, pH 8) was used. For the ligation of DNA 1 × DNA Ligase buffer (66 mM TRIS-HCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 7.5 % polyethylene glycol 6000, pH 7.6) was used. For restriction digestion reactions NEB 2 buffer (10 mM TRIS-HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9) and NEB 3 buffer (50 mM TRIS-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7) were used. For transformation protocols standard transformation buffer (10 mM MES pH 6.3 with KOH), 100 mM KCl, 45 mM $MnCl_2$, 10 mM $CaCl_2$, 3 mM $HACoCl_3$) and Dimethyl sulfoxide (DMSO) and Dithiothreitol (DTT) (DND) (0.9 ml DMSO, 0.153 g DTT, 100 µl 1 M Potassium Acetate, pH 7.5 were used.

2.1.10.3 Mimetic orange purification buffers

For the purification of PETNR using mimetic orange mimetic orange loading buffer/ equilibration buffer (10 mM potassium phosphate pH 7), and mimetic orange elution buffer (50 mM potassium phosphate, 250 mM NaCl, pH 7) were used.

2.1.10.4 Q-Sepharose purification buffers

For the purification of PETNR using Q-Sepharose, Q-Sepharose loading buffer (20 mM TRIS-HCl pH 8), Q-Sepharose low salt elution buffer (20 mM TRIS-HCl, 50 mM NaCl, pH 8), Q-Sepharose high salt elution buffer (20 mM TRIS-HCl, 500 mM NaCl, pH 8) were used.

2.1.10.5 Ni - NTA purification buffers

For the purification of $PETNR_{His}$ using Ni - NTA agarose, Ni - NTA loading buffer (50 mM potassium phosphate, 300 mM NaCl, 10 mM imidazole pH 8), Ni - NTA Wash buffer 1 (50 mM potassium phosphate, 40 mM imidazole, pH 8), Ni - NTA wash buffer 2 (50 mM potassium phosphate, 40 mM imidazole, 0.5 M NaCl pH 8), Ni - NTA wash buffer (50 mM potassium phosphate, 300 mM NaCl, 20 mM imidazole, pH 8), Ni - NTA elution buffer 1 (50 mM potassium phosphate, 300 mM NaCl, 250 mM imidazole, pH 8), Ni - NTA elution buffer 2 (50 mM potassium phosphate, 300 mM NaCl, 500 mM imidazole pH 8) were used.

2.1.10.6 Monomeric avidin purification buffers

Phosphate Buffered Saline (PBS; 0.1 M sodium phosphate, 0.15 M sodium chloride pH 7), Biotin Elution buffer (2 mM D-biotin in PBS).

2.1.10.7 Crystallography buffers

PETNR mother liquor (100 mM sodium cacodylate pH 6.2, 100 mM sodium acetate, 19 – 24 % PEG 3000, 12 – 18 % isopropanol).

2.1.10.8 Robotics buffers

For the robotics protocols a number of buffers were used including PBS, 2 × glycerol stock buffer (20 % glycerol in PBS), magnetic bead storage buffer (20 mM

potassium phosphate, 150 mM NaCl, 0.1 % triton X-100, pH 7.0), magnetic bead charge buffer (50 mM $NiSO_4$), magnetic bead binding buffer (50 mM potassium phosphate, 0.3 M NaCl, 5 mM imidazole, 40 % (w/v) sorbitol, pH 7.9), magnetic bead strip buffer (20 mM potassium phosphate, 0.5 M NaCl, 100 mM EDTA, pH 7.9), magnetic bead lysis buffer (1 × bugbuster, 50 mM potassium phosphate, 100 μg ml^{-1} lysozyme, 5 U ml^{-1} Benzonase, EDTA free protease inhibitor tablets, 100 mM FMN, pH 7), magnetic bead equilibration buffer (5 ×) (1.5 M NaCl, 0.2 M imidazole, pH 7.9), magnetic bead wash buffer (50 mM potassium phosphate, 0.3 M NaCl, 40 mM imidazole, pH 7.9), magnetic bead elution buffer (50 mM potassium phosphate, 0.3 M NaCl, 250 mM imidazole, pH 7) and enzyme dilution/assay/kinetic buffer (50 mM potassium phosphate, pH 7).

2.2 Methods

2.2.1 General DNA methods

2.2.1.1 DNA agarose gel electrophoresis

Agarose gel electrophoresis for the separation of DNA fragments was performed using an Owl B1 DNA Gel apparatus. Gels were made by boiling 0.6 - 1.5 % (w/v) agarose in 1 × TAE buffer followed by the addition of 0.001 % Ethidium Bromide (EtBr) (v/v). The gel was poured into the casting apparatus and a comb was added to generate sample loading wells. Samples were prepared by DNA gel loading buffer to 1 ×, and were loaded into the wells. The gel was covered with 1 × TAE buffer and run at 100 v (constant voltage) for 45 - 60 min, dependent on the concentration of agarose and size of the gel.

2.2.1.2 Plasmid DNA purification (miniprep)

For the miniprep of plasmid DNA buffers P1, P2, N3, PB, PE, EB were used from the miniprep kit as prepared by Qiagen. *E. coli* containing the desired plasmid was grown in LB medium (5 ml) containing 100 µg ml^{-1} ampicillin at 37 ˚C with 200 rpm shaking using a glycerol stock as the inoculum. Cultures were harvested by centrifugation at 13,300 *g* and the pellet was resuspended in 250 µl of Buffer P1. Buffer P2 (250 µl), containing LyseBlue (1 µl ml^{-1}), was added and gently mixed to lyse the cells. Lysate was neutralised and protein precipitated by the addition of 350 µl of Buffer N3. The lysate was centrifuged at 13300 *g* for 10 min to pellet the precipitated cell debris and precipitated protein/SDS. The clarified supernatant was applied to the QIAprep spin column and centrifuged at 13300 *g* for 1 min to allow the DNA to bind and the supernatant was discarded. The column was washed with 0.5 ml PB buffer followed by 0.75 ml PE buffer with intermediate centrifugation steps at 13300 *g* for 1 min. An additional centrifugation step was included after supernatant removal to ensure complete removal of residual ethanol. Plasmid DNA was eluted into a sterile 1 ml tube with 50 µl EB buffer, allowed to stand for 1 m before being centrifuged at 13300 *g* for 1 min. Plasmid DNA was stored at -80 ˚c until required.

2.2.1.3 Gel purification

For the gel purification of plasmid DNA buffers QG, PE, EB were used from the gel purification kit as prepared by Qiagen.DNA bands of interest in agarose gels were excised using a scalpel whilst the gel was visualised using a UV trans-illuminator. The excised fragment was weighed and solubilised in 3 volumes (w/v) of buffer QG and incubated at 50 ˚C for 10 min until completely soluble. Isopropanol was added until equivalent in volume to the mass of the excised fragment and the solution was applied to a QIAquick column and centrifuged at 13300 *g* for 1 min. The column was washed with 500 µl QG buffer then 750 µl PE buffer with centrifugation at 13300 *g* for 1 min and supernatant removal after each step. An additional centrifugation step was included after supernatant removal to ensure complete removal of residual ethanol. The DNA was eluted into a sterile 1 ml tube with 50 µl EB buffer, allowed to stand for 1 min before being centrifuged at 13300 *g* for 1 min.

2.2.1.4 DNA quantitation (Qubit)

Plasmid DNA was quantified using a Qubit fluorometer in conjunction with the Qubit dsDNA BR or HS assay kit. The following method describes the procedure for the BR kit. The method for the more sensitive HS kit was identical except for the substitution of the appropriate reagents/buffers/standards. All reagents and buffers were equilibrated at room temperature before use. The 200 × fluorescent dye was diluted to 1 × in the dilution reagent and aliquoted into clear 500 µl reaction tubes to a volume of 200 µl minus DNA volume (1 – 20 µl). Standards (10 µl aliquots of 0 ng μl^{-1} and 100 ng μl^{-1}) were added to the calibration tubes and 1-20 µl of sample DNA was added to the remaining sample tubes. All tubes were mixed by vortexing before incubation at room temperature for 5 min. The Qubit Fluorometer was calibrated using the standard reactions, before reading the sample DNA Q values. Final dsDNA concentration was calculated using the equation below (Equation 2.1).

Equation 2.1. Calculation to derive the concentration of dsDNA using the Qubit quantitation method. Q is the experimentally derived parameter. Final concentration is in the same units as Q, which is included in the output from the fluorometer.

2.2.1.5 Sequencing reactions

Sequencing of experimentally derived DNA was performed by Eurofins MWG operon. Samples were processed using the value read tube service grade. Samples were prepared to a concentration of 75 ng μl^{-1} to a minimum volume of 15 µl with sterile water. Sequencing could be performed from standard or from custom synthesised oligonucleotides, details of which are included in Appendix 1. Results were aligned using the Clustal online tool (European Bioinformatics Institute, EBI).

2.2.1.6 Phusion PCR

PCR reactions were carried out in thin walled PCR tubes according to the standard reaction with Phusion DNA polymerase configuration detailed below (Table 2.1). Reactions were performed in a total volume of 50 µl using sterile PCR grade water. PCR reactions were performed using standard cycling parameters detailed below (Table 2.2) unless otherwise stated. Oligonucleotide primers were designed to have a T_m ~5 - 10 ˚C above the annealing temperature and were analysed for the presence of undesirable secondary structure using mfold (Integrated DNA technologies, IDT). Primers were synthesised by Eurofins MWG operon.

Component	Final concentration
5× Phusion GC buffer	1 ×
dNTP (dATP, dTTP, dGTP, dCTP)	200 µM (Each)
Primers	0.5 µM (Each)
PETNRHis template DNA	1 ng μl^{-1}
Phusion DNA Polymerase	0.02 U μl^{-1}
Water	Up to 50 µl

Fig 2.1. Composition of reaction mixture of standard Phusion PCR reactions.

Temperature (˚C)	Time (s)	Cycles (×)
98	120	1
98	10	20
72	90	
72	300	1

Table 2.2. Reaction cycles of standard Phusion PCR reactions.

2.2.1.7 Taq PCR

PCR reactions were carried out in thin walled PCR tubes according to the standard reaction with *Taq* DNA polymerase configuration detailed below (Table 2.3). Reactions were performed in a total volume of 50 µl using sterile PCR grade water. PCR reactions were performed using standard cycling parameters detailed below (Table 2.4) unless otherwise stated. Oligonucleotide primers were designed as discussed in section 2.2.1.6.

Component	Final concentration
10× standard *Taq* buffer	1 ×
dNTP (dATP, dTTP, dGTP, dCTP)	0.2 mM (Each dNTP)
Primers	0.2 µM (Each primer)
PETNR template DNA	1 µl lysate
Taq DNA Polymerase	0.025 U μl^{-1}

Table 2.3. Composition of reaction mixture of standard *Taq* PCR reactions.

Temperature (˚C)	Time (s)	Cycles (×)
94	120	1
94	15	
54	30	35
68	30	
68	120	1

Table 2.4. Reaction cycles of standard *Taq* PCR reactions.

2.2.1.8 Kod PCR

PCR reactions were carried out in thin walled PCR tubes according to the standard reaction with *Kod* DNA polymerase configuration detailed below (Table 2.5). Reactions were performed in a total volume of 50 µl using sterile PCR grade water. PCR reactions were performed using standard cycling parameters detailed below (Table 2.6) unless otherwise stated. Oligonucleotide primers were designed as discussed in section 2.2.1.6.

Component	Final concentration
10× *Kod* buffer	1 ×
$MgCl_2$	1.5 mM
dNTP (dATP, dTTP, dGTP, dCTP)	0.2 mM (Each dNTP)
Primers	0.4 µM (Each primer)
Template DNA	0.2 ng μl^{-1}
Kod DNA Polymerase	0.02 U μl^{-1}

Fig 2.5. Composition of reaction mixture of standard *Kod* PCR reactions.

Temperature (˚C)	Time (s)	Cycles (×)
95	120	1
95	20	
65	10	20
70	100	

Fig 2.6. Reaction cycles of standard *Kod* PCR reactions.

2.2.1.9 DpnI digestion

To increase the fraction of PCR products before transformation into *E. Coli,* template (miniprep) DNA was digested with DpnI at 0.4 U μl^{-1} for 1 h at 37 ˚C. The restriction enzyme *DpnI* is specific for methylated DNA and so is only able to cut paternal DNA not PCR product.

2.2.1.10 PCR purification

To facilitate the clean-up of PCR reactions not analysed by electrophoresis, the Qiagen PCR purification kit was used according to the manufacturer's protocol. The PCR product was added to 5 equivalent volumes of PBI buffer, and centrifuged in a QIAquick spin column at 13300 *g* for 1 min. The column was washed with 750 µl PE buffer, and the centrifugation step was repeated to remove residual ethanol. The DNA was then eluted into a sterile 1 ml tube with 50 µl EB, allowed to stand for 1 min before being centrifuged at 13300 *g* for 1 min.

2.2.1.11 Oligonucleotide annealing

Complementary pairs of oligonucleotides (400 nM each) were mixed in STE buffer and overlaid with sterile mineral oil to prevent evaporation. The pairs of oligonucleotides were re-annealed by heating to 94 °C in a dry block heater followed by slow cooling to room temperature over a 90 min period with the dry block removed from the heat.

2.2.1.12 Oligonucleotide phosphorylation

The pairs of primers to be annealed were phosphorylated separately using T4 Polynucleotide Kinase (PNK) at 37 ˚C for 30 min. The reaction mixture was composed of 1 × DNA Ligase buffer 120 nM of each of the annealed primers and 10 U PNK.

2.2.1.13 Plasmid digestion using BsmI

Plasmid was digested in a 50 µl reaction with the restriction enzyme *BsmI*. The reaction mixture was composed of 1 × NEB 2 buffer, ~180 µg ml^{-1} plasmid DNA and

10 U of *BsmI*. The mixture was covered with a layer of sterile mineral oil, to prevent evaporation, and incubated at 65 °C for 12 h.

2.2.1.14 Dephosphorylation of digested plasmid

The dephosphorylation reaction was catalysed by the enzyme Antarctic phosphatase (AP). The 50 µl reaction mixture was composed of 1 × NEB 3 buffer, ~ 100 µg ml^{-1} digested plasmid DNA and 10 U of AP. The reaction was incubated at 37 ˚C for 15 min.

2.2.1.15 DNA ligation

The ligation reaction was carried out according to the standard protocol for quick ligase. The 10 µl reaction mixture was composed of 1 × DNA Ligase buffer, ~50 ng digested plasmid DNA, and ~ 150-300 ng phosphorylated DNA insert and 1 µl of Quick Ligation enzyme. The reaction was incubated at room temperature for 5 min before placing on ice.

2.2.1.16 Generation of competent E. coli

An overnight starter culture growth (10 ml) in LB medium of the *E. coli* strain AVB100 was grown at 37 ˚C with shaking at 200 rpm. These were used as an inoculum for 100 ml LB growth, to OD_{550} = 0.5, under the same conditions. Cells were incubated on ice for 15 min before being centrifuged at 3900 *g* at 4 ˚c for 15 min and the supernatant discarded. The cell pellet was resuspended in 25 ml standard transformation buffer (TFB). Cells were centrifuged at 3900 *g* at 4 ˚C for 15 min and resuspended in 25 ml TFB 3 times, before finally being resuspended in 4 ml TFB and frozen at -80 ˚C.

2.2.1.17 Pre-treatment of competent cells

The protocol for the pre-treatment of competent cells varies according to cell type. For AVB 100 7 µl of DMSO and DTT (DND) was added to 200 µl of the pre-treated cells and incubated on ice for 10 min, a second addition of DND gives a final concentration of 7 % (v/v). Following this addition the cells were incubated for 20 min. In the case of XL10 Gold 4 µl of the included 2-mercaptoethanol mix was

added to 100 µl of cells, the mixture was then incubated on ice for 10 min with mixing every 2 min.

2.2.1.18 Transformation of competent cells

In the case of JM109 no pre-treatment was required. Following appropriate pre-treatment 50 ng of plasmid DNA was added to the aliquots of cells, the transformation reaction was mixed by flicking the tube gently. Reactions were incubated on ice for 10 min before a heat shock of 50 s at 42 ˚C, following heat shock transformations were incubated on ice for 2 min. 900 µl SOC was then added and the reactions were incubated at 37 ˚C, 225 rpm. Transformants were then plated at a range of concentrations on selective plates.

2.2.1.19 Small scale lysis method for Colony Pick PCR (CPPCR)

Individual colonies of clones of interest grown in selective LB agar were picked and inoculated into 5 ml of LB media containing 100 µg ml^{-1} ampicillin. These colonies were grown at 37 ˚C, 180 rpm for 12 h. Glycerol stocks were produced by combining 0.5 ml of the culture with an equal volume of glycerol stock buffer. Stocks were frozen on dry ice, and stored at -80 ˚C. The remaining cells were harvested by centrifugation at 1000 *g* for 1 min and the supernatant was discarded. The pellet was resuspended in 50 µl of sterile water and boiled for 5 min at 100 ˚C to lyse the cells. The extracts were centrifuged at 1000 *g* for 1 min and the supernatant was transferred to clean sterile reaction tubes for use in PCR reactions.

2.2.2 General protein methods

2.2.2.1 SDS PAGE

SDS PAGE analysis was performed using Bio-Rad ready gel Tris-HCl precast gels with standard SDS running buffer. Samples were prepared by addition of equal volumes of sample and 2 × SDS buffer and boiled at 95 ˚C for 5 min. The gel was run for 45 min at 180 V (constant voltage). Once running was complete the gel was removed from the tank and extracted from the casing. This was incubated with shaking in

coomassie stain for 3-4 h followed by incubation with destain until bands are visible.

2.2.3 General analytical methods

2.2.3.1 Anaerobic methods

2.2.3.1.1 Preparation of anaerobic buffers/reaction components

Anaerobic buffers were prepared by bubbling solutions with oxygen free nitrogen (OFN) at 2 bar for ~30 - 60 min. Buffers were placed inside the glove box and left exposed to the OFN (at < 5 ppm O_2) for at least 12 hours prior to use.

Substrates were dissolved in anaerobic solvents and/or buffer solutions within an anaerobic glove box.

2.2.3.1.2 Preparation of anaerobic enzyme solutions

Traces of oxygen in the enzyme solution were removed by passing oxidised enzyme solution through a Bio-Rad Econopac 10 DG column pre-equilibrated with anaerobic buffer. Yellow fractions were collected and the concentration of enzyme was determined using the extinction co-efficient of ε_{464} = 11,300 M^{-1} cm^{-1}.

2.2.3.1.3 Preparation of chemically reduced enzyme

In order to study the oxidative half reaction of PETNR in single turnover experiments pre-reduced enzyme is required. Reduced enzyme was generated from the reductive titration of PETNR in an anaerobic cuvette sealed with a size 13 Subaseal, sodium dithionite was added using a 25 µl Hamilton syringe, with the flavin absorbance monitored spectrophotometrically.

2.2.3.1.4 Anaerobic steady state/single turnover assays

Anaerobic assays were performed within an anaerobic glove box (Belle technology) to prevent the re-oxidation of the reduced enzyme by molecular oxygen. The anaerobic atmosphere (< 5 ppm oxygen) was maintained using an atmosphere of OFN combined with external gas cycling and oxygen removing catalysts.

2.2.3.2 Spectrophotometric steady state assays

Steady state analyses of enzyme variants were performed on a Cary UV-50 Bio UV–visible scanning spectrophotometer using a 1 ml quartz cuvette (Hellma) housed in a Belle technology anaerobic glove box. Reactions (1 ml) were composed of kinetic buffer pH 7 containing 0 - 1 mM NADPH, 0 - 5 µM PETNR and 0 - 50 mM oxidative substrate. The reaction was usually initiated by addition of the oxidative substrate and $A_{340/365}$ (characteristic absorbance of NADPH) was monitored continuously. Initial rates were determined using Cary WinUV software and fitted to the Michaelis Menten steady state kinetic equations (Equations 2.2 and 2.3) using Origin 7 software.

Equation 2.2. Michaelis Menten equation describing the relating of the initial rate of reaction with the concentration of substrate under standard conditions. v_0 represents the initial rate, [S] is the substrate concentration, v_{max} represents the maximal rate under a set of conditions and K_M is the Michaelis constant.

Equation 2.3. Michaelis Menten equation describing the relating of the initial rate of reaction with the concentration of substrate under conditions of substrate inhibition. v_0 represents the initial rate, [S] is the substrate concentration, v_{max} represents the maximal rate under a set of conditions, K_M is the Michaelis constant and K_i is the inhibition constant.

2.2.3.3 Stopped flow pre steady state assays

Data were recorded on an Applied Photophysics SX18MV stopped flow housed in a Belle technology anaerobic glove box. Concentrated (2 x) solutions of enzyme and substrate were added to the syringes of the stopped flow apparatus. Solutions were injected into the measurement cell by firing the pressurised drive syringe. Data were recorded using the single wavelength detection module set to 464 nm to monitor FMN reduction/oxidation. Exponential curves were fitted to a single exponential expression (Equation 2.4) using spectra kinetics software from Applied

Photophysics. These data were then plotted in Origin 7 and fitted to the Strickland equation (Equation 2.5, [154]).

Equation 2.4. Single exponential equation for determining k_{obs} experimentally derived absorbance changes. is the observed absorbance change, is the amplitude, is the observed rate and t is the time.

Equation 2.5. Strickland equation for the analysis of single turnover kinetic results. is the observed rate from equation 2.4, is the limiting rate constant for reduction and is the dissociation constant.

2.2.3.4 Plate reader kinetics

Plate reader steady-state analysis was performed in a Biotek Gen 5 plate reader, 96 well microtitre plates were used in the place of individual cuvettes. Reactions were monitored in the UV region of the spectrum (340 nm) therefore, UV transparent 96 well microplates were used (Greiner, UV-star). Wavelength was determined using the plate reader monochromator. For reactions monitored in the visible wavelength spectrum, standard polypropylene 96 well microplates were used (Greiner).

Reactions (300 µl volume) were composed of anaerobic kinetic buffer containing 0 - 1 mM NADPH, 0 - 5 µM PETNR and 0 - 50 mM substrate. The reaction was initiated by the addition of the oxidative substrate followed by a 10 cycles of mixing with the pipette at maximum available speed. After allowing 10 s for the surface of the well to settle following mixing, the kinetic transient in 1 s intervals was monitored at 340 nm to follow the oxidation of the cofactor. Initial rates were derived from numbers exported by Biotek Gen 5 software using excels 'SLOPE' function and fitted to the appropriate Michalis Menten steady state kinetic equations using Origin 7 (Equation 2.2 and 2.3).

2.2.3.5 GC reactions

2.2.3.5.1 Reaction conditions

Reactions (1 ml) were composed of assay buffer containing 5 % ethanol, 5 mM oxidative substrate, 6 mM NADPH and 2 µM enzyme. Reactions were prepared in 2 ml vials under anaerobic conditions and sealed with parafilm before being incubated at 37 ˚C for 4 - 24 h with shaking at 200 rpm.

2.2.3.5.2 Solvent extraction

Products were extracted in ethyl acetate (900 µl) containing 0.5 % (v/v) limonene, the latter acting as an internal standard for GC. The phases were mixed vigorously by inversion and vortexing repeatedly, followed by centrifugation at 13300 *g* for 5 min. This enabled the phases to separate and 700 µl of the solvent phase was removed and transferred to glass vials.

Substrates containing carboxylic acid functionalities required further treatment entailing acidification prior to extraction by the addition of 50 µl of 1 M HCl. Once extraction was completed methylation reaction was performed by the addition of 200 µl methanol and 10 µl 2 M trimethylsllyldiazomethane in hexane, followed by incubation in a fume hood at room temperature for 10 min. Excess catalyst was inactivated by addition of 5 µl glacial acetic acid.

2.2.3.5.3 Quantitative GC

Quantitative detection of products was performed on a Varian CP-3800 gas chromatograph with a CombiPAL auto sampler using a DB-Wax column (30 m, 0.25 mm). Samples were vaporized using a split injector in a ratio and loaded onto the capillary column using a stream of helium (BOC) carrier gas. Samples were analyzed using temperature methods as detailed in table 2.7 consisting of a pause at a lower start temperature, before a temperature ramping segment, followed by a pause at a higher finishing temperature. Eluted compounds were detected using the Flame Ionization Detector (FID).

Substrate and product peaks were processed using star chromatography workstation software (Varian) and identified by comparison with verified authentic standards. Conversion and yield values were calculated from integration of the

peaks according to equation 2.6, 2.7 and 2.8, which takes into account calibration curves of each substrate product mixture.

Equation 2.6. Calculations used to derive concentrations of substrate and product used for quantitative GC. Concentrations derived from these calculations were used to determine conversion (%) of substrate and yield (%) of product.

Equation 2.7. Calculation used to derive a value for conversion of substrate in biotransformation assays.

Equation 2.8. Calculation used to derive a value for yield of product in biotransformation assays.

2.2.3.5.4 Chiral GC

Quantification of the enantiomeric excess (e.e.) from biotransformation reactions was performed on a Varian CP-3800 gas chromatograph with a CombiPAL auto sampler using a CP Chirasil DEX CB column (25 m, 0.32 mm, 0.25 µm film). Samples were vaporized, run and analysed as described previously (Section 2.2.3.5.3). Peaks were integrated and identified by comparison with authentic chiral standards e.e. was calculated from equation 2.9.

Equation 2.9. Calculation used to determine enantiomeric excess of biotransformation assays.

Analyte	Split injector ratio (1:x)	Flow rate (ml $min^{-1)}$	Injector temp (˚C)	Start Temp (˚C)	Start Temp Hold (min)	End Temp (˚C)	End Temp Hold (min)	Temp ramping (˚C/min)	Detector temp (˚C)
2-Cyclohexenone	20	1	220	40	10	200	1	20	250
Trans-cinnamic acid	20	1.7	250	100	2	230	3	20	250
Cinnamonitrile	20	1	250	100	2	240	3	20	250
1-Cyclohexene-1-carboxylic acid	20	1	220	60	5	200	1	20	250
Cyclohexene-1-carbonitrile	20	1	220	60	5	200	1	20	250
Trans-cinnamaldehyde	20	1	250	100	2	200	2	20	250
α-Methyl-trans-cinnamaldehyde	20	2	250	100	2	220	2	20	250
α-Methyl-trans-cinnamaldehyde (Chiral)	100	3	250	90	30	180	1	20	250
methyl *trans*-cinnamate	20	1.7	250	100	2	200	2	20	250
3-Methyl-2-cyclohexenone	20	1	220	40	10	200	1	20	250
2-Methyl-2-cyclopentenone	20	1	220	70	10	200	2	20	250
2-Methyl-2-cyclopentenone (Chiral)	100	2.5	250	50	25	180	1	20	250
3-Methyl-2-cyclopentenone	20	1	220	40	10	200	3	15	250

Table 2.7. Temperature methods used to analyze conversion and yield of biotransformation reactions with mutant PETNR. Chiral methods are highlighted in grey.

2.2.3.6 Protein crystallography

2.2.3.6.1 Crystal growth

PETNR crystals were grown using the sitting drop method in crystallography trays. The enzyme crystallised in drops composed of 2 µl enzyme solution (7 mg ml^{-1}) and 2 µl mother liquor with 1 ml of the latter reagents located in the central wells of the tray. Drops were covered with crystallography tape and incubated at 20 ˚C for 2 - 5 days. Crystals were mounted in standard crystallography loops in the absence of additional cryoprotectant and frozen in liquid nitrogen

2.2.3.6.2 Collection of diffraction data

A full $PETNR_{His}$ (1.4 Å) x-ray diffraction dataset was collected from a single crystal at the European Synchrotron Radiation Facility (ESRF; Grenoble, France) on station ID 14.4 (wavelength 1.07 Å; 100K) using an ADSC CCD detector.

2.2.3.6.3 Model generation

Crystallographic data sets were processed and scaled using the programmes MOSFLM [155] and Scala [156]. The structures were solved via molecular replacement using the coordinates for the acetate bound PETNR reductase (PDB 1H50 [40]). Model rebuilding and water addition was performed automatically using REFMAC5 [156] combined with ARP/warp [157]. Positional and anisotropic B-factor refinement was performed using REFMAC5 [156], (hydrogens included in the refinement) with alternate rounds of manual rebuilding of the model in coot [158].

2.2.4 Thesis specific methods

2.2.4.1 Wild type PETNR ($PETNR_{WT}$) purification

PETNR was purified from the *E. coli* strain JM109 expressing the PETNR gene from a pBluescriptII SK+ plasmid under the control of an IPTG inducible promoter. Starter cultures (5 ml then 100 ml LB) containing 100 µg ml^{-1} ampicillin were grown from a glycerol stock at 37 ˚C mixing at 180 rpm for 12 h. Larger scale cultures (1 – 12 L) of Double yeast tryptone (DYT) medium containing 100 µg ml^{-1} ampicillin were inoculated from the starters and grown at 37 °C mixing at 180 rpm until they

reached an OD_{600} of 0.7. Protein expression was induced by IPTG addition to a final concentration of 0.4 mM followed by growth at 37 ˚C mixing at 180 rpm for 12 h. Alternatively cells were grown at 37 ˚C mixing at 180 rpm for 18 h in TBAIM containing 100 μg ml^{-1} ampicillin.

Cells were harvested by centrifugation at 7,000 *g* at 4 ˚C for 20 min. The pellet was re-suspended in Mimetic Orange loading buffer (50 mM potassium phosphate pH 7) to a concentration of 1 g ml^{-1}. A small amount of lysozyme and DNAase were added to the cell slurry to initiate cell lysis and DNA degradation, the latter of which was essential for decreasing the viscosity of the slurry. An EDTA free protease inhibitor tablet (Roche) containing a cocktail of protease inhibitors was added to prevent degradation of the expressed protein by cellular proteases. Further cell lysis was achieved by passing the enzyme repeatedly through a french press at 4 ˚C. The cell lysate was spun at 20,000 *g* at 4 ˚C for 20 min, and the supernatant was retained. The extract was dialysed into 12 L Mimetic Orange loading buffer using 10 kDa MW cut off dialysis tubing at 4 ˚C.

The dialysed extract was applied to a 25 ml mimetic orange column, pre-equilibrated with 500 ml Mimetic Orange loading buffer. The column was washed with 500 ml equilibration buffer followed by enzyme elution in Mimetic Orange elution buffer (50 mM potassium phosphate, 250 mM NaCl, pH 7). PETNR containing fractions were pooled and dialysed using 12 L Q-Sepharose loading buffer (20 mM TRIS-HCl buffer pH 8) using 10 kDa MW cut off dialysis tubing at 4˚C. Dialysed eluates were concentrated on a Vivaspin 20 (Vivascience) spin filter with a 30 kDa MW cut off at 3000 *g* at 4 ˚C.

PETNR was further purified by application to a 25 ml Q-Sepharose column pre-equilibrated with 500 ml Q-Sepharose loading buffer. The column was washed with 500 ml loading buffer followed by enzyme elution using a gradient of 50 to 500 mM NaCl in 20 mM TRIS-HCl buffer pH 8.

Purified PETNR was defined as having an A_{280}:A_{464} ≤ 10 as well as a single band on SDS PAGE. Active fractions were pooled and dialysed into kinetic buffer and concentrated using 10 kDa MW cut off vivaspin concentrator at 3000 *g* at 4 ˚C.

Protein (~14 mg ml^{-1}) was frozen in small aliquots in liquid nitrogen and stored at -80 ˚C.

2.2.4.2 His tagged PETNR ($PETNR_{His}$) generation

Due to the nature of the PETNR construct, enzyme was expressed at very high levels (~50 mg L^{-1}). In order to maintain the high levels of expression the His tag was added in the same vector to the 3' end of the gene. Addition of a His tag required the insertion of 15 bp, CATCATCACCATCACCAT, to the gene in a multi-step PCR method. The mutation was located at the 5' or the 3' end of the two non-complementary primers which anneal flanking the insertion site, just before the $PETNR_{WT}$ STOP codon. The sequence to be included was included as a tail at opposite ends of each primer. This resulted in the generation of the tagged enzyme construct with staggered nicks. A graphical summary of the first PCR reaction is shown in Fig 2.1.

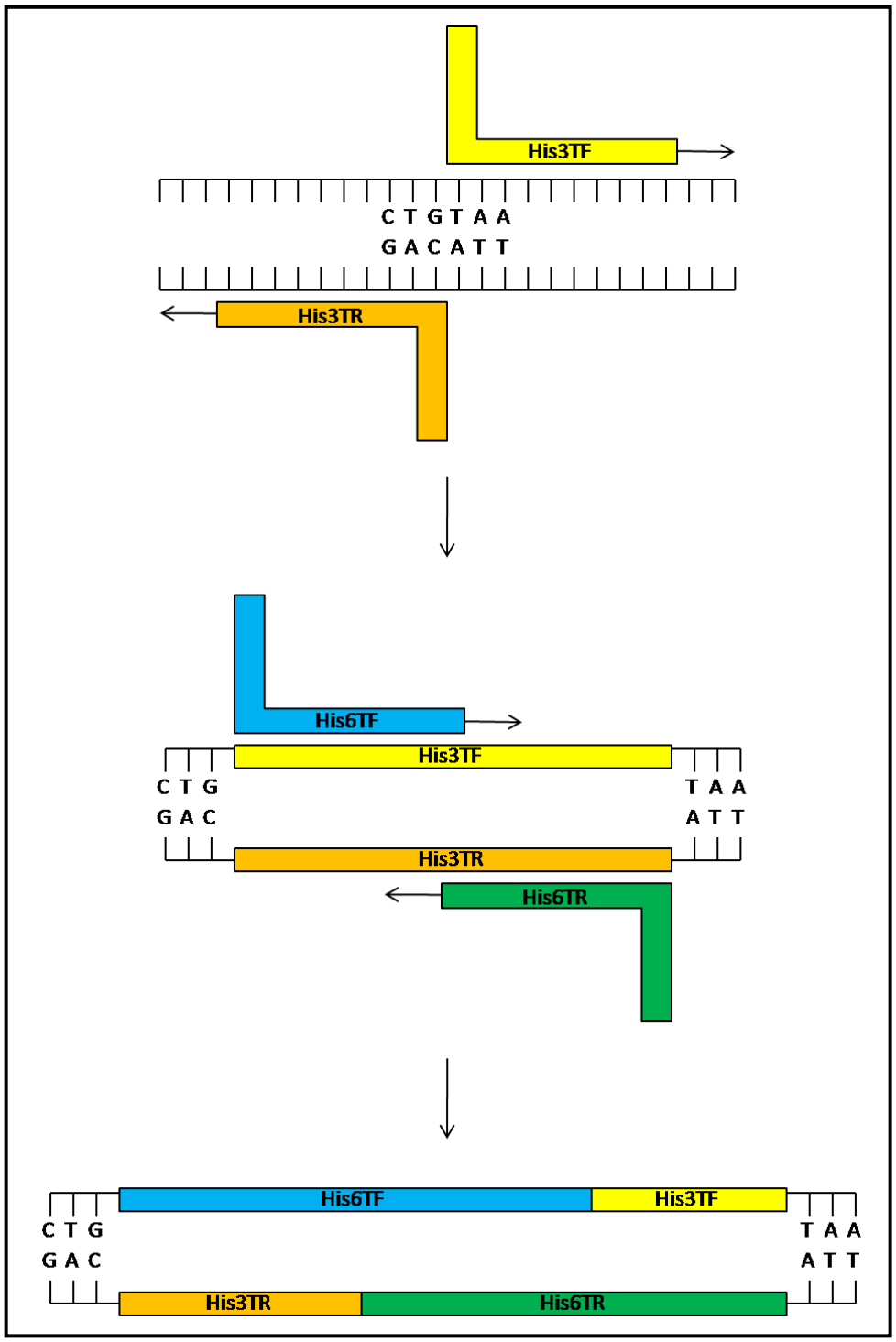


Fig 2.1. Summary of His tag addition to PETNR, top image shows the binding of the oligonucleotide primers to either side of the insertion site (selected nucleotides are labelled), arrows indicate the direction of DNA synthesis from the primer. The His_x tag sequence is derived from the 3' and 5' overhangs, and the resulting nicked plasmid is shown in the lower part of the figure. The first PCR reaction is represented in the figure, the second reaction is performed along similar principles, however the existing part of the tag from the first reaction is included in the primers.

Incorporation of the His tag to PETNR was performed in 2 reactions. The primers, named His3T and His6T, for these PCR reactions are detailed in Appendix 1. Concentrations for the PCR reaction and cycling parameters are detailed in Section 2.2.1.6.

The final PCR product was gel purified according to Section 2.2.1.3 and transformed into XL10 Gold *E. coli* according to the protocol in Section 2.2.1.17 and 2.2.1.18. Resulting colonies were screened for the presence of the His tag by colony pick PCR, according to the method in Section 2.2.1.7 and 2.2.1.19, and sequenced, Section 2.2.1.5, to confirm the identity.

2.2.4.3 Biotin Tagged PETNR ($PETNR_{Bio}$) generation

Immobilisation using the biotin method required the addition of a 45 bp insert (GGCCTGAACGATATTTTTGAAGCGCAGAAAATTGAATGGCATGAA) corresponding to a 15 amino acid recognition sequence (GLNDIFEAQ**K**IEWHE). This is the recognition sequence for the enzyme Biotin Ligase to covalently attach a biotin molecule on the highlighted K residue of the sequence.

In order to incorporate the biotin tag, 2 pairs of oligonucleotides Bio3 and 5, sequences of which are detailed in Appendix 1, were annealed separately according to the protocol discussed in Section 2.2.1.11. These oligos, when assembled, encode the biotin ligase recognition sequence including complementary ends. Annealed oligonucleotide pairs were then phosphorylated to enable joining of the sequences according to the protocol in Section 2.2.1.12. Also prepared was a gel purified BsmI digested PETNR construct, according to the protocols in Section 2.2.1.3 and 2.2.1.13, which was subsequently dephosphorylated to prevent religation without the insert, Section 2.2.1.14. The half inserts and the digested PETNR were ligated together, according to the method in Section 2.2.1.15, to give the final $PETNR_{Bio}$ construct. A simplified graphical representation is given below in Figure 2.2.

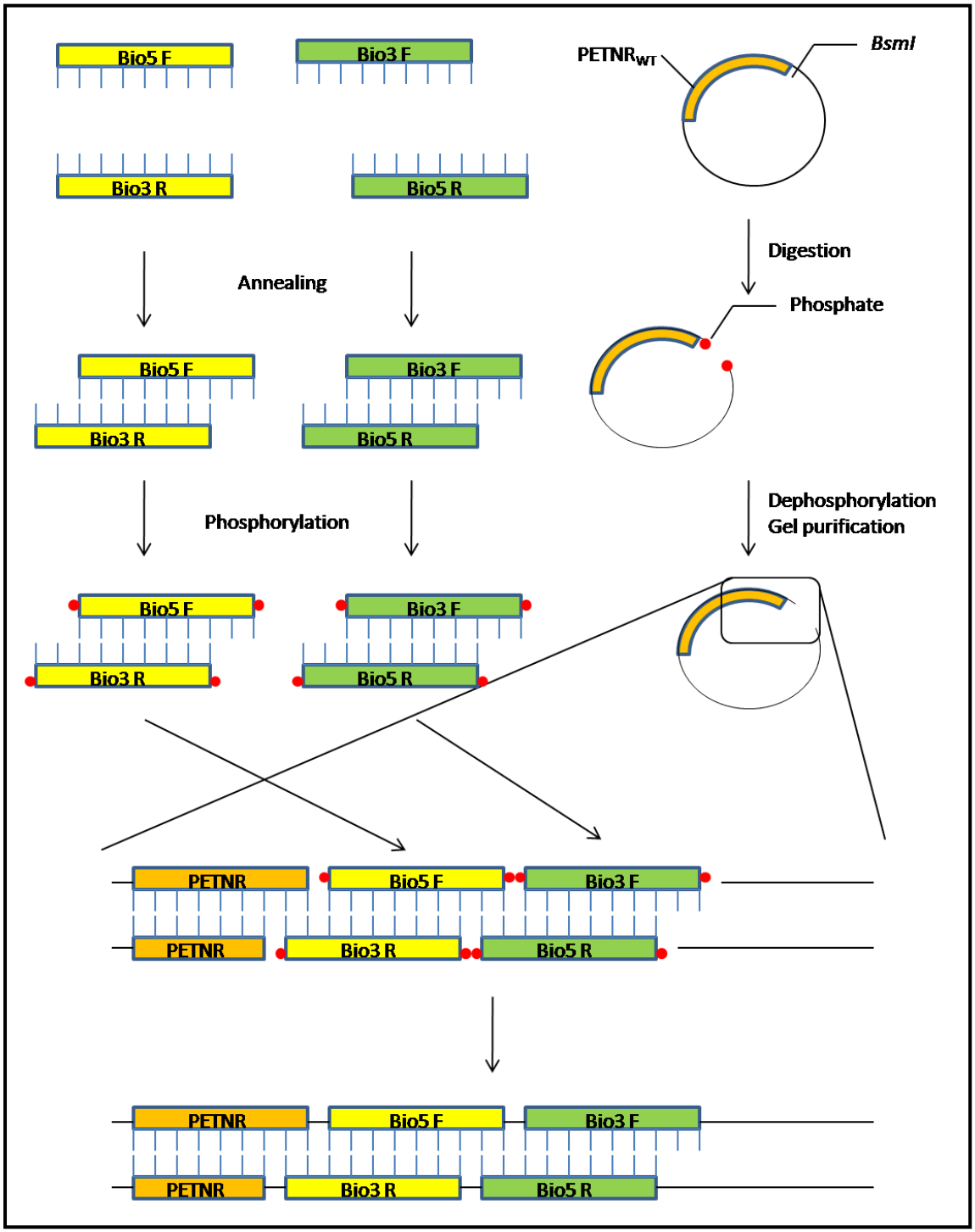


Fig 2.2. Summary of biotin tag (biotin ligase recognition sequence) addition to PETNR, The top image illustrates the mixing of complementary pairs of primers, these were then heated and cooled slowly to anneal correctly as shown in the second image, the final image shows the incorporation of annealed oligonucleotide cassette into the BsmI digested vector ready for ligation.

Following the ligation into the vector, clones which have potentially taken up the tag were detected by colony pick PCR, according to the methods discussed in Section 2.2.1.8 and 2.2.1.19, and sequenced, Section 2.2.1.5, to confirm the identity.

2.2.4.4 Ni - NTA batch purification

$PETNR_{His}$ was purified from the *E. coli* strain JM109 expressing the $PETNR_{His}$ gene construct from the same plasmid as the wild type. Cultures, 5 ml then 100 ml, were inoculated from a glycerol stock into TBAIM containing 100 µg ml^{-1} ampicillin and grown at 37 ˚C mixing at 180 rpm for 18 h. Cells were then harvested by centrifugation at 7,000 *g* for 20 min. Cell lysis was performed in Ni - NTA loading buffer as for the wild type enzyme (Section 2.2.4.1). The lysate was clarified at 20,000 *g* at 4 ˚C for 20 min. Clarified lysate (1 ml) was added to a 1.5 ml tube containing the 250 µl of 50 % Ni - NTA agarose beads. The beads were agitated for 1 min then centrifuged at 13,300 *g* at 4 ˚C for 1 min. The supernatant was discarded and the loading steps were repeated until the desired quantity of PETNR was bound to the resin. Beads were split into equal aliquots and washed with either 3 × 1.5 ml Wash buffer 1 or wash buffer 2. Elution was performed sequentially with 300 µl 50 mM potassium phosphate pH 7, ± 0.5 M NaCl, with a range of imidazole concentrations from 100 mM to 500 mM. Following addition of each wash/elution step, beads were agitated and incubated for 1 min before centrifugation at 13,300 *g* at 4 ˚C for 1 min and buffer removal.

2.2.4.5 Large scale purification of $PETNR_{His}$

Cultures, 5 ml then 100 ml, were inoculated from a glycerol stock into TBAIM containing 100 µg ml^{-1} ampicillin and grown at 37 ˚C mixing at 180 rpm for 18 h, cells were then harvested by centrifugation at 7,000 *g* for 20 min. Cell lysis was performed in Ni - NTA loading buffer as for the wild type enzyme (Section 2.2.4.1). The lysate was clarified at 20,000 *g* at 4 ˚C for 20 min. The extract was loaded directly onto Ni - NTA agarose column, pre-equilibrated with 500 ml Ni - NTA loading buffer. The column was washed with 100 ml Ni - NTA loading buffer

followed by 500 ml Ni - NTA Wash Buffer. Enzyme elution was performed using Elution Buffer 1, with a final elution step with Elution Buffer 2.

PETNR-containing fractions, as described previously (Section 2.2.4.1), were then pooled and dialysed overnight into Kinetic Assay buffer, using 10 kDa MW cut off dialysis tubing. Pooled eluates were concentrated on a Vivaspin 20 spin filter with a 30 kDa MW cut off at 3000 *g* at 4 ˚C. Protein (~14 mg ml^{-1}) was frozen in small aliquots in liquid nitrogen and stored at -80 ˚C.

2.2.4.6 $PETNR_{Bio}$ induction trials

$PETNR_{Bio}$ was purified from the *E. coli* strain AVB100 expressing the PETNR gene from our pBluescriptII KS+ plasmid, containing the $PETNR_{Bio}$ construct under the control of an IPTG inducible promoter and the biotin ligase gene encoded chromosomally under the control of an arabinose inducible promoter. Starter cultures (5 ml then 100 ml) LB containing 100 µg ml^{-1} ampicillin were grown from a glycerol stock at 37 ˚C at 180 rpm for 12 h, were used to inoculate larger cultures of Terrific Broth (TB) medium 100 µg ml^{-1} ampicillin. Cultures were grown at 37 °C at 180 rpm until they reached an OD_{600} of 0.7. The culture was divided into equal aliquots to test a variety of induction conditions. Biotin was added to all cultures to a final concentration of 50 µM. Induction was performed by the addition of 0 – 0.5 % (w/v) arabinose and 0 – 0.4 mM IPTG, Followed by growth at 37 ˚C at 180 rpm for 18 h, cells were harvested by centrifugation at 7,000 *g* at 4 ˚C for 20 min. Cells were lysed in 1 × bugbuster, 7.5 kU ml^{-1} *r*Lysozyme, 50 U ml^{-1} Benzonase and protease inhibitors for 1 h at 37 ˚C. The lysate was clarified at 20,000 *g* at 4 ˚C for 20 min. Clarified lysate (1 ml) was added to a 1.5 ml tube containing the 200 µl of 50 % monomeric avidin beads pre-equilibrated with PBS. Immobilisation was incubated at 25 ˚C at 180 rpm for 1 h. Following this, incubation resin was centrifuged at 13,300 *g* at 4 ˚C for 1 min, before the supernatant was discarded. Resin was washed twice with 1 ml Phosphate Buffered Saline (PBS), before being eluted with 200 µl monomeric avidin elution buffer, all fractions were stored for analysis by SDS PAGE.

2.2.4.7 High throughput lysis trials

Lysis trials were performed on 1 ml cultures of $PETNR_{His}$ purified in 96 well format with Ni - NTA agarose beads. Columns of $PETNR_{His}$ from the same culture were aliquoted into 4 columns of a DWB, centrifuged at 1107 *g* at 4 ˚C for 20 min. Pellets from each row were resuspended in 250 µl of a different lysis buffer from; Cell lytic B, Bugbuster or 50 mM potassium phosphate, 300 mM NaCl, 10 mM Imidazole pH 8 containing either 0.05 % Tween or 0.1 % Triton-X 100. Resuspended pellets were shaken at 180 rpm for 1 h. Following centrifugation at 1107 *g* at 4 ˚C for 20 min, the supernatant was transferred to a new DWB, before the addition of 50 µl 5 × equilibration buffer (1.5 M NaCl, 0.2 M Imidazole, 1 mM FMN pH 7.9). Ni - NTA agarose beads (200 µl, 50 %) were added to the cell lysate and shaken at 180 rpm for 30 min. Supernatant was removed following separated of the Ni - NTA beads by centrifugation at 1107 *g* at 4 ˚C for 1 min. The same procedure was repeated to perform 3 × 500 µl washes with Ni - NTA wash buffer followed by a 100 µl elution with Ni - NTA elution buffer. The resulting eluates were analysed by SDS PAGE and spectrophotometric activity assays.

2.2.4.8 Purification trials 96 well format

2.2.4.8.1 His-Mag/Talon

Both His-Mag and Talon Magnetic beads (1 ml, 50 %) were resuspended in 7 ml binding buffer, and 50 µl was aliquoted per well. $PETNR_{His}$ (250 µl of 5 µM) was added to each well and the slurries were shaken for 1 min to allow enzyme binding. The supernatant was removed by magnetic bead pull down using the magnetic block. The beads were washed with 3 × 500 µl magnetic bead wash buffer using the procedure of addition, agitation, incubation and supernatant removal as previously. Enzyme was eluted by the addition of 50 µl elution buffer. To assess the amount of enzyme purified, 50 µl of eluate was assayed with 6.5 mM 2CH, 0.1 mM NADPH in a 96 well format as described in Section 2.2.4.9.5.

2.2.4.8.2 His-Grab HBC plates

For each plate 250 µl of 5 µM purified $PETNR_{His}$ in pH 7 potassium phosphate buffer was added to each well. The plates were shaken for 1 h at 25 ˚C allowing maximal binding of enzyme to the plate. The plate was washed 3 times with Phosphate Buffered Saline (PBS), and a further 2 times with 50 mM potassium phosphate buffer pH 7 to avoid nonspecific binding observed in early trials. Assays were performed as in Section 2.2.4.9.5. without the enzyme addition steps.

2.2.4.9 Library generation PCR

Mutagenic libraries were generated by the binding of degenerate oligonucleotide primers to the WT template during a PCR reaction and the reaction proceeds to generate the remainder of the plasmid as described in figure 2.3.

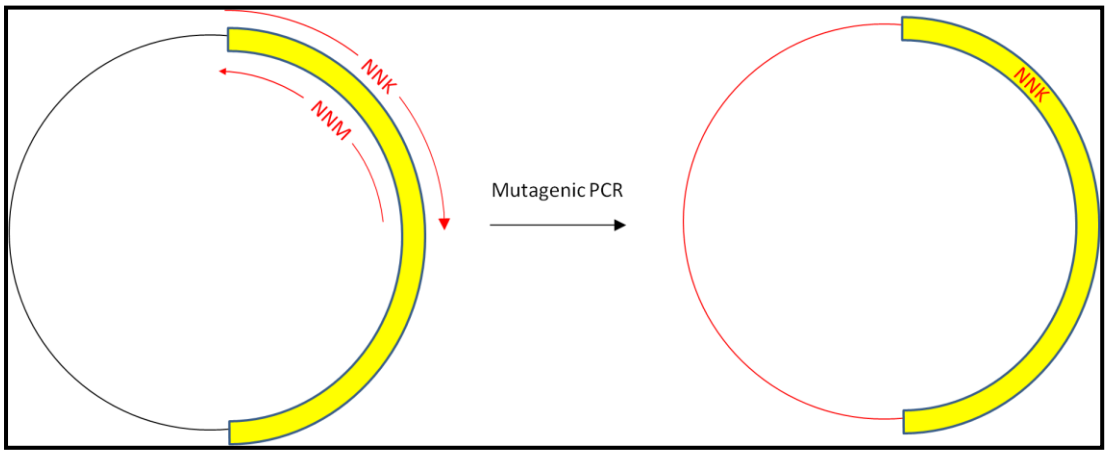


Fig 2.3 Method of generating mutagenic libraries using degenerate oligonucleotides. Degenerate oligonucleotides bind template DNA and are extended, replicating the remainder of the plasmid and incorporating the mutation specified by the degenerate oligonucleotide. Following PCR, nicked plasmids are transformed into *E. coli* and ligated.

Single site saturated libraries were generated at a number of positions around the active site of PETNR by PCR mutagenesis with KOD DNA polymerase (Section 2.2.1.8), using NNK degenerate oligonucleotides (Appendix 1). All products were *DpnI* digested (Section 2.2.1.9) and analysed by DNA gel electrophoresis (Section 2.2.1.1). The correct sized band was gel purified using a Qiagen gel purification kit (Section 2.2.1.3) and quantified using Qubit fluorometer (Section 2.2.1.4). Library

DNA was transformed into the *E. coli* strain JM109 (Section 2.2.1.18) and grown on LB agar containing 100 µg ml^{-1} ampicillin at 37 ˚C for 12 h.

2.2.4.10 Robotics methods

2.2.4.10.1 Colony pick method

Deep well blocks (96 wells, 2.2 ml each) containing 1 ml per well TB with 100 µg ml^{-1} ampicillin were manually prepared in a PCR workstation UV sterile hood and introduced to the robotics glove box along with LB agar plates containing the correct density of *E. coli* colonies (~150 per plate) for colony picking. The robotics platform picked 96 using an 8 channel pipetting system with disposable tips which were transferred into the wells of the media containing DWB. The block was removed from the anaerobic environment and covered with a sterile, air permeable, self adhesive film. Cultures were incubated at 37 °C at 1050 rpm for 18 h in a Titramax 1000 incubator.

Following the growth period 50 µl of overnight growth were added to 50 µl of glycerol stock buffer (80 % v/v PBS, 20 % v/v Glycerol) in a 96 well PCR plate (Starlabs) and covered with a 96 well silicone sealing mat (Starlabs). Glycerol stocks were then frozen on dry ice for 30 minutes before being transferred to a -80 °C freezer for medium-long term storage.

2.2.4.10.2 96 well growth method – AIM

Glycerol stocks of clones in 96 well format were used to inoculate 1 ml cultures in DWB of TB AIM containing 100 µg ml^{-1} ampicillin. Cultures were grown at 37 °C at 1050 rpm for 24 h, which allows additional time for auto induction of protein production.

2.2.4.10.3 Harvesting of 96 well growth

Following growth, the breathable seal was replaced with a Biomek aluminium foil seal and cultures were centrifuged at 1109 *g* at 4 °C for 20 min. The supernatant was discarded and pellets were frozen at -20 °C for short term storage.

2.2.4.10.4 96 well purification

Cell pellets were lysed by the addition of 200 µl of lysis buffer, followed by 10 × 150 µl mixing cycles. The DWB was moved to the shaker and agitated for 5 min. Following lysis, 50 µl of equilibration buffer was added to the lysate to ensure maximal flavination of mutant enzymes and to idealise conditions for immobilising enzyme to the magnetic beads. The magnetic bead slurry (50 µl of a 7 % stock in binding buffer) was added to the cell lysate by the robot. The bead slurry was shaken on a shaker for 3 min to allow binding of the tagged enzymes to the beads. The DWB was removed from the shaker and placed on the magnet plate to immobilise the beads and enable supernatant removal. Three wash steps were then performed consisting of resuspention of magnetic beads in 700 µl (or 200 µl for subsequent washes) wash buffer, immobilisation of the magnetic beads on the magnet (30 sec) and removal of the wash buffer to waste. The enzyme was eluted in 120 µl of elution buffer and the DWB was shaken for 3 min to maximise protein elution. The eluate was removed further 180 µl of dilution buffer in a 96 well microplate in order to reduce the final imidazole concentration.

2.2.4.10.5 Bead cleaning and regeneration

Following each use magnetic beads were stripped and regenerated using 3 aliquots of strip buffer followed by 3 aliquots of charge buffer, each followed by rinsing with 3 aliquots of water. This procedure was usually performed on the large magnet with pooled magnetic bead solutions and aliquots of 15 ml.

2.2.4.10.6 Robotic assay

Library assays were performed in 96 well UV transparent - UV Star - microplates. Each library extraction provided enough purified enzyme for up to 3 assays. Assay buffer (150 µl) was added to all wells of the reaction plate followed by the reductive substrate NADPH (50 µl of 600 µM in buffer) and the enzyme eluant to a final volume of 250 µl. To start the assay, 50 µl of a 6 × stock, in 30 % EtOH, of the oxidative substrate was added to the reaction. The robotics platform transferred

the assay plate to an anaerobic plate reader, followed by shaking for ~ 10 s, before NADPH consumption was monitored at 340 nm for 10 min.

2.2.4.10.7 Calculation of Z factor

Data derived from screening experiments were analysed using a method of non controls based normalisation, called Z factor scoring [159].This method involves the calculation of z scores from experimental parameters according to Equation 2.10.

Equation 2.10. Calculation of z factor for screening data according to the method of [159]. is the z value for a given well, is the experimental data for a given well, is the mean for the entire sample and is the standard deviation for the sample.

3 Generation, purification and characterisation of novel tagged PETNR variants and robotic growth, lysis, purification and assay trials.

3.1 Introduction

The overall focus of the work described in this chapter was the generation and robotic screening of libraries of PETNR mutants to generate novel biocatalysts with improved activity towards a range of industrially relevant substrates (Figure 3.1). To achieve this, the PETNR construct was altered to incorporate C-terminal tags, to facilitate rapid and efficient protein purification in a 96 well format on a robotics platform. A number of immobilisation technologies have been demonstrated including the immobilisation of a poly-histidine tag using metal affinity chromatography [160] and the immobilisation of biotin tagged protein to resins related to the bacterial protein streptavidin [161]. Options for the robotic high throughput purification include the use of coated microplates or the use of magnetic beads.

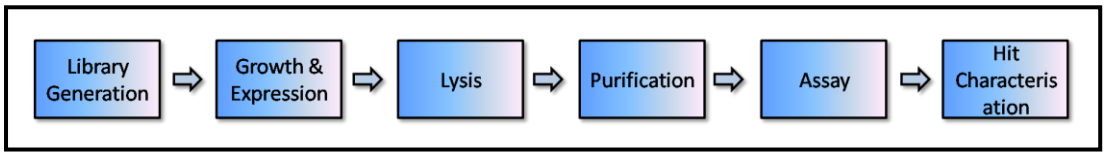


Figure 3.1. Experimental outline of protein engineering experiments to change the properties of PETNR for more adventitious use as a biocatalyst.

In order to accurately assess the potential usefulness of each tagged enzyme construct, none tagged enzyme ($PETNR_{WT}$; [21]) was prepared and kinetically characterised [162]. The activity and structure of the tagged enzymes was determined and compared to $PETNR_{WT}$ to ensure the tags had no major effects on biocatalytic potential, as has been observed previously in the case of other OYE homologues [79]. Successful tagged constructs were then used for optimising robotic screening and validation procedures such as protein expression and purification.

A number of components for robotic screening of enzyme variants were trialled and optimised, including the development of a colony picking methodology, cell growth

and PETNR expression in a 96 well format, culture lysis, protein purification and activity detection.

After optimisation of the robotics protocols, trial studies were performed with all $PETNR_{WT}$ cultures to determine the range of errors inherent in the new protocols, such as variation in protein expression, levels of protein purified and enzyme activity assays.

3.2 PETNR Wild Type ($PETNR_{WT}$)

Initially, $PETNR_{WT}$ was expressed, purified and characterised kinetically by determining the reductive half-reaction using transient (single turnover) kinetic methods and steady-state reaction rates. These data were generated to provide a basis for comparison with the tagged enzymes to investigate if the presence of the tags compromised activity.

3.2.1 $PETNR_{WT}$ purification

$PETNR_{WT}$ (encoded by gene pONR-1 [3]) was expressed in large quantities (~50 mgL^{-1}) from the *Lac* promoter of pBluescript SK+ [21]. $PETNR_{WT}$ was purified according to the previously published protocol incorporating affinity separation using mimetic orange resin followed by ion exchange chromatography using Q-Sepharose resin [40], as described in Section 2.2.4.1. Purification was followed by monitoring the relationship between the total protein aromatic residue content (estimated from absorbance measurements at 280 nm) and the oxidised FMN content (measured at the flavin peak absorbance at 464 nm). The extent of purification was monitored visually by SDS-PAGE, an example of which is presented in Figure 3.2.

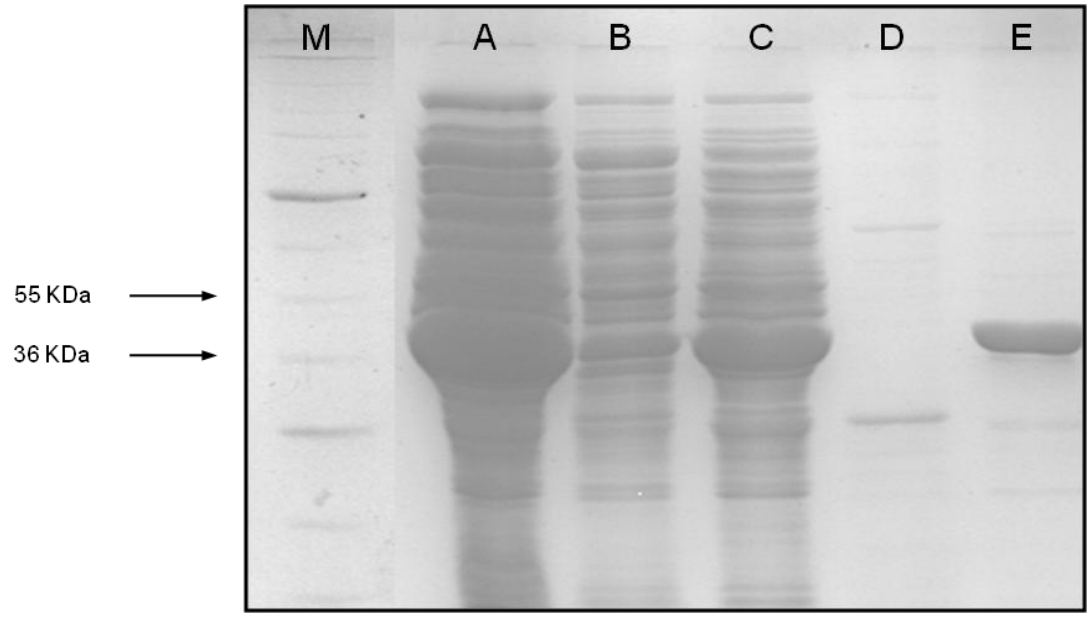


Figure 3.2. SDS-PAGE analysis of $PETNR_{WT}$ purification protocol. Lane M, Pageruler plus marker (Fermentas; 10, 17, 28, 36, 55, 72, 95, 130 and 250 kDa); Lane A, cell extract; Lane B, unbound fractions from the mimetic orange column; Lane C, mimetic orange column high salt eluates; Lane D, unbound fractions from the Q-Sepharose column; Lane E, Q-Sepharose column high salt eluates; Analysis was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained with Coomassie Brilliant Blue R250.

3.2.2 Kinetic behaviour of $PETNR_{WT}$

Steady-state and single turnover kinetic data were determined for $PETNR_{WT}$ with the model substrates NADPH and cyclohex-2-enone (CH). Data for the steady-state reactions are shown in Figures 3.3 and 3.4 and data for the pre steady-state reactions are shown in Figure 3.5.

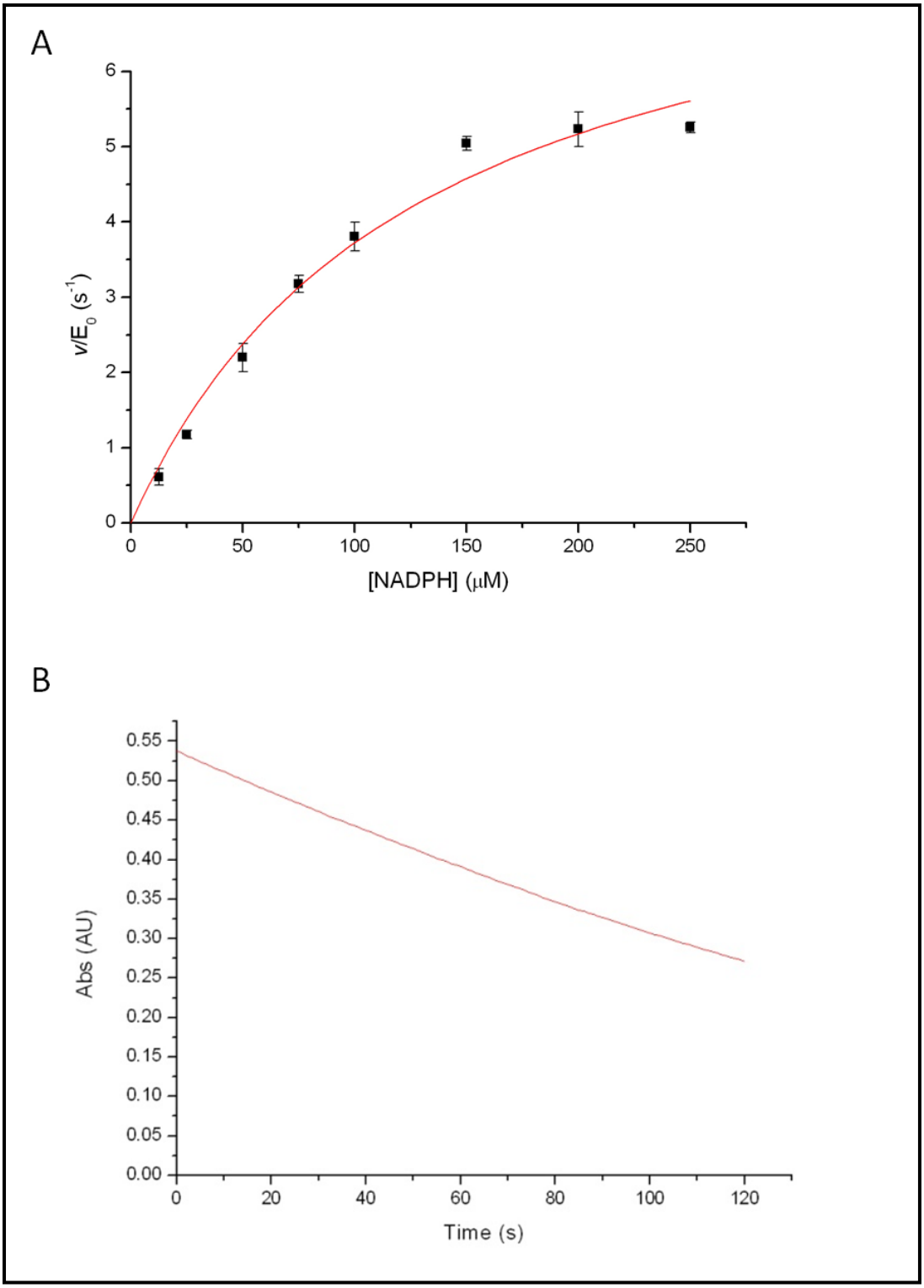


Figure 3.3. The steady-state rate as a function of NADPH concentration of $PETNR_{WT}$. Data were recorded by monitoring the rate of NADPH oxidation at 340 nm over a maximum period of 120 s. Data were recorded at 0.1 μM PETNR, 6.5 mM CH, 25 °C in 50 mM KH_2PO_4/K_2HPO_4 pH 7. Panel A, non-linear least squares fit to the Michalis-Menten equation (Section 2.2.3.2, Equation 2.2) and kinetic parameters were estimated; Panel B, example trace obtained for an assay performed with 100 μM NADPH.

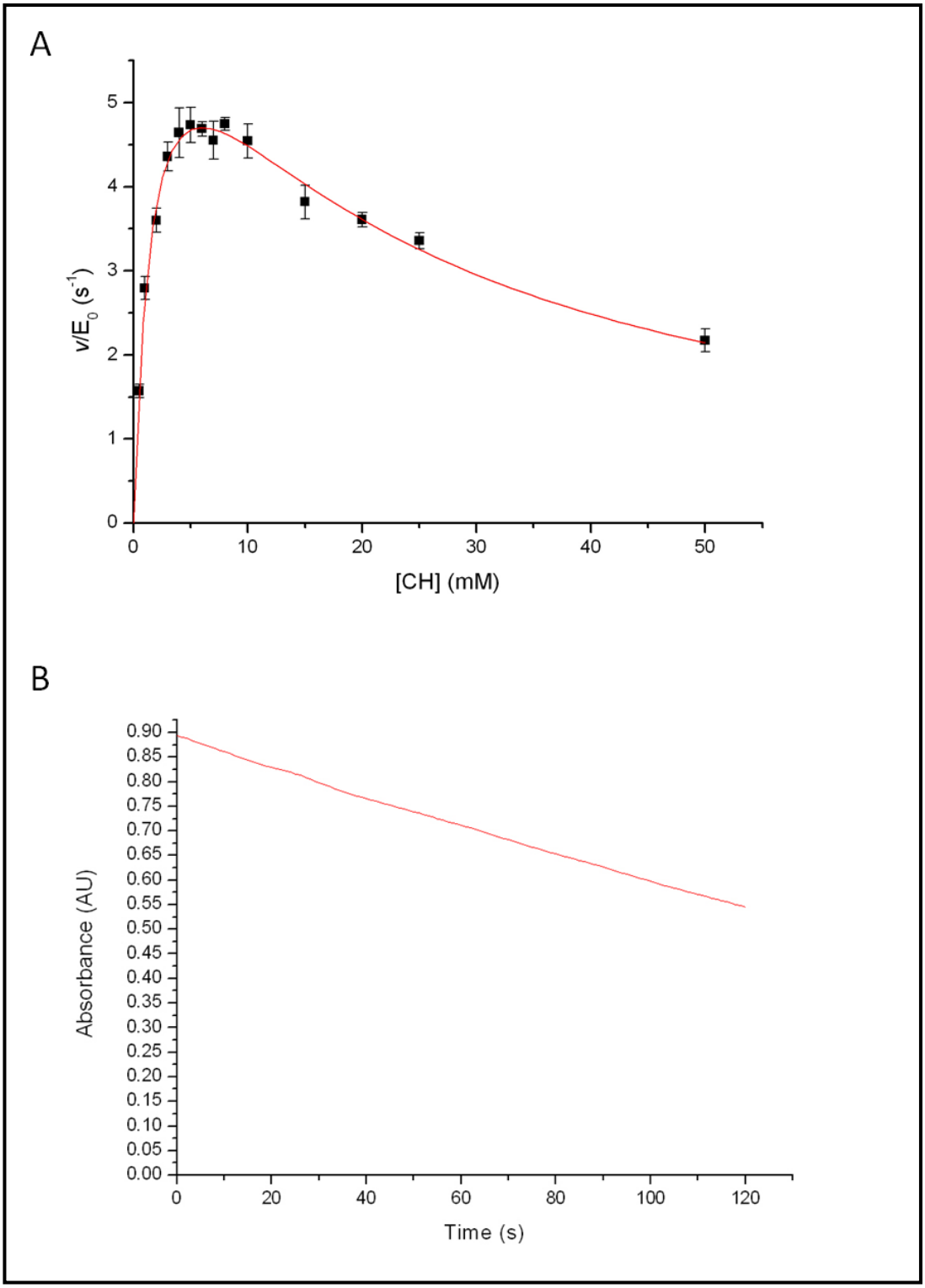


Figure 3.4. The steady-state rate as a function of CH concentration of $PETNR_{WT}$. Data were recorded by monitoring the rate of NADPH oxidation at 340 nm over a maximum period of 120 s. Data were recorded at 0.1 µM PETNR, 150 µM NADPH, 25 °C in 50 mM KH_2PO_4/K_2HPO_4 pH 7. Panel A, non-linear least squares fit to the Michalis-Menten equation incorporating substrate inhibition (Section 2.2.3.2, Equation 2.3) and kinetic parameters were estimated; Panel B, example trace obtained for an assay performed with 6 mM CH.

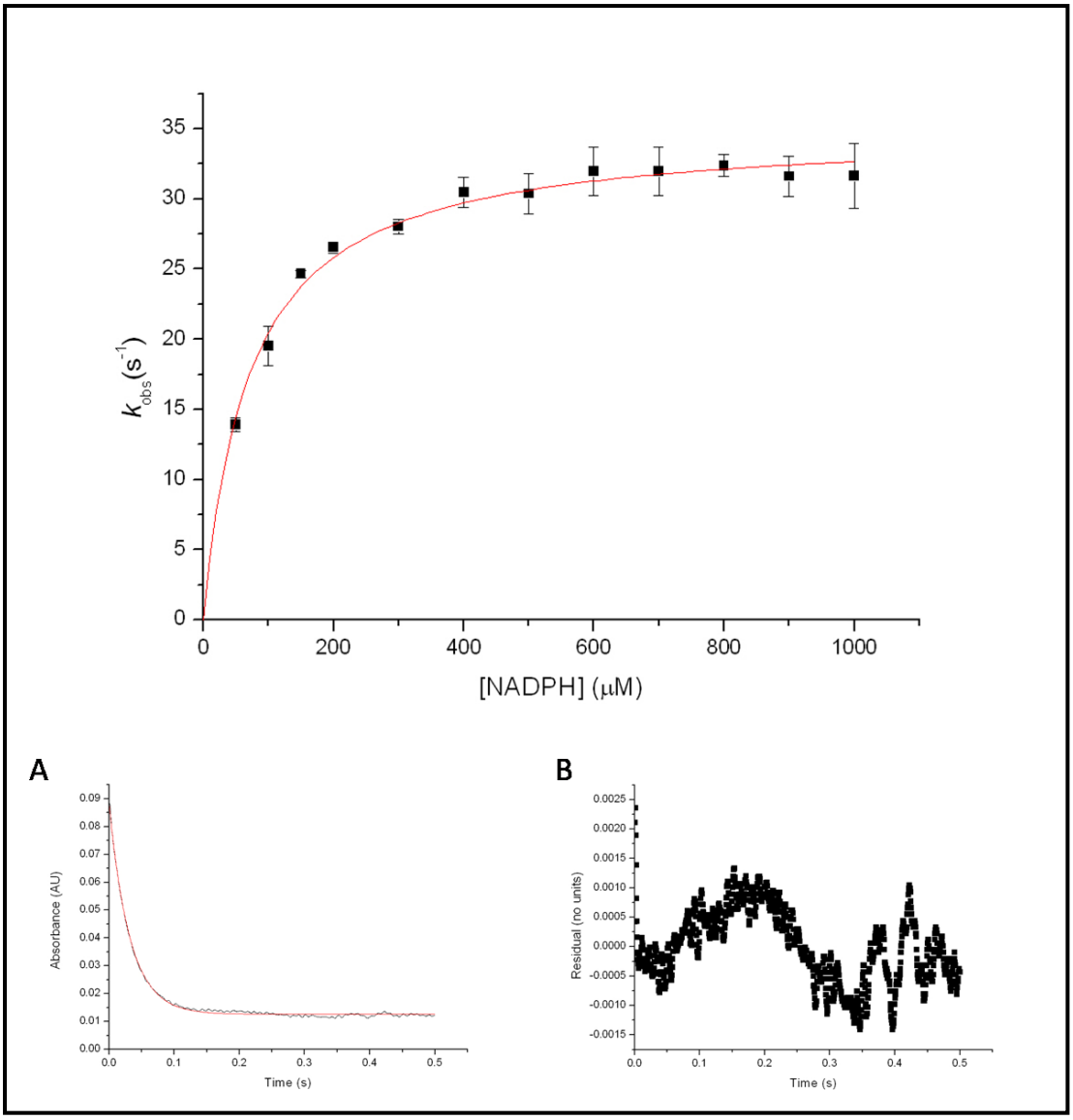


Figure 3.5. The pre steady-state reductive half-reaction of $PETNR_{WT}$. Data were recorded by monitoring the rate of FMN reduction at 464 nm over a period of 1 s following rapid mixing of oxidised enzyme with NADPH. Data were recorded for reactions containing 10 μM PETNR, in 50 mM pH 7 KH_2PO_4/K_2HPO_4 buffer at 25 °C. Panel A shows a non-linear least squares fit to the Strickland equation (Section 2.2.3.3, Equation 2.5) and kinetic parameters were estimated. Panel B, data obtained for a reaction with 800 μM NADPH, red line shows the fit to a single exponential equation (Section 2.2.3.3, Equation 2.4); Panel B, residuals plot of the of the fit shown in panel A.

From the above data, the apparent kinetic parameters associated with the reaction of PETNR with NADPH and CH were estimated as k_{cat} = 8.5 ± 0.8 s^{-1}, K_M^{NADPH} = 128.1 ± 24.5 μM for the steady-state turnover at fixed CH concentration (6.5 mM). For steady-state reactions at fixed NADPH concentrations (150 μM) k_{cat} = 7.5 ± 0.3 s^{-1}, K_M^{CH} = 1.8 ± 0.2 mM. Substrate inhibition was observed and the value of K_i^{CH} = 20.4

± 1.9 mM. From the single turnover stopped-flow studies of the reductive half-reaction (i.e. reduction of enzyme by NADPH) the flavin reduction rate, k_{red} was calculated as k_{red} = 34.9 ± 0.4 s^{-1} and the apparent dissociation constant, K_D, for the E-NADPH complex was calculated as K_D = 70.9 ± 4.6 µM. The results above are consistent with published data where they are available (Table 3.1). Other parameters have not been determined previously and provide additional data for comparison with tagged enzymes.

Experiment		Parameter	This study	Published value	Ref
Steady-state	Fixed CH	k_{cat} (s^{-1})	8.5 ± 0.8	*nd*	*n/a*
		K_M^{NADPH} (µM)	128.1 ± 24.5		
	Fixed NADPH	k_{cat} (s^{-1})	7.5 ± 0.3	*nd*	*n/a*
		K_M^{CH} (mM)	1.8 ± 0.2		
		K_i^{CH} (mM)	20.4 ± 1.9		
Single turnover	Reductive half-reaction	k_{red} (s^{-1})	34.9 ± 0.5	34 ± 0.4	[162]
		K_D (µM)	70.9 ± 4.6	73 ± 4	

Table 3.1. Summary table of $PETNR_{WT}$ kinetic data. Comparison between the experimentally derived data for $PETNR_{WT}$ and published values where appropriate.

3.3 His Tagged PETNR ($PETNR_{His}$)

A crucial factor in the robotic protocol is the requirement for small-scale enzyme purification in a 96 well format. The only feasible approach is the use of affinity purification via the incorporation of an affinity tag to the enzyme.

The poly histidine tag is a commonly used N- and C- terminal protein tag [160]. The number of residues in the tag can vary, but is most commonly between 6 and 12 residues long [160]. Proteins with this amino acid sequence can coordinate to divalent metal ions such as Ni and Co as demonstrated in Figure 3.6.

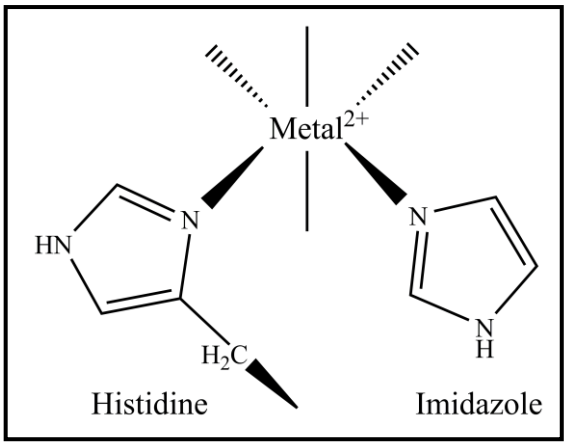


Figure 3.6. Histidine side chain interacting with a divalent metal ion in immobilised metal affinity chromatography. High concentrations of imidazole are required to compete with His residues in the target protein for the metal ion, leading to elution of the His-tagged protein.

Once immobilised, the protein of interest can be eluted using a high concentration of imidazole, which then competes with the His residues in the protein in binding the metal ion. One notable problem of this technique is the potential for other *E. coli* proteins to co-purify with the protein of interest, such as FK506 - binding proteins [160], including SlyD [163].

It has been observed that addition of a His-tag to the C terminal region of YqjM impairs protein function dramatically [79]. However, this loss of function is probably due to the impairment of oligomerisation [43] and is not expected to be a problem for PETNR as it is thought to function as a monomer [3]. The tag was localised on the C-terminal end as it was thought that an N-terminal tag might disrupt or block access to the active site channel.

3.3.1 $PETNR_{His}$ construction

The poly-His tag was added in two sequential PCR reactions, each incorporating either 3 or 6 histidine residues by incorporating an overhanging tail (as discussed in Section 2.2.4.2). Following each PCR, the product was analysed by agarose gel electrophoresis, and the amplified product was purified from the gel to yield the template for the next reaction (first round) or the final product (second round). Agarose gel electrophoretic analysis of the 'tagged' PCR product is shown in Figure 3.7. The template (first round product) encoding $PETNR_{3T}$ was purified from the agarose gel and used as a template for the second PCR reaction using the '6T primers' to incorporate additional bases that encode the final 3 residues of the His-tag. In addition, a second PCR reaction was carried out with the 'His6 primers' using the gene encoding wild-type PETNR as the template. Surprisingly, the expected 4.5 kb PCR product, encoding the $PETNR_{His}$ protein, was obtained from both templates negating the need to perform the initial His_3 incorporation step. It is likely that the T_m of the complementary part of the 'His_6 primers' was sufficiently high to allow annealing of the larger primer to the wild-type template at 72°C, most likely due to its high GC content.

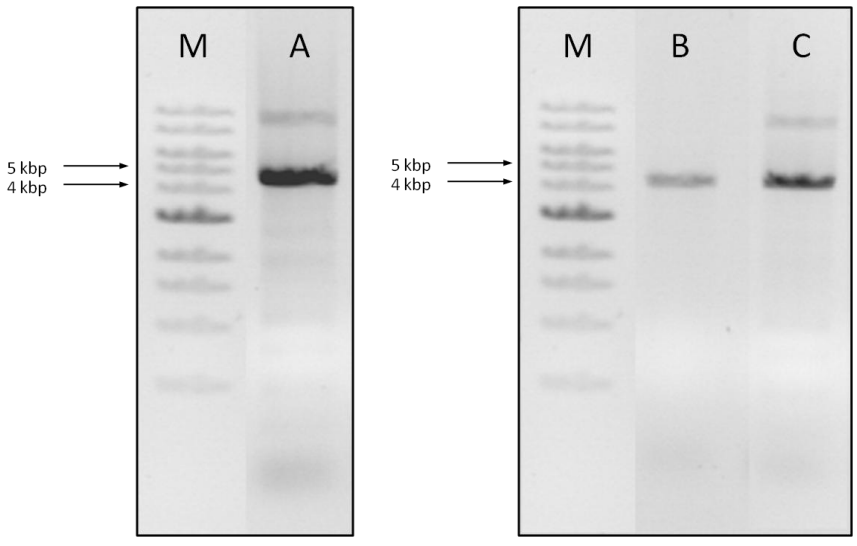


Figure 3.7. Agarose gel showing the construction of $PETNR_{His}$. Lanes M, 1 kbp ladder (NEB; 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 kbp); Lane A, reaction with $PETNR_{WT}$ template and 3T primers; Lane B, reaction with $PETNR_{3T}$ template and 6T primers; Lane C, reaction with $PETNR_{WT}$ template and 6T primers. The major band at 4.5 kb corresponds to the expected size of the PETNR construct. Analysis was performed on a 0.6 % agarose gel electrophoresed at 100 V for 45 min.

Agarose gels of PCR products amplified from individual colonies (Colony pick PCR, CPPCR) containing the $PETNR_{His}$ encoding plasmid were electrophoresed for 1 - 1.5 h. This enabled separation of DNA fragments with small differences in base pair content (18 bp). To overcome problems of uneven lane running due, attributed to heat generation during the long electrophoresis running time, control ($PETNR_{WT}$) DNA samples were electrophoresed at regular intervals throughout the gel. An example is given in Figure 3.8. Positive results from the CPPCR were further characterised by DNA sequencing to confirm the presence of nucleotides encoding the poly-histidine tag.

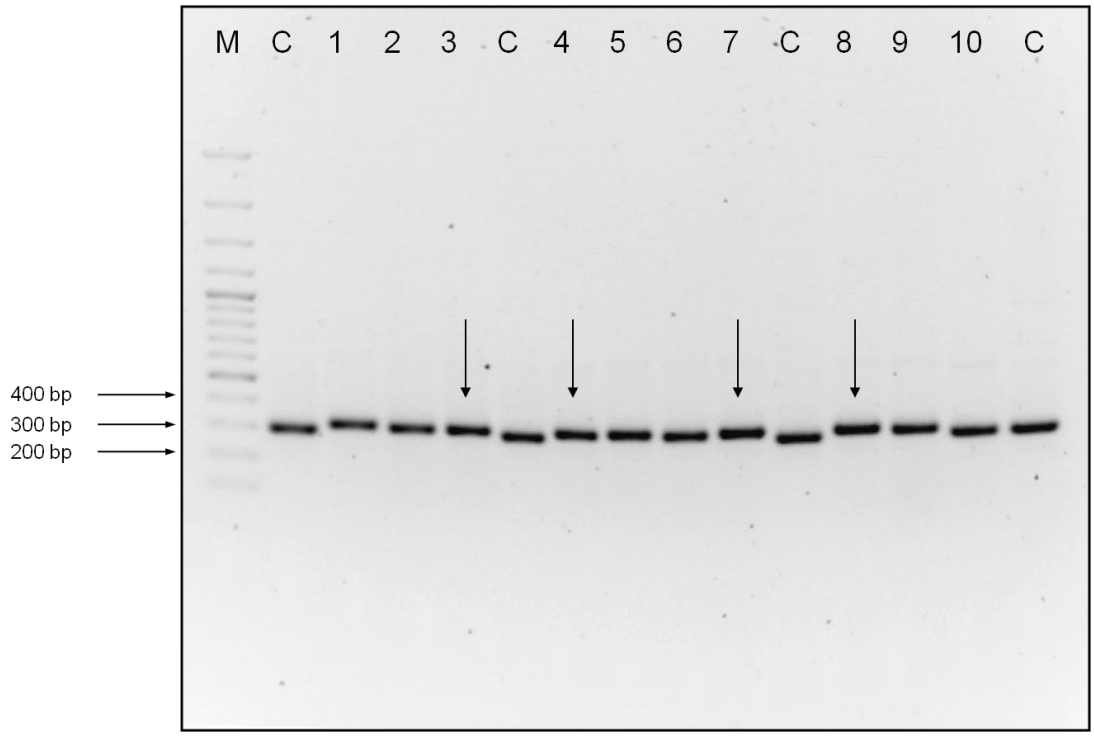


Figure 3.8. Agarose gel showing the PCR products from 'colony pick PCR' of $PETNR_{His}$ clones. Lane M, 100 bp Ladder (NEB; 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bp, 1.2, 1.5, 2, 3 kbp); Lane C, control reaction with wild-type template; Numbered lanes, sample PCR products Arrows indicate lanes sequenced. Analysis was performed on a 1.3 % agarose gel electrophoresed at 100 V for 1 – 1.5 h.

Clones shown in lanes 7 and 8 (hereafter termed His7 and His8) of Figure 3.8 show the greatest difference in migration in the agarose gel compared to the control product. Sequencing of these clones showed that both contained DNA encoding an in-frame poly-histidine tail immediately before the stop codon. However, it was

found that both clones encoded eight instead of the expected six histidines. Figure 3.9 shows an alignment of the sequence of the 3' terminus of clone His8 compared to the wild-type sequence. It is likely that the two additional codons were incorporated as a result of primer mis-annealing. However, as tags in the range of 6 - 12 histidine residues are frequently used to bind to Nickel affinity resins, the His8 clone was used for further studies.

Clone His8 was further sequenced and was found to contain an additional mutation A921G compared to the PETNR deposited sequence [21]. This corresponds to a silent mutation at residue 307 changing GAA to GAG. However, sequencing of wild-type PETNR template revealed the presence of this mutation. This silent mutation could be a natural isoform of the gene, or more likely is a mistake in the initial sequencing of the gene. In any circumstance there are no anticipated consequences for the proposed scheme of research.

```
Predicted     TGATTATCGGGGCGGGTGCGTATACGGCAGAAAAGCCGAGGATTTGATCGGTAAAGGCC
PETNR His 8   TGATTATCGGGGCGGGTGCGTATACGGCAGAGAAAGCCGAGGATTTGATCGGTAAAGGCC
              ****************************** ****************************

Predicted     TGATCGACGCCGTGGCCTTTGGCCGTGACTACATTGCTAACCCGGATCTGGTTGCCCGTT
PETNR His 8   TGATCGACGCCGTGGCCTTTGGCCGTGACTACATTGCTAACCCGGATCTGGTTGCCCGTT
              ************************************************************

Predicted     TGCAGAAAAAAGCCGAACTGAACCCGCAGCGTCCTGAAAGCTTCTATGGCGGCGGCGCGG
PETNR His 8   TGCAGAAAAAAGCCGAACTGAACCCGCAGCGTCCTGAAAGCTTCTATGGCGGCGGCGCGG
              ************************************************************

Predicted     AAGGTTATACCGACTACCCTTCACTGCATCATCACCATCACCAT------TAATCCCGCT
PETNR His 8   AAGGTTATACCGACTACCCTTCACTGCATCATCACCATCACCATCACCATTAATCCCGCT
              ****************************************      **********
```

Figure 3.9. Clustal alignment of the 3' sequence of the deposited $PETNR_{WT}$ encoding DNA sequence with the resultant His8 sequence [164]. Green, DNA encoding the expected His_6 tag; cyan, additional bases in the experimentally derived His8 tag encoding region; yellow, triplet affected by the silent mutation; red text, the silent mutation in PETNR; underline, stop codon.

3.3.2 Trial batch purification of $PETNR_{His}$ with Ni - NTA agarose

Before any further studies with $PETNR_{His}$ variant could be conducted, optimisation of the purification was required. Standard protocols for the purification of poly histidine-tagged enzymes include varying concentrations of imidazole (100 - 500 mM) and NaCl concentration (0 – 0.5 M; Figures 3.10 and 3.11) during elution [165].

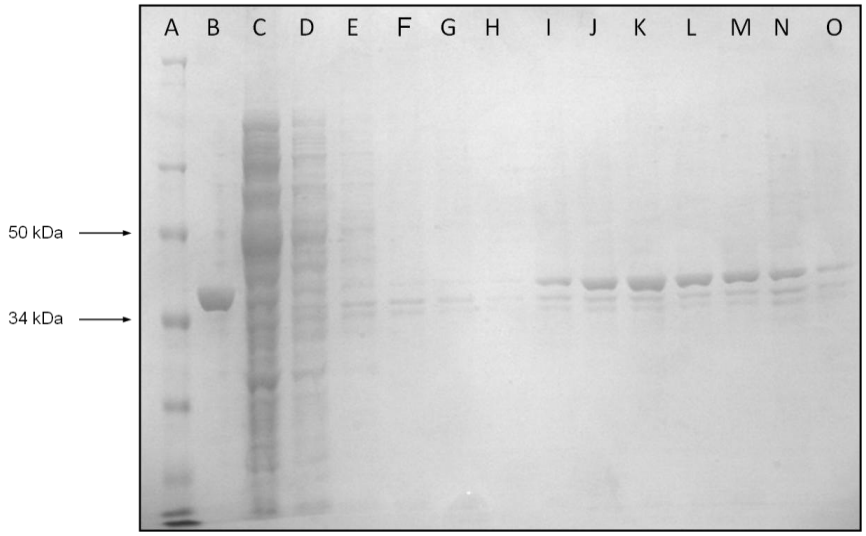


Figure 3.10. Purification of $PETNR_{His}$ in the absence of NaCl. Lane A, 'low range' prestained protein ladder (bio-rad; 21, 29, 34, 50, 77, 103 kDa); Lane B, $PETNR_{WT}$ authentic standard; Lane C, flow through; Lanes D-H, washes 1-5; Lane I, 100 mM imidazole eluate; Lane J, 150 mM imidazole eluate; Lane K, 200 mM imidazole eluate; Lane L, 250 mM eluate; Lane M, 300 mM imidazole eluate; Lane N, 400 mM imidazole eluate; Lane O, 500 mM imidazole eluate. Analysis was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained with Coomassie Brilliant Blue R250.

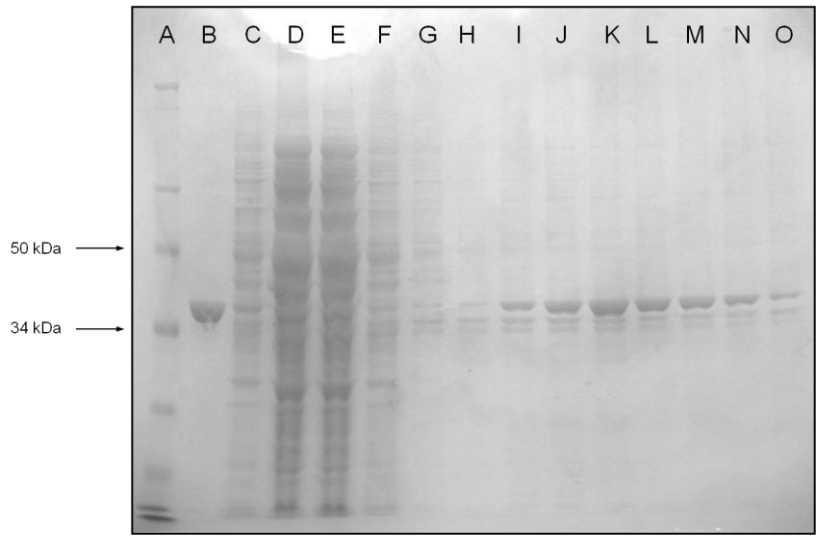


Figure 3.11. Purification of $PETNR_{His}$ in the presence of 0.5 M NaCl. Lane A, 'low range' prestained protein ladder (bio-rad; 21, 29, 34, 50, 77, 103 kDa); Lane B, $PETNR_{WT}$ authentic standard; Lane C, flow through; Lanes D-H, washes 1-5; Lane I, 100 mM imidazole eluate; Lane J, 150 mM imidazole eluate; Lane K, 200 mM imidazole eluate; Lane L, 250 mM eluate; Lane M, 300 mM imidazole eluate; Lane N, 400 mM imidazole eluate; Lane O, 500 mM imidazole eluate. Analysis was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained with Coomassie Brilliant Blue R250.

SDS PAGE analysis of the batch purification of $PETNR_{His}$ using Ni - NTA agarose in the presence or absence of salt indicated identical purification results, this observation suggests salt can be included in the purification to interrupt potential non specific binding of redox active proteins interfering with the assay. Results indicated that the minimal amount of imidazole required for significant elution of $PETNR_{His}$ was around 200 mM. This is similar to the standard six His tagged protein purification protocols [165]. Therefore 0.5 M NaCl and 200 mM imidazole were used throughout the purification of $PETNR_{His}$ in all subsequent studies [165].

Compared to the standard $PETNR_{WT}$ protocol, $PETNR_{His}$ purification can be accomplished in a single preparative step, consistent with the requirement for incorporation into a robotic screening process.

3.3.3 Preparative scale purification of $PETNR_{His}$

A larger scale purification of $PETNR_{His}$ was performed from a 1L culture of recombinant *E. coli* cells using the optimised protocol and a 25 ml Ni - NTA agarose column. At various stages of the purification samples were taken for analysis by SDS-PAGE (Figure 3.12) and UV/visible spectrophotometry. The ratio of absorbance at 280 nm and 464 nm (A_{280}:A_{464}) was determined at different stages of the purification to assess purity.

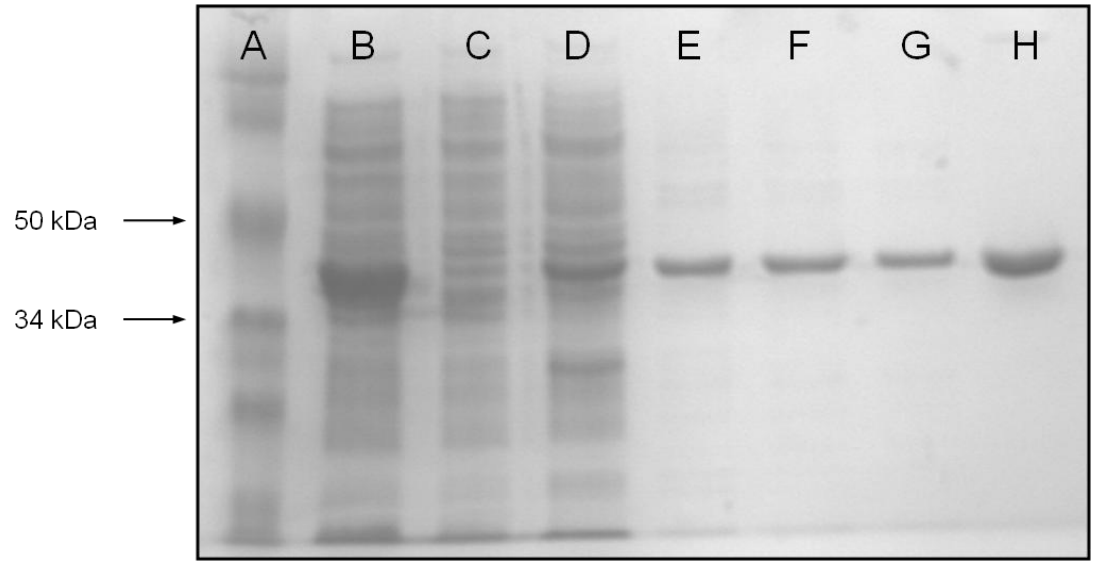


Figure 3.12. Ni - NTA purification of $PETNR_{His}$. Lane A, low range prestained protein ladder (bio-rad; 21, 29, 34, 50, 77, 103 kDa); Lane B, cell lysate (10%); Lane C, flow through (10%); Lane D – G, wash fraction 1 – 4; Lane H, eluate (10%). Analysis was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained with Coomassie Brilliant Blue R250.

A small amount of $PETNR_{His}$ is leached from the resin with every wash step, but when compared to the intensity of the eluted bands the amount of protein lost is minor. The final lane (lane H, Figure 3.12) indicates that $PETNR_{His}$ can be isolated in pure form using this purification method.

3.3.4 Analysis of kinetic behaviour of $PETNR_{His}$

The addition of a poly-histidine tag to the OYE YqjM has previously been demonstrated to be problematic [79]. To investigate the kinetic effects of incorporating a poly histidine tag to the C-terminus of PETNR, a kinetic profile for the tagged enzyme was determined under steady-state and pre steady-state conditions. Such an analysis also provides baseline data for comparison with variant forms of $PETNR_{His}$. Data for the steady-state reactions are shown in Figures 3.13 and 3.14 and data for the pre steady-state reactions are shown in Figure 3.15.

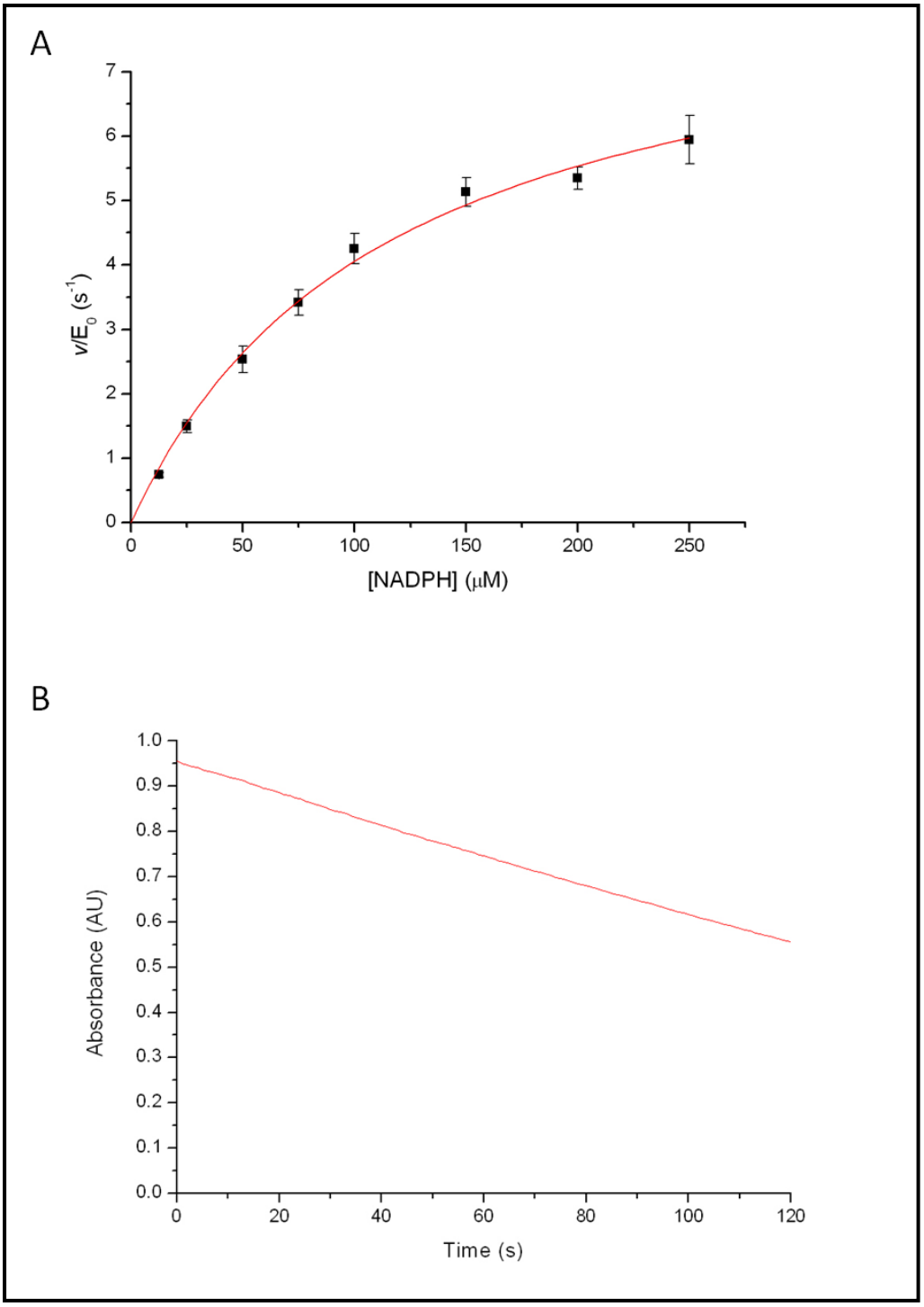


Figure 3.13. The steady-state rate as a function of NADPH concentration of $PETNR_{WT}$. Data were recorded by monitoring the rate of NADPH oxidation at 340 nm over a maximum period of 120 s. Data were recorded at 0.1 µM PETNR, 6.5 mM CH, 25 °C in 50 mM KH_2PO_4/K_2HPO_4 pH 7. Panel A, non-linear least squares fit to the Michalis-Menten equation (Section 2.2.3.2, Equation 2.2) and kinetic parameters were estimated; Panel B, example trace obtained for an assay performed with 150 µM NADPH.

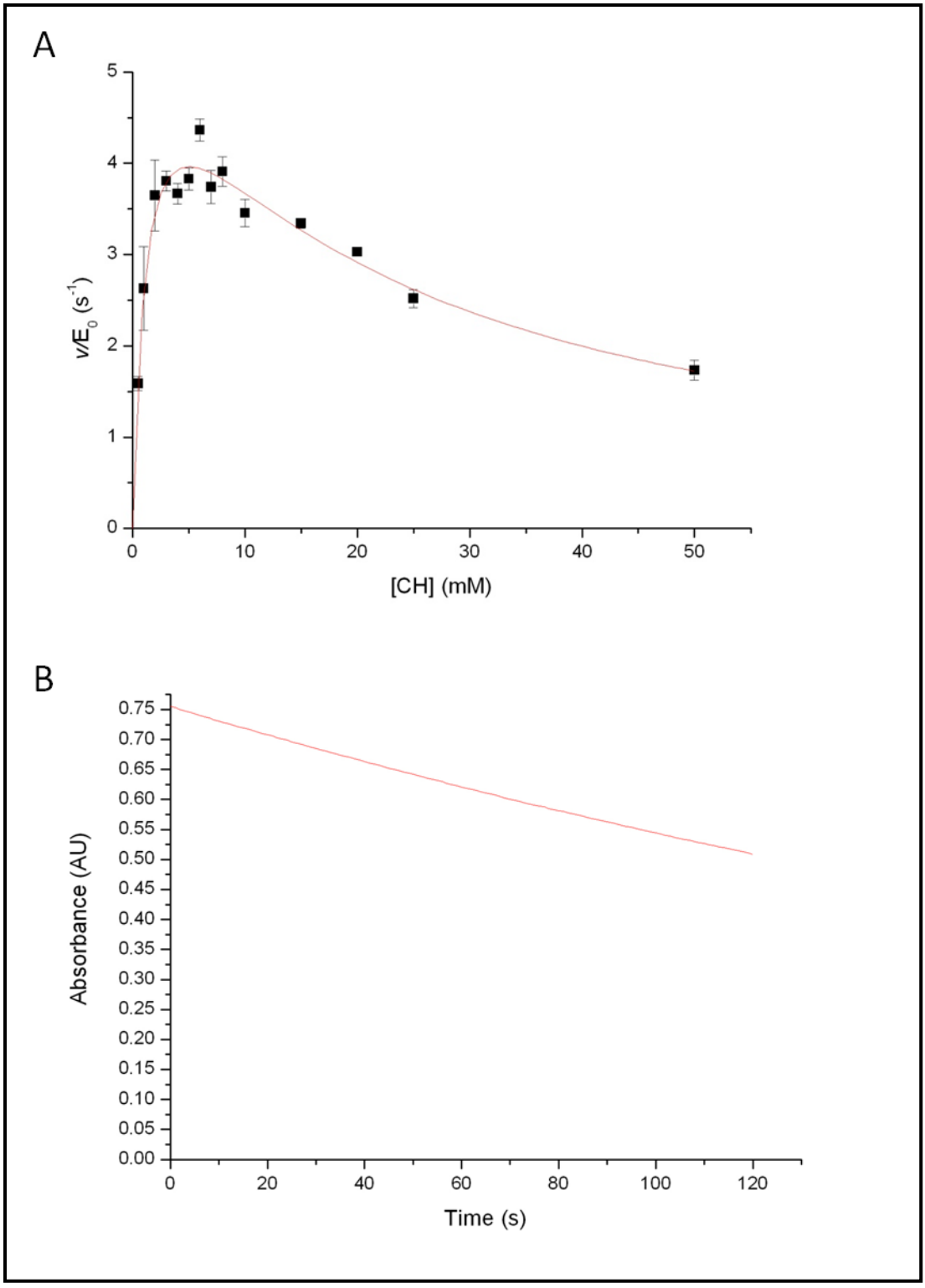


Figure 3.14. The steady-state rate as a function of CH concentration of $PETNR_{WT}$. Data were recorded by monitoring the rate of NADPH oxidation at 340 nm over a maximum period of 120 s. Data were recorded at 0.1 µM PETNR, 150 µM NADPH, 25 °C in 50 mM KH_2PO_4/K_2HPO_4 pH 7. Panel A, non-linear least squares fit to the Michalis-Menten equation incorporating substrate inhibition (Section 2.2.3.2, Equation 2.3) and kinetic parameters were estimated; Panel B, example trace obtained for an assay performed with 6 mM CH.

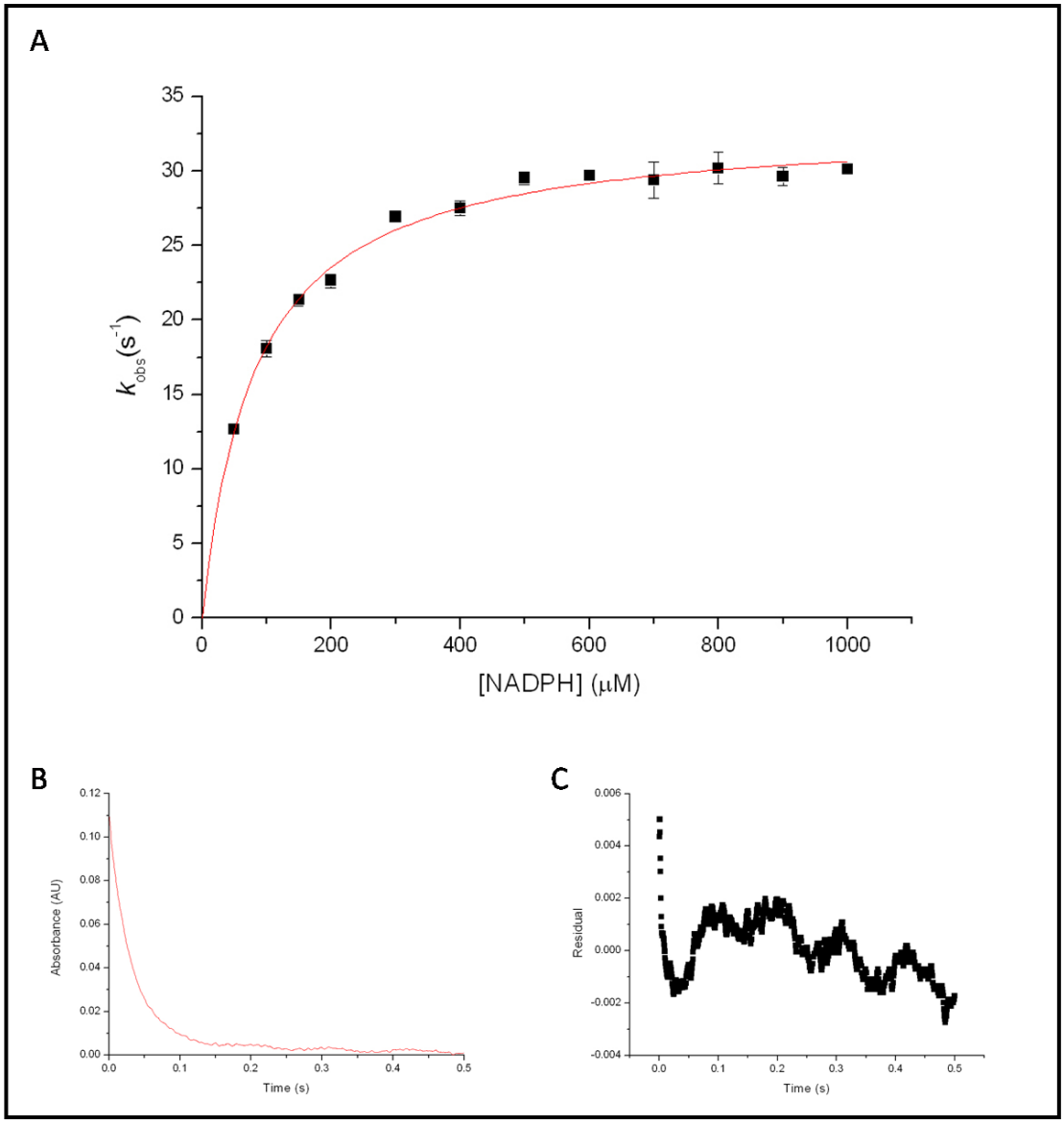


Figure 3.15. The pre steady-state reductive half-reaction of $PETNR_{WT}$. Data were recorded by monitoring the rate of FMN reduction at 464 nm over a period of 1 s following rapid mixing of oxidised enzyme with NADPH. Data were recorded for reactions containing 10 µM PETNR, in 50 mM pH 7 KH_2PO_4/K_2HPO_4 buffer at 25 °C. Panel A shows a non-linear least squares fit to the Strickland equation (Section 2.2.3.3, Equation 2.5) and kinetic parameters were estimated. Panel B, data obtained for a reaction with 800 µM NADPH, red line shows the fit to a single exponential equation (Section 2.2.3.3, Equation 2.4); Panel B, residuals plot of the of the fit shown in panel A.

From the above data, the apparent kinetic parameters associated with the reaction of $PETNR_{His}$ with NADPH and CH were estimated as k_{cat} = 8.7 ± 0.4 s^{-1}, K_M^{NADPH} = 115.2 ± 11.3 µM for the steady-state turnover at fixed CH concentration (6.5 mM). For steady-state reactions at fixed NADPH concentrations (150 µM) k_{cat} = 5.9 ± 0.4 s^{-1}, K_M^{CH} = 1.2 ± 0.2 mM. Substrate inhibition was observed and the value of K_i^{CH} =

21.0 ± 3.5 mM. From the single turnover stopped-flow studies of the reductive half-reaction the flavin reduction rate was calculated as k_{red} = 33.1 ± 0.4 s^{-1} and the apparent dissociation constant was calculated as K_D = 81.8 ± 4.7 μM. The results above are consistent with the data derived previously for $PETNR_{WT}$ (Table 3.2) suggesting the enzyme variants are largely catalytically indistinguishable

Experiment		Parameter	$PETNR_{WT}$	$PETNR_{His}$
Steady-state	Fixed CH	k_{cat} (s^{-1})	8.5 ± 0.8	8.7 ± 0.4
		K_M^{NADPH} (μM)	128.1 ± 24.5	115.2 ± 11.3
	Fixed NADPH	k_{cat} (s^{-1})	7.5 ± 0.3	5.9 ± 0.4
		K_M^{CH} (mM)	1.8 ± 0.2	1.2 ± 0.2
		K_i^{CH} (mM)	20.4 ± 1.9	21.0 ± 3.5
Single turnover	Reductive half-reaction	k_{red} (s^{-1})	34.9 ± 0.4	33.1 ± 0.4
		K_D (μM)	70.9 ± 4.6	81.8 ± 4.7

Table 3.2. Summary table comparing kinetic parameters for $PETNR_{WT}$ and $PETNR_{His}$

3.3.5 $PETNR_{His}$ crystal structure

In order to observe any changes in the structure of $PETNR_{His}$, which could lead to subtle changes in catalytic profile, the structure of the $PETNR_{His}$ variant was determined to 1.4 Å resolution (statistics described in Table 3.3). This work was performed in collaboration with Dr Helen S. Toogood (University of Manchester).

Apart from the obvious exception of the C-terminal residues the structure of $PETNR_{His}$ is almost identical to the previously solved structure of $PETNR_{WT}$ [40] (Figure 3.18). From the omit map $|F_0|-|F_c|$ (Figure 3.16), it is clear that on removing the His tag residues from the model aberrant density is present which corresponds to one full and one partial histidine residue. Hence, the first two residues of the tags were included in the structure. A number of interactions are observed incorporating tag residues, solvent water molecules, other PETNR residues and residues from symmetry related molecules in the crystal. An overview of the PETNR active site (Figure 3.18) shows very little change in any of the active site residues, which would potentially lead to large changes in the catalytic profile.

Parameters		$PETNR_{His}$
Space group		$P2_12_12_1$
Cell dimensions	a, b, c (Å)	58.96, 68.95,88.48
	α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)		37.2-1.4 (1.48-1.40)
R_{merge} (%)		6.7 (29.0)
$I/\sigma I$		12.7 (3.8)
Completeness (%)		99.5 (99.9)
Redundancy		3.4 (3.4)
Unique reflections		71252
R_{work}/R_{free}		12.6/15.6 (16.2/20.2)
RMS deviations	Bond angles (°)	1.436
	Bond lengths (Å)	0.012
Ramachandran plot	Allowed region (%)	95.7
	Additionally allowed region (%)	4.0

$R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl)-[I(hkl)]|/\Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th observation of unique reflection *hkl*.
Redundancy = total number of reflections / total unique reflections.
$R_{work} = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are observed and model structure factors, respectively. R_{free} was calculated by using a randomly selected set (5%) of reflections.

Table 3.3. X-ray crystallographic data collection and refinement statistics for $PETNR_{His}$.

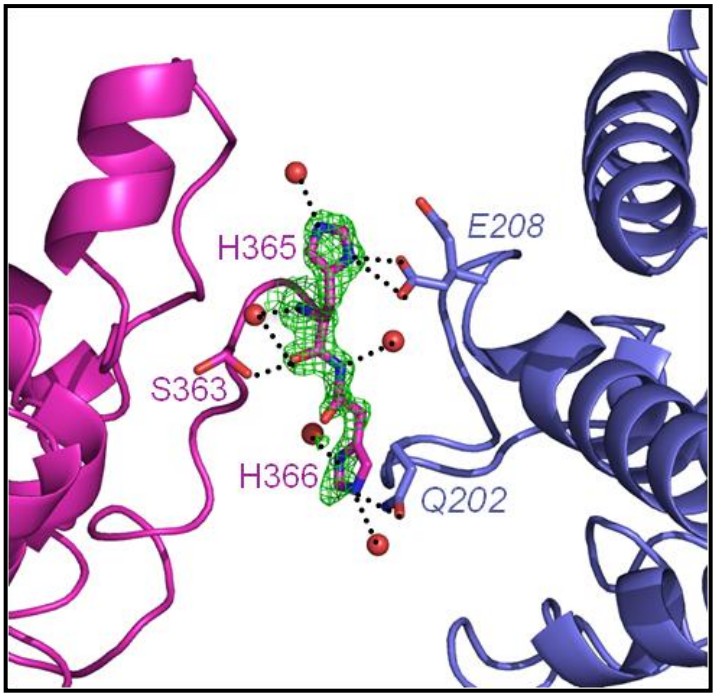


Figure 3.16. His tag model including electron density and its interactions with PETNR monomer and its symmetry related molecule in the crystal. Magenta, PETNR monomer; blue, symmetry related PETNR monomer; green, $|F_0|$ - $|F_\sigma|$ map showing aberrant electron density associated with the histidine tag.

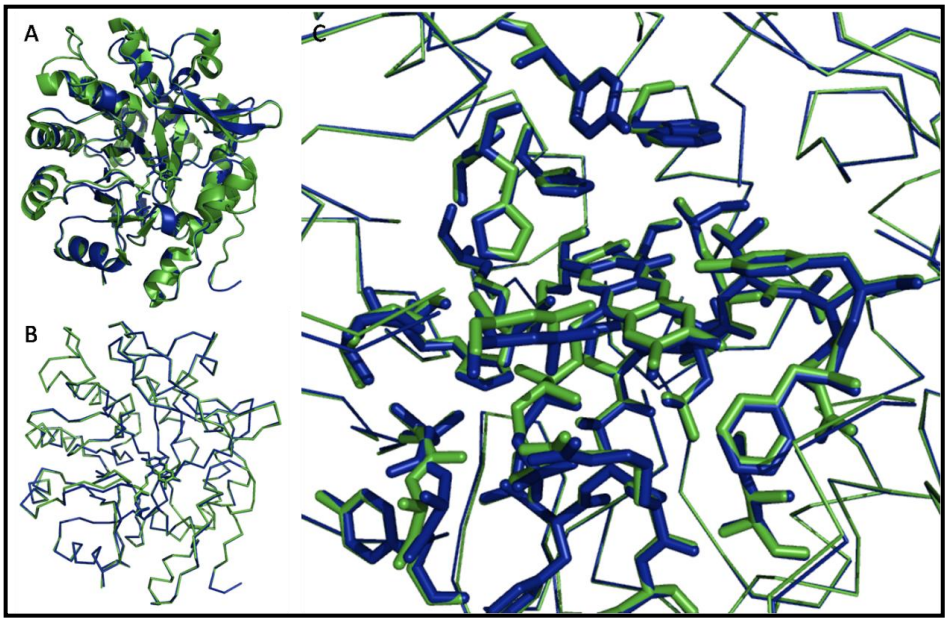


Figure 3.17. Alignments of $PETNR_{His}$ (blue, as derived in this study) and $PETNR_{WT}$ (green, 1GVQ [38]). Panel A, alignment of cartoon representations of the structure; Panel B, alignment presented as C_α backbone ribbon showing tiny variations in the backbone; Panel C, active site residues and FMN revealing virtually no change in residues implicated in catalysis and mechanism.

3.4 Biotin Tagged PETNR ($PETNR_{Bio}$)

A PETNR construct with the biotin ligase recognition sequence was constructed to provide an alternative to $PETNR_{His}$ for the robotic purification of library clones. Additional advantages include purification in a plate format potentially allowing the immobilisation and assay of library clones in a single microplate.

The biotin tag is a specific sequence of amino acids which undergoes a post-translational biotinylation reaction catalysed by the enzyme biotin ligase [166]. Biotin is added to the terminal amino group of a specific lysine residue surrounded by the consensus sequence GLNDIFEAQ**K**IEWHE [166]. Bacterial strains are available that contain the biotin ligase gene stably inserted into the hosts genomic DNA under the control of an arabinose-inducible vector. Transformation of the gene of interest into these strains, in an IPTG-inducible vector, enables independent control of expression of both the biotin ligase and the target protein (as described at http://www.avidity.com/). Supplementing the growth media with biotin enables the post-translational biotinylation reaction to proceed *in vivo*.

Biotin-tagged proteins can be affinity purified using streptavidin-agarose beads on the basis of the strong interaction between biotin and the egg white protein avidin (or its bacterial counterpart streptavidin) [161]. The association constant for this interaction is in the region of 10^{15} M^{-1}. Following immobilisation, the protein of interest can be eluted using competing concentrations of soluble biotin.

3.4.1 $PETNR_{Bio}$ construction

Due to the length of the biotin tag, cassette-based incorporation (as described in Section 2.2.4.3), was chosen to avoid extensive PCR reactions, screening of colonies and sequencing. Following annealing and ligation steps, only one clone from a total of 41 colonies produced a significantly larger CPPCR product (Figure 3.18). This PCR product is approximately 350 bp, corresponding to the larger expected size of a full-length biotin tag incorporated PETNR construct.

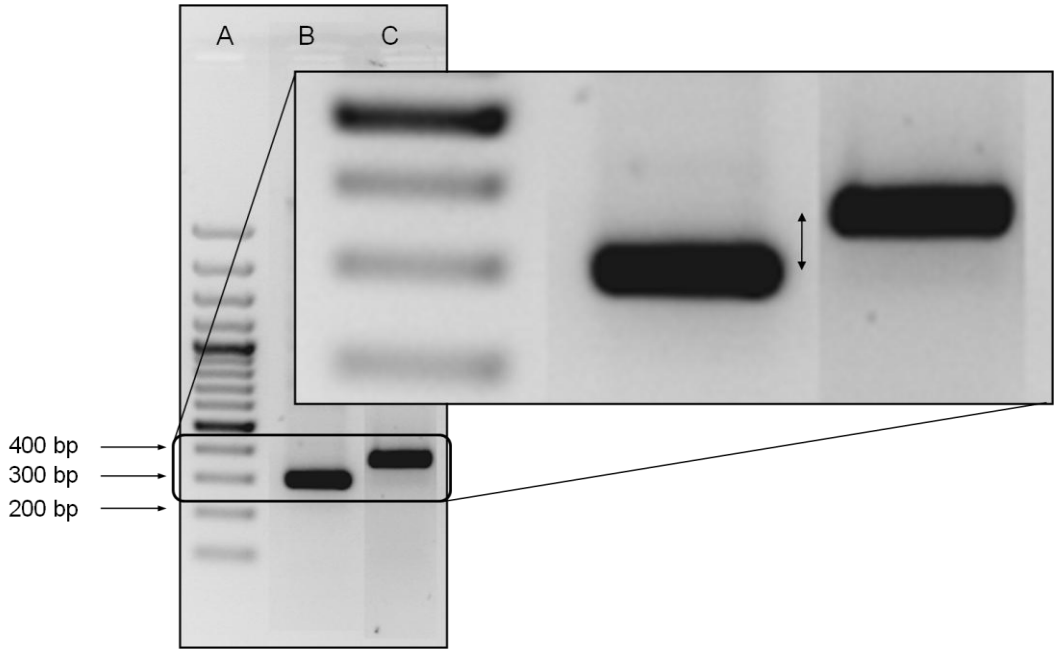


Figure 3.18. Agarose gel showing the PCR products from CPPCR of one biotin tag clone. Lane A, 100 bp Ladder (NEB; 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bp, 1.2, 1.5, 2, 3 kbp); Lane B, control reaction with wild-type template (pONR1, pBluescript SK+); Lane C, Biotin7 clone PCR product. Analysis was performed on a 1.3 % agarose gel electrophoresed at 100 V for 1 - 1.5 hr.

As the biotin tag cassette contains complementary sticky ends to the same restriction enzyme, there is a 50 % chance the insert could be incorporated the wrong way around. Sequencing reactions of the Biotin 7 clone at the 3' end of the gene showed the presence of an 18 amino acid tag, incorporated in the correct orientation (Figure 3.20). However, it was found that this clone also contained a mutation, G1138A, which corresponded to an amino acid change of E380K. This mutation is located within the incorporated biotin tag, suggesting the mutation may be the result of a mistake in the sequence of the incorporated primers. As this tag sequence has been optimised for biotin incorporation [166], and post-translational biotin incorporation occurs at the lysine residue only two amino acids away from this mutation, the addition of a second lysine residue close to the key lysine may affect biotin incorporation. Therefore this clone was used as a template for site-directed mutagenesis to correct the mutation. Biotin7 was further sequenced and was found to contain the additional silent mutation/isoform A921G, as discussed in Section 3.3.1.

```
Nucleotide sequence

Predicted      GACTACCCTTCACTGGAATGCCAGGGCCTGAACGATATTTTTGAAGCGCAGAAAATTGAA
PETNR Biotin7  GACTACCCTTCACTGGAATGCCAGGGCCTGAACGATATTTTTGAAGCGCAGAAAATTAAA
               ********************************************************** **

Predicted      TGGCATGAATAATCCCGCTTTGTACATTGATAGCGGCGACCTTTCGCCGCTATACTAAAA
PETNR Biotin7  TGGCATGAATAATCCCGCTTTGTACATTGATAGCGGCGACCTTTCGCCGCTATACTAAAA
               ************************************************************

Amino acid sequence

Predicted      D Y P S L E C Q G L N D I F E A Q K I E W H E Stop
PETNR Biotin7  D Y P S L E C Q G L N D I F E A Q K I K W H E Stop
```

Figure 3.19. Clustal alignment of the 3' sequence of the deposited $PETNR_{WT}$ encoding DNA sequence with the resultant Biotin7 sequence [164]. Cyan, tag nucleotide sequence; red, error in the tag sequence; underline, stop codon.

Following mutagenesis to repair of the unexpected mutation, the product was isolated from the agarose gel (Figure 3.20), and sequencing was repeated to confirm the correct incorporation of the full recognition sequence (Figure 3.21).

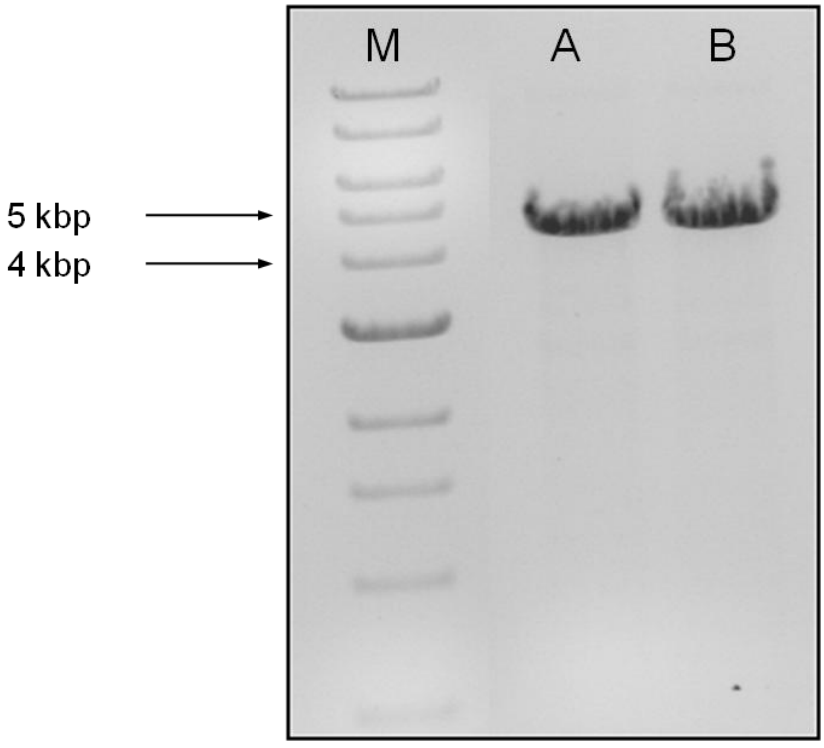


Figure 3.20. Agarose gel showing the PCR product from the PCR designed to correct the sequence of the biotin tag. Lane A, 1 kbp ladder (NEB; 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 kbp); Lane B, $PETNR_{WT}$ plasmid control; Lane C, Biotin repair PCR product. Analysis was performed on a 0.6 % agarose gel at 100 V for 45 min.

Figure 3.21. Clustal alignment of the 3' sequence of the deposited $PETNR_{WT}$ encoding DNA sequence with the Biotin repair PCR product [164]. Cyan, tag nucleotide sequence; red, the position of the error in the tag sequence; underline, stop codon.

3.4.2 $PETNR_{Bio}$ double induction trials

Expression trials of $PETNR_{Bio}$ in *E. coli* strain AVB100 were carried out by varying both the concentration of IPTG and arabinose in search of conditions optimal for protein expression and *in vivo* biotinylation, respectively (Figure 3.22). Samples were grown with induction ranging from 0 – 0.4 mM IPTG and from 0 – 0.5 % arabinose followed by the immobilisation of tagged protein using avidin beads using a batch purification protocol (Section 2.2.4.6).

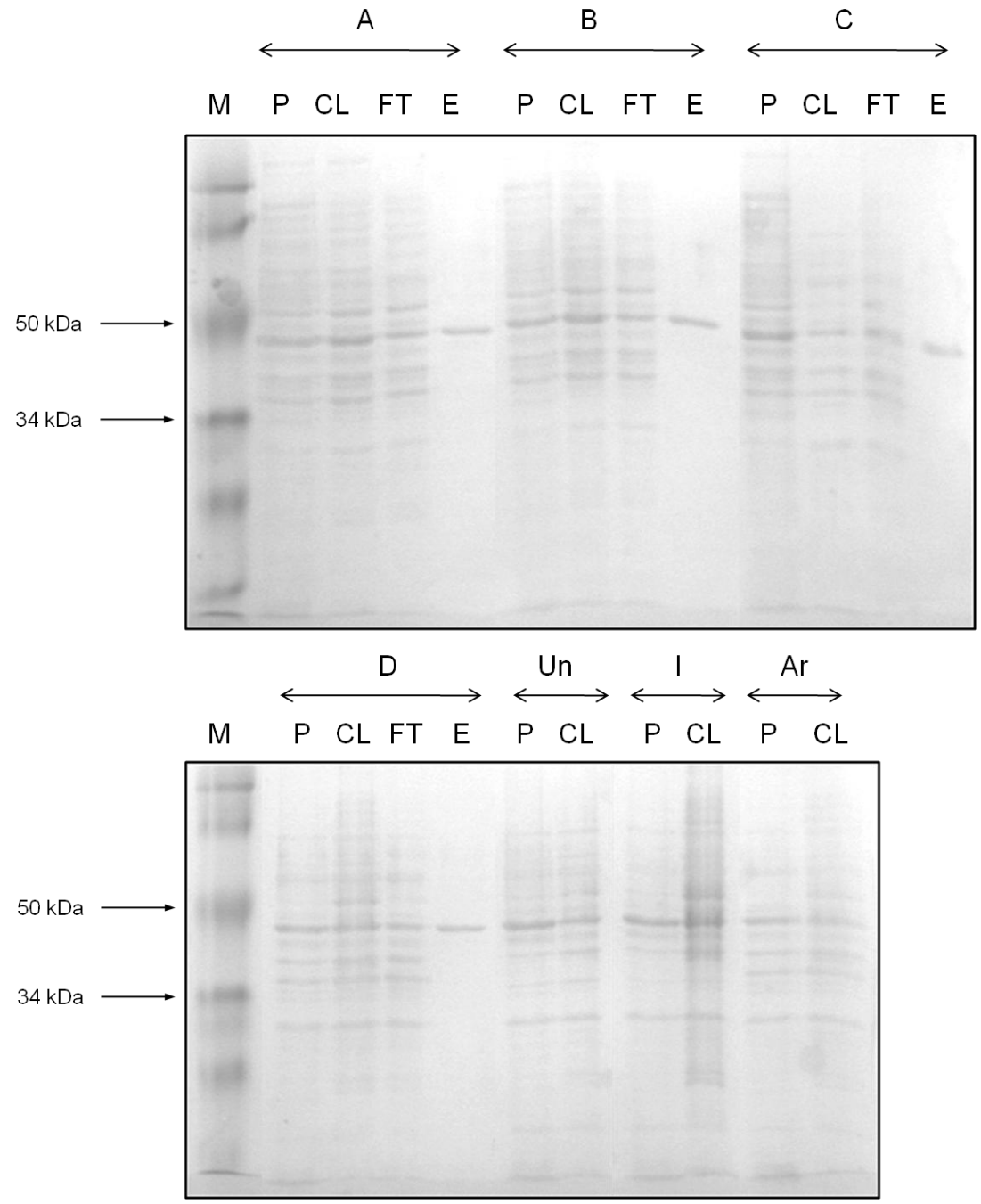


Figure 3.23. Expression and purification trials of $PETNR_{Bio}$ at variable IPTG (induction of $PETNR_{Bio}$) and arabinose (induction of Biotin Ligase) concentrations. Lane M, low range prestained protein ladder (bio-rad; 21, 29, 34, 50, 77, 103 kDa); Lane A, 0.05 mM IPTG, 0.1 % Arabinose; Lane B, 0.05 mM IPTG, 0.5 % Arabinose; Lane C, 0.4 mM IPTG, 0.1 % Arabinose; Lane D, 0.4 mM IPTG, 0.5 % Arabinose; Un, Un-induced; I, IPTG only 0.4 mM; Ar, Arabinose only 0.5 %. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

In spite of the range of conditions trialled, $PETNR_{Bio}$ was poorly expressed in all cases compared to both $PETNR_{WT}$ and $PETNR_{His}$. The levels of $PETNR_{Bio}$ in the monomeric avidin eluate were used to estimate the levels of biotinylation across different induction condition. The biotinylation levels were found to be approximately invariant for different growth conditions as demonstrated by the presence of similar amounts of $PETNR_{Bio}$ in the eluate and the pellet of all experiments. Thus, fine control as discussed in the literature [167] does not appear to have much influence on the biotinylation levels achieved. Biotin ligase (~30 kDa), the enzyme responsible for the *in vitro* biotinylation, is not visible in any fractions derived from the purification protocol suggesting it is expressed at low levels.

Due to the lack of success in expressing $PETNR_{Bio}$ at high levels and given the apparent success of the initial $PETNR_{His}$ purification trials, the His-tagged variant was carried forward for further study.

3.5 Robotic methods optimisation

3.5.1 Robotic colony picking

Colony picking on the robotics platform was fully automated using the 8 channel pipettor to transfer colonies from agar plates to 96 well deep well blocks pre-prepared with media. Efficient robotic colony picking was achieved when the agar plates contained well separated colonies (~150 – 200 per plate) of uniform size and shape. This enabled the software to pick the 'best' 96 colonies as defined by the colony selection criteria (min/max colony size and closeness to an 'ideal' circular colony). Due to variation in the transformation reactions, consistent numbers of colonies were difficult to obtain, so library transformation slurries were plated by serial dilution, and the plate with the appropriate colony density was selected. Libraries were grown and archived as glycerol stocks as described in Section 2.2.4.10.1.

3.5.2 Optimisation of *E. coli* growth in 96 well format

Critical to the success of the screening strategy is the growth and expression of protein libraries. To ensure experimental conditions are optimal for expression, a range of growth media and *E. coli* strains were trialled. Initially, 96 well cultures were grown for 12 h at 37 ˚C followed by a 4 hour induction with IPTG at a final concentration of 0.4 mM. This approach was considered sub-optimal due to a number of factors. First, inducing cells after a fixed time period does not take into account the variation of cell densities of individual variants at the time of induction which is often critical in determining the level of over expression. Second, short induction times often produce large amounts of non-flavinated protein as *in vivo* flavin production is slower than protein production. Third, mutations are often toxic and strong induction protocols often retard cell growth and lead to cell lysis.

To potentially counteract these problems, auto induction media was trialled. Auto induction media provides a blend of glucose and lactose to the growing cells. During initial periods of growth glucose is used preferentially until its supply is exhausted as glucose is a known *Lac* operon repressor. This prevents protein expression at low

cell density. Once sufficient cell density is obtained, indicated by the utilisation of all glucose, the cells then switch metabolism over to utilise lactose thus activating the *Lac* operon and inducing the protein of interest [168]. This type of medium provides a slow, steady means of protein expression tailored to each individual culture, increasing the likelihood of producing flavin-containing enzymes and minimising the occurrence of inclusion bodies.

In addition, a variety of *E. coli* strains, including some designed specifically for protein expression, were trialled to determine the strain providing the highest and most consistent levels of expression of $PETNR_{His}$.

3.5.2.1 Library expression with auto induction media

Following initial un-induced growth and glycerol stock storage as detailed in Section 2.2.4.10.1, libraries were re-grown, by inoculation from glycerol stock, using auto induction media. Trials of different auto induction media (Section 2.1.2) were conducted to determine the best media required to generate consistent protein expression levels. Protein expression levels were characterised by SDS PAGE analysis (Figure 3.24).

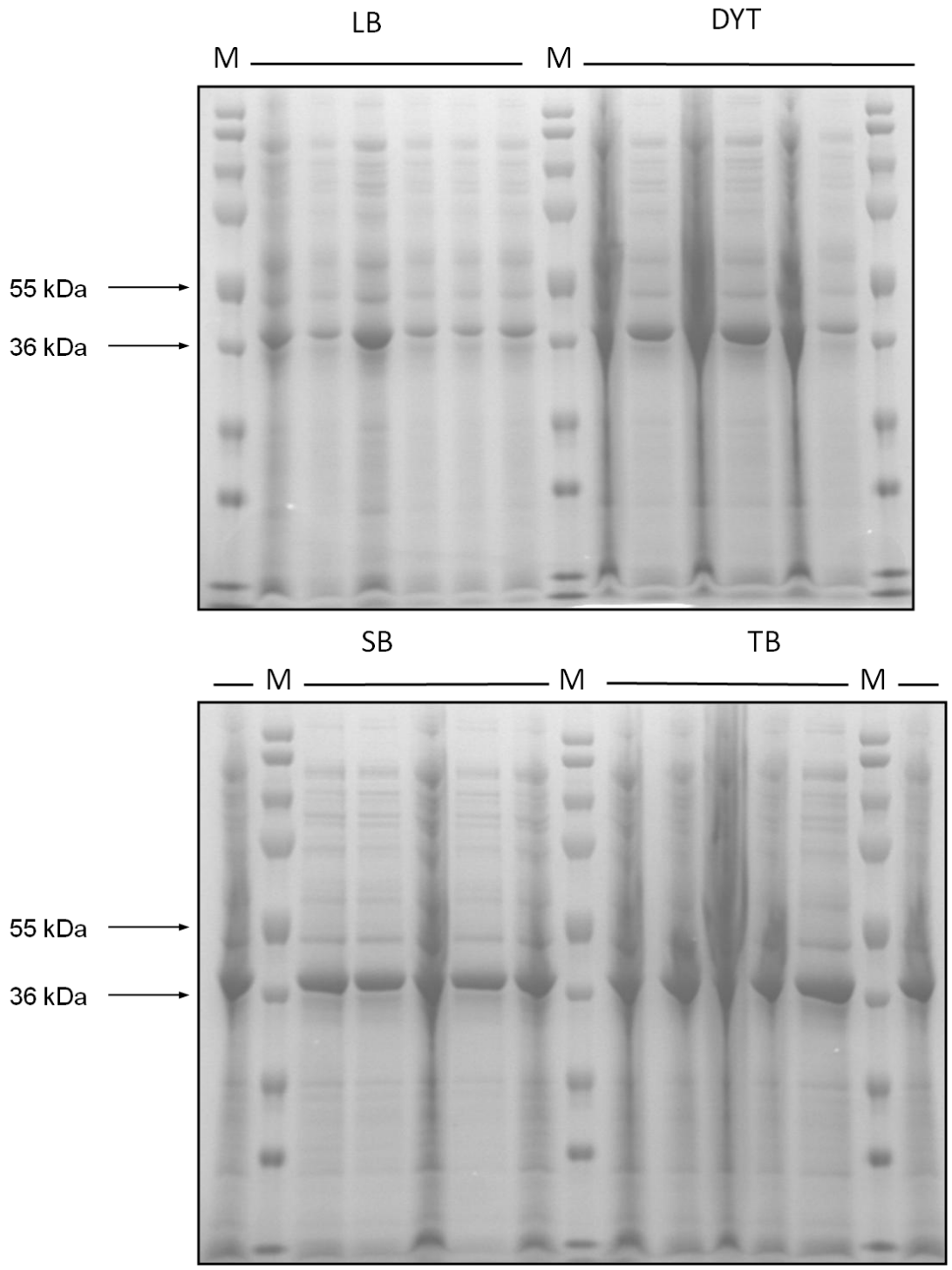


Figure 3.24. Expression trial testing a variety of auto-induction media for use in growing mutant libraries. Lane M, Pageruler plus prestained protein ladder (Fermentas, 10, 17, 28, 36, 55, 72, 95, 130, 250 kDa); LB, Luria Broth AIM growths; DYT, Double Yeast Tryptone AIM growths; SB, Super Broth AIM growths; TB, Terrific Broth AIM growths. All cultures were $PETNR_{His}$ in JM109 cells grown for 24 h. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

The results indicate that all media tested yield expression of $PETNR_{His}$. LB AIM being the least rich of the media showed significantly weaker expression levels than the other media. Growth in DYT AIM and SB AIM media show vastly enhanced expression levels compared to LB AIM. However DYT AIM showed a large degree of variation in expression. TB AIM appears to show a further enhanced degree of expression. Due to the far greater expression levels, TB was selected for further investigation in the optimisation of the cell growth and protein expression trials.

3.5.2.2 Library expression from different E. coli strains

Library DNA (Section 2.2.4.9) was transformed into a range of *E. coli* strains before colonies were picked individually and grown in a 96 well format. Individual colonies were picked from transformation plates of the T26X library and grown in 1 ml LB AIM for 24 h at 37˚C. A random selection of colonies from each plate were analysed by SDS PAGE (Figure 3.25). It is important to note that this experiment was performed with library DNA to assess the consistency of transformation and expression with a range of mutants.

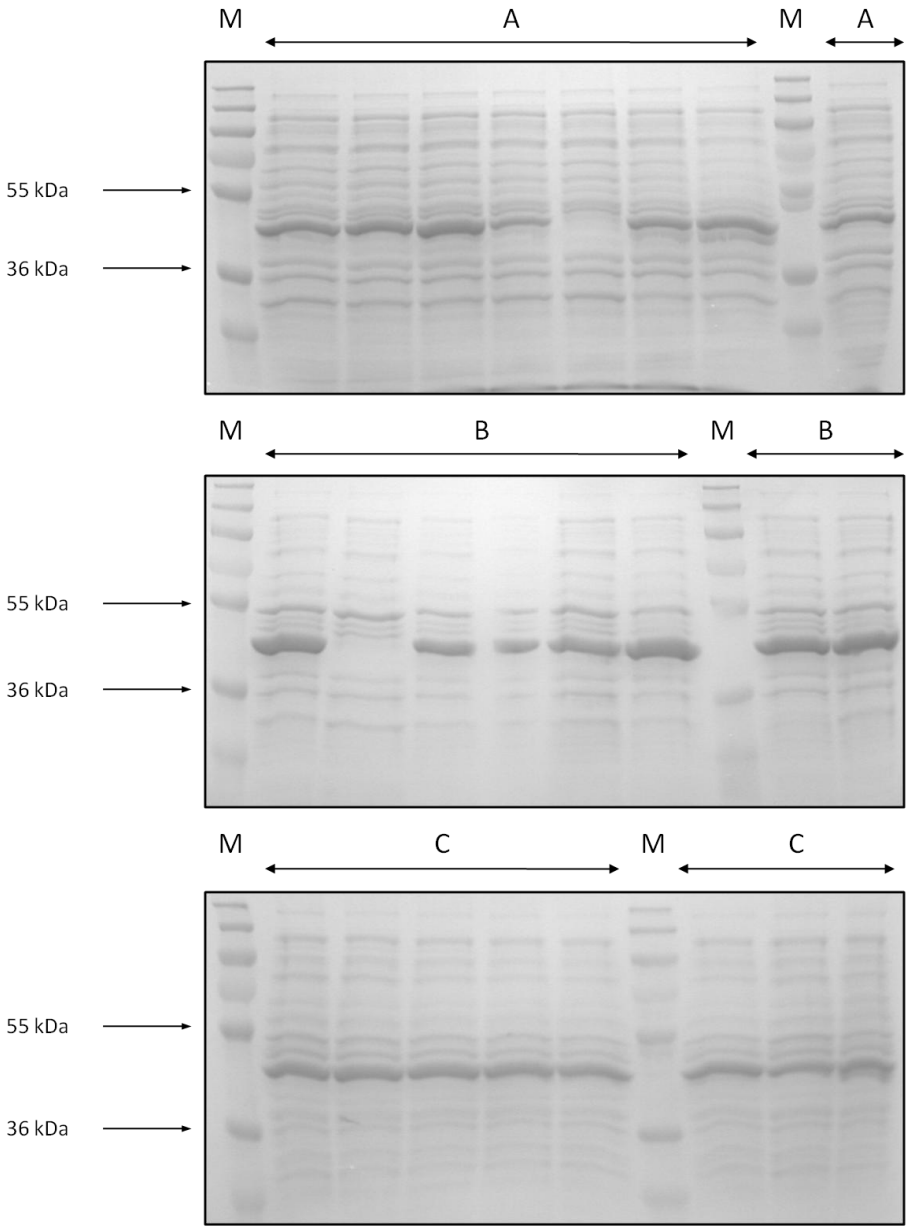


Figure 3.25. Expression trial of a variety of *E. coli* expression strains transformed with T26X library with TB AIM. Lane M, Pageruler plus prestained protein ladder (Fermentas, 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); lanes A, XL10 Gold; lanes B, XL1 Blue; lanes C, JM109. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

All three cell strains demonstrated substantial over-expression in the majority of cases, with XL1 Blue showing the best levels of expression. In the case of XL1 Blue, protein expression was variable with some clones not producing an observable protein band for PETNR. Strain JM109 shows a strong over-expressed band for $PETNR_{His}$ and expression at these levels is consistent across the sampled colonies,

consequently this strain was taken forward to further studies. It is important to note that all three strains are individual transformations of the same library DNA and as such individual lanes do not correspond to the same variant, hence individual lanes are not necessarily comparable between gels. In this case, it is assumed that the colonies picked at random represent an average of the entire library.

3.5.3 Robotic cell lysis and clarification optimisation

An appropriate cell lysis method needed to be developed for the robotic protocols. Conventional techniques such as sonication or French press with accompanying centrifugation steps are not amenable to an anaerobic 96 well protocol. Therefore the use of chemical lysis approaches resulting in the production of lysates not requiring clarification by centrifugation was trialled to optimise protein recovery.

A variety of commercial and tailored lysis buffers were trialled and assessed for their ability to completely lyse cells to yield the active protein. All buffers contained additionally lysozyme, benzonase and protease inhibitor tablets. Lysates were purified using a standard Ni - NTA batch purification protocol (Section 2.2.4.4), analysed by SDS PAGE to assess the available protein content of the eluate (Figure 3.26), and assayed for activity with the PETNR model substrates NADPH and cyclohexenone.

The presence of $PETNR_{His}$ was detected in all the pellets and all the cell lysates indicating that the same level of PETNR was present in all induced cells. No visible pellet was present after lysis with commercial lysis buffers (bugbuster, Cell Lytic B). However, with both of the tailored solutions (Tween 20 and Triton X-100) small, but significant, orange pellets were observed following lysis, indicating the need for a centrifugation step. SDS PAGE analysis of all lysates showed that large quantities of protein were liberated from cells and were available for binding to the resin.

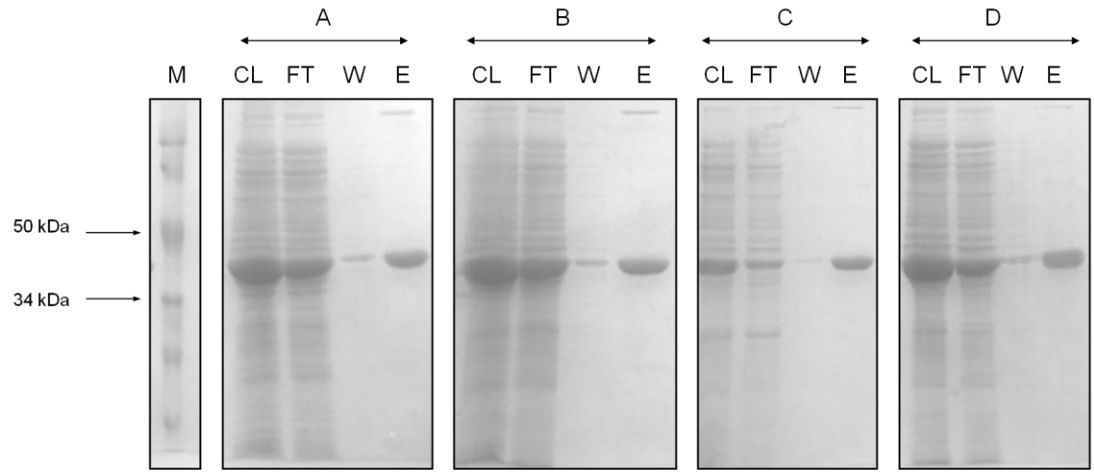


Figure 3.26. The effectiveness of different lysis buffers to produce clarified lysates. Lane M, low range prestained protein ladder (bio-rad; 21, 29, 34, 50, 77, 103 kDa); Lanes A, bugbuster; Lanes B, cell lytic B; Lanes C, Tween 20 buffer; Lanes D, Triton X-100 buffer. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

Kinetic assays of the eluates from the purification experiment indicated that there was only a partial correlation between visual protein levels and activity detection. Bugbuster was found to produce the highest level of active protein, so was selected as the lysis buffer for the robotic protocols.

Lysis buffer	Initial rate (µM s^{-1})
Water	0.019
Tween – 20	0.009
Triton X – 100	0.085
Cell lytic A	0.173
Bugbuster (50 %)	0.231

Table 3.4. Anaerobic turnover assays with 6 mM CH and 150 µM NADPH. Data were recorded for the reaction catalysed by 20 µl of cell lysate, except for bugbuster since the reaction with this volume was too fast to be recorded (10 µl was used in this case).

3.5.4 Robotic 96 well purification trial

Following optimal lysis it was next necessary to develop a micro-scale purification method which yielded the substantial amounts of purified protein to ensure the maximal amount of assays per purification. Importantly consistency in purification

is also needed to ensure results are comparable across the 96 well plate. Potential types of purification assessed included purification with magnetic beads (Section 2.2.8.1) and immobilisation onto a coated microplate surface (Section 2.2.8.2).

In the case of plate purifications, assays were added directly into the wells of the microplate and performed according to the method discussed previously with buffer added instead of enzyme (Section 2.2.4.10.6).

	Average rate (nM s^{-1})	Standard Deviation (nM s^{-1})	Error (%)
His Mag magnetic beads (Novagen)	206.8	55.4	27
Talon Magnetic beads (Clonetech)	92.0	44.9	49
His grab HBC plates (Pierce)	13.9	3.5	25

Figure 3.5. Anaerobic turnover assays with 6 mM CH and 150 µM NADPH with 50 µl eluate (from beads) and assays were performed in the purification plate (for plate purification). Values represent average and standard deviation of 96 wells.

A further consideration is that plate purification methods are not reusable and only provide enough protein for a single assay. However beads provide a volume of eluate which can potentially be used for multiple reactions and the beads can be regenerated and reused for multiple experiments.

The results indicated that the plate purification method resulted in lower activity detection compared with the bead methods. 'His-Mag' magnetic beads (Novagen) provided a far superior purification option with a 2-fold increase in the activity when compared with the next best method, represented by the talon beads, with a lower error when considered in the context of total enzyme.

This leads to the identification of His-Mag beads as the most practical method for micro-scale purification.

3.5.5 Robotic 96 well assay trial

Ultimately, the identification of improved variants is via the detection of an increased steady-state reaction rate. Therefore, the 96 well plate assay method must be optimised to detect both fast and very weak activities. The latter is

necessary when looking for activity with very poor substrates of the wild-type enzyme. To ensure this, a relatively long reaction time was employed. Activity was determined by monitoring NADPH oxidation at 340 nm in the presence of an oxidising substrate.

3.5.5.1 Microplate apparent path length calculation

Due to the requirement to calculate accurate rates for assay, it was necessary to derive an accurate value for the apparent path length used in enzyme assays. The absorbance of known concentrations of NADPH in a constant volume (300 µl) were recorded to determine the apparent path length (l_{app}) of solution in microplate wells (Figure 3.27).

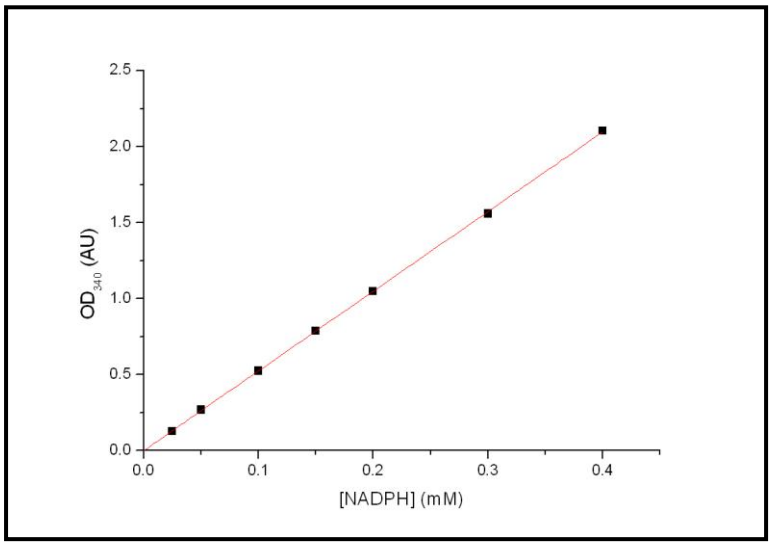


Figure 3.27. Standard curve of known NADPH concentrations and the absorbance given in the plate reader.

Analysis of the linear fit to the data and applying beers law ($A = \varepsilon c l$) yielded a l_{app} of 0.84 cm, which was used in all subsequent analyses.

3.5.5.2 PETNR$_{His}$ concentration dependence of reaction rate

In order to allow estimation of the concentration of PETNR present in purifications, a standard curve was recorded (Figure 3.28) of the steady-state rates as a function of enzyme concentration. This graph should allow the estimation of the enzyme

concentration from purifications by performing standard assays with NADPH and CH.

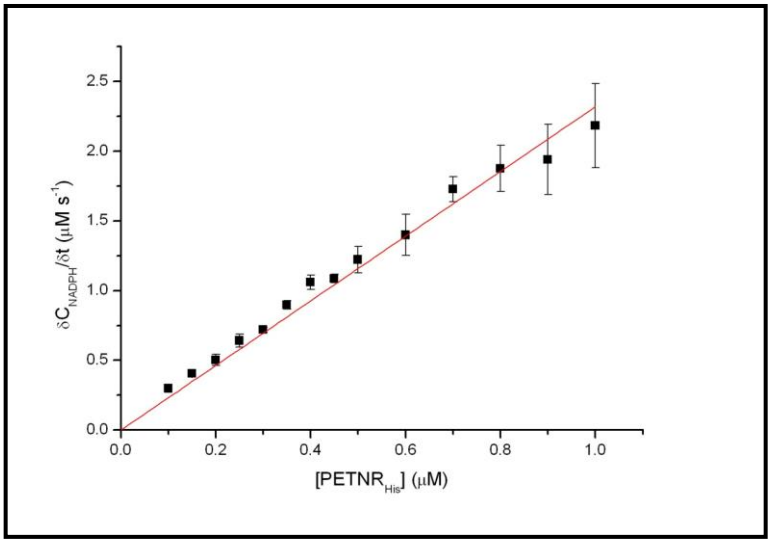


Figure 3.28. Calibration curve used to estimate the concentration of $PETNR_{His}$. Rates were determined using 0.1 mM NADPH, 6.5 mM CH in 50 mM pH 7 KH_2PO_4/K_2HPO_4 buffer for 10 min at 25 ˚C.

Results indicated a strong linear correlation between the amount of enzyme present in the reaction and the steady-state reaction rates from the turnover of NADPH and CH and will be used in later sections to extrapolate and estimate of the protein concentration.

3.5.5.3 Steady-state rate as a function of CH concentration of $PETNR_{His}$ recorded in a 96 well format on a microplate reader.

To assess the accuracy and reproducibility of microtitre plate assays, data were recorded for the steady-state reaction rates as a function of CH concentration under identical conditions to the initial characterisation conditions described in Figures 3.4 and 3.14. Calculations were adjusted to account for the differences in path length between conventional spectrophotometric assays and microplate assays.

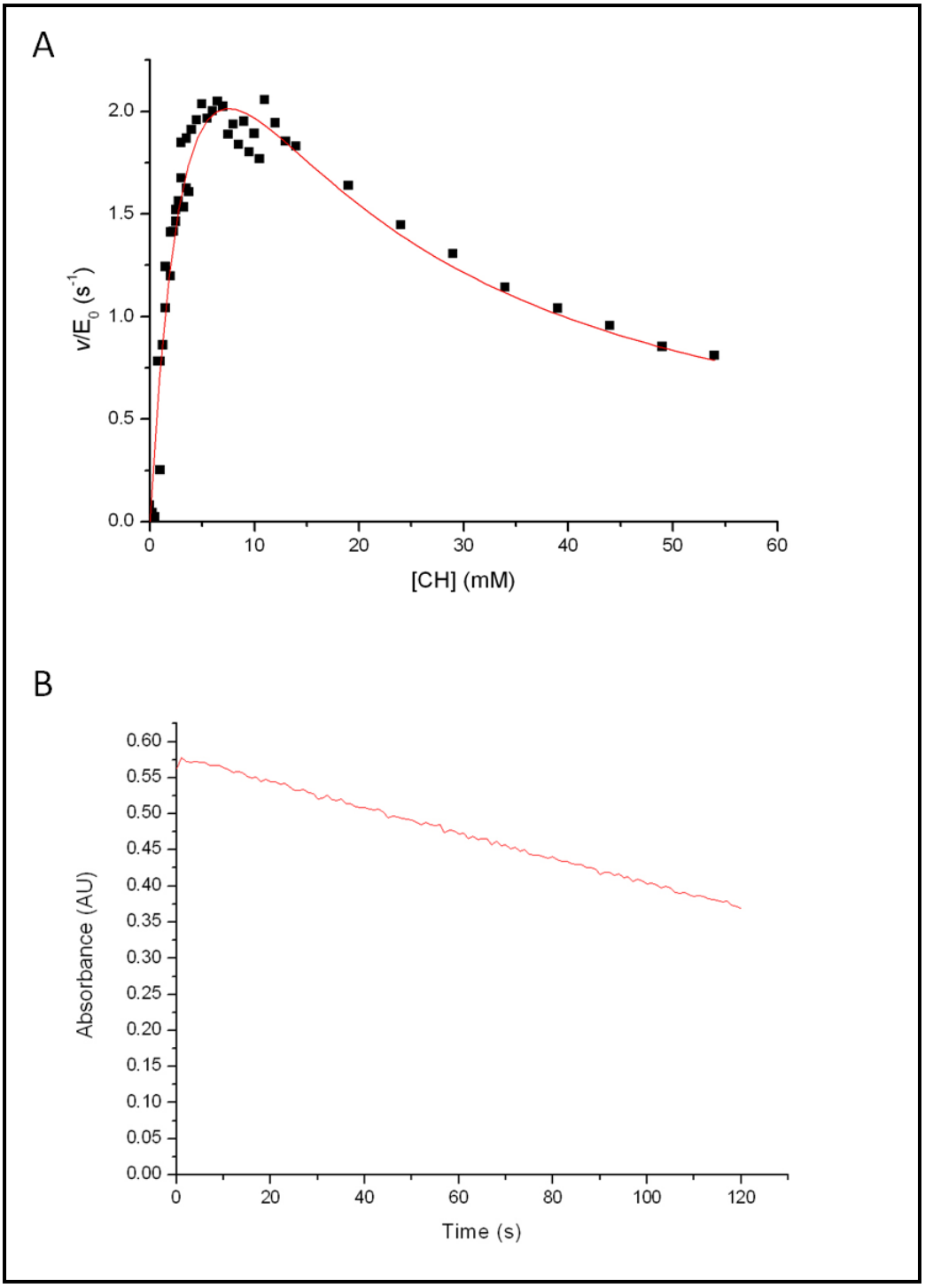


Figure 3.29. The steady-state rate as a function of CH concentration of $PETNR_{His}$ recorded using an anaerobic plate reader. Data were recorded by monitoring the rate of NADPH oxidation at 340 nm over a maximum period of 120 s. Data were recorded at 0.1 µM PETNR, 150 µM NADPH, 25 °C in 50 mM KH_2PO_4/K_2HPO_4 pH 7. Panel A, non-linear least squares fit to the Michalis-Menten equation incorporating substrate inhibition (Section 2.2.3.2, Equation 2.3) and kinetic parameters were estimated; Panel B, example trace obtained for an assay performed with 6 mM CH.

From the above data, the apparent kinetic parameters associated with the steady-state reaction of $PETNR_{His}$ were estimated as k_{cat} = 4.9 ± 0.7 s^{-1}, K_M^{CH} = 5.4 ± 1.1 mM for the steady-state turnover at fixed NADPH concentration (150 µM). Substrate inhibition was again observed and the value of K_i^{CH} = 10.5 ± 2.4 mM, which bears reasonable similarity to the previously recorded data from Section 3.3.4. There are however some significant differences to note, most obviously the maximum recorded rate and k_{cat} are lower than those recorded using the conventional spectrophotometric method. This is most likely due to an increase in the time taken to begin reading reactions. In the spectrophotometric methods reaction initiation, mixing and initial measurements occur in very close order. However, in the plate reader there are several steps which contribute to an increased delay. First there are several steps in the initiation of reaction that are limited by the speed of the plate readers mechanical parts, such as moving the plate into the reader. Second, mixing using the plate reader is less efficient, hence takes longer to complete, than mixing by inversion.

This additional delay results in an increase in the time between initiation of reaction and observation of first data points. This results in the underestimation of the true reaction rate. This underestimation results in the underestimation of v/E_0 and as a result the underestimation of k_{cat}. Those data points with a faster rate are particularly sensitive to this phenomenon.

Lowering the concentration of enzyme or decreasing the mixing time could potentially minimise these problems however generate further problems of their own including too small a change in absorbance to be reliable or poorly mixed reactions respectively.

3.5.6 Control robotic method trials with $PETNR_{His}$

In order to validate the approach and assess the accuracy and reproducibility of the robotics protocols, $PETNR_{His}$ colonies were tested and protein levels, degree of purification and kinetic activity were detected and statistically analysed.

3.5.6.1 Expression trial

Highly consistent protein expression was detected in cultures grown after robotic colony picking and expression in TB AIM (Figure 3.30)

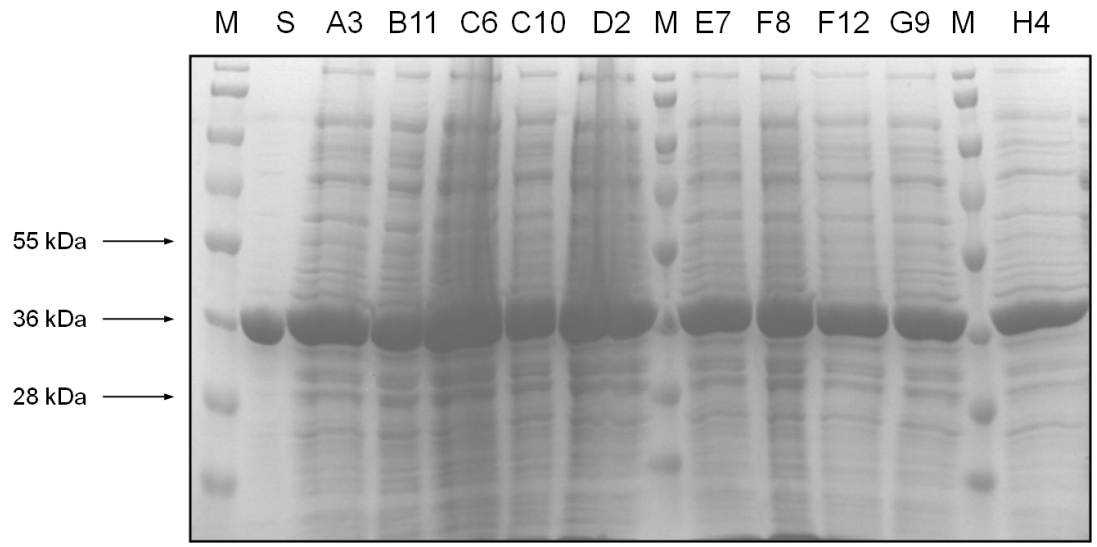


Figure 3.30. Expression trial gel showing the cell pellet of 1 ml cultures $PETNR_{His}$ WT grown in a 96 well format. Lane M, Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250); Lane S, $PETNR_{WT}$ standard; other labels, well from which each sample was taken. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

3.5.6.2 Purification trial

Robotic lysis and purification were performed, as described in the protocols discussed in Section 2.2.4.10, for $PETNR_{His}$ clones. Eluates from the purification were subjected to analysis by SDS PAGE, protein quantitation and standard assay with the NADPH/CH couple.

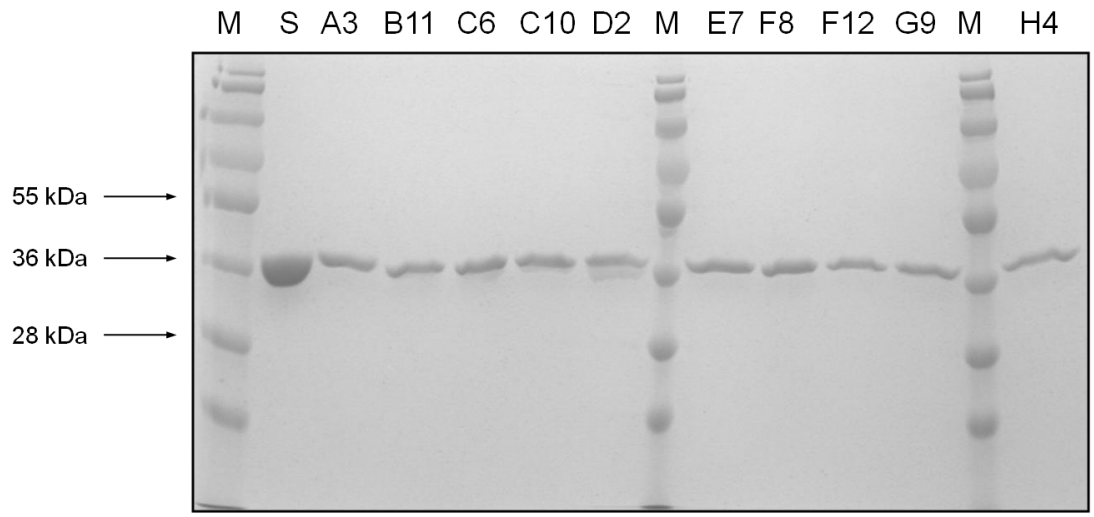


Figure 3.31. Eluate gel from the purification of 1 ml cultures $PETNR_{His}$ WT grown in a 96 well format. Lane M, Pageruler plus prestained protein ladder (Fermentas; 10, 17, 28, 36, 55, 72, 95, 130, 250); Lane S, $PETNR_{WT}$ standard; other labels, well from which each sample was taken. SDS PAGE was performed on a 12 % SDS PAGE gel electrophoresed at 180 V for 45 min. The gel was stained. The gel was stained with Coomassie Brilliant Blue R250.

SDS PAGE analysis indicated an approximately constant level of pure protein in the eluates, which following quantitation was found to be 148 ± 20 µg ml^{-1}. This represents an error associated with the protein concentration of ~14 %. This value is thought to account for errors in the protocol, attributed to growth and purification, and to some extent protein expression.

Protein eluates were assayed using a standard 96 well protocol using NADPH/CH to assess the intra block variability (Figure 3.32). The results demonstrated an average rate of 207 ± 55 nM s^{-1} overall representing a total error of 27 %. A large proportion of this error has been accounted for in the purification error. However, this experiment also suggests that an approximately equal contribution comes from variations within the assay.

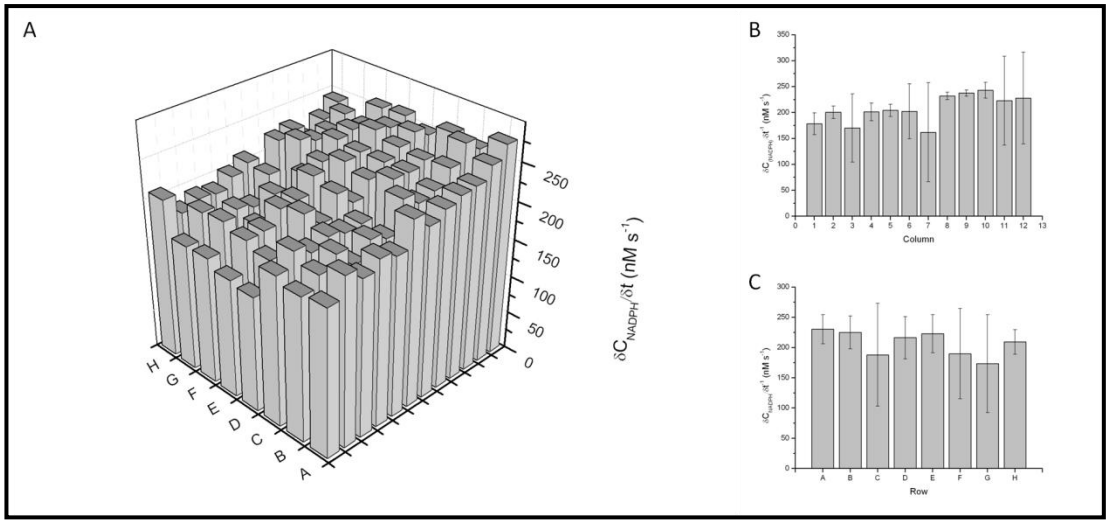


Figure 3.32. Panel A shows the initial rates for screening reactions in individual wells of the 96 well plate. Panel B shows the initial rate for screening reactions grouped by row. Panel C shows the initial rate for screening reactions grouped by column.

An important conclusion from this experiment is to demonstrate the minimisation of a source of systematic error inherent in the purification protocol. Early purification trials with magnetic beads demonstrated a large bias, resulting in wells pipetted later received a larger amount of magnetic beads attributable to them settling during distribution. Bead settling was counteracted by the addition of sucrose to the magnetic bead buffer However, as is demonstrated in Figure 3.32, panels B and C, whilst this effect has been minimised it has not been completely eliminated with a maximal variation of up to 22 % in the average across the columns of the plate; no trend in variation was observed when classified by row. Ideally, this type of error would be best classified by assessment of the protein concentration but insufficient data were available from protein qubit concentration determination to analyse in this way.

Potential sources of error associated with the growth, purification and assay in 96 well format include; i) the accuracy of robotic colony picking, ensuring that the best growing colonies are picked and that no bubbles or marks in the agar are mistaken for colonies; ii) the uniformity of growth rate between individual wells, variation in the rate of growth of individual cultures leads to variety in the volume of the pellet harvested, this includes the potential loss of pellets during removal of growth media; iii) the uniformity of purification conditions, this includes variation in the

amount of magnetic beads, the volume of individual buffers and the volume of final eluate associated with errors in pipetting on the robotics platform; iv) the uniformity of assay conditions, largely consisting of error in pipetting, errors in timing of reaction initiation and reading, errors associated with spectrophotometer reading. Given the number of potential sources of error in these experiments the level of constancy in the wild type assay is remarkable.

3.6 Summary

In summary, $PETNR_{WT}$ was purified and characterised to provide baseline values with which to compare data of the tagged enzymes. A poly-histidine tagged version ($PETNR_{His}$) was generated which expressed extremely well with simple IPTG induction, purified simply in a one step protocol and was kinetically almost indistinguishable from $PETNR_{WT}$. A biotin tagged variant of PETNR was generated ($PETNR_{Bio}$), which required complex induction conditions and was demonstrated to be expressed to a low level and was characterised no further because of this difficulty.

Following generation and characterisation of $PETNR_{His}$ of the next step was to develop a consistent 96 well purification and assay for the high throughput growth purification and screening of mutant libraries. This chapter has outlined preliminary experiments which enabled the finalisation of a protocol for the library colony picking, growth, lysis, purification, and assay using a Hamilton robotics platform. The rationale and optimisation of the methods of colony picking, AIM growth, lysis, purification, assay and *E. coli* strain selection have been demonstrated. Initial full scale trials of the robotics procedures with $PETNR_{His}$ enabled an estimation of the accuracy and reproducibility of the procedures and a discussion of potential sources of error.

4 Generation and initial characterisation of PETNR$_{His}$ libraries

Following the development of a repeatable, reliable library generation, growth, purification and assay protocol as discussed in Chapter 3. The first stage of the experimental outline (Figure 3.1) involved the generation of diversity in the gene encoding the protein of interest. Following library generation, it is important to ensure that; libraries have been sufficiently randomised, that libraries encode proteins which are expressed, that the purification protocol performs correctly and to characterise the baseline activity with a model substrate couple. Following satisfactory verification of these libraries they can then be included in reactions screening for acceptance of a range of novel substrates which will be described in Chapter 5.

The design of a series of mutant libraries must include consideration of a number of issues including, the method of library generation and the residue position targeted for mutagenesis.

First, one needs to consider a number of options available regarding mutation strategy to employ. This ranges from error prone PCR methods (Section 1.6.2.2, [85, 87, 90]), covering large amounts of sequence space, to small targeted single site saturation methods (Section 1.6.2.3, [92, 96, 98]), which can be screened with a high degree of coverage.

In the case of OYE homologues, in particular PETNR, a large amount of information is available concerning the mechanism in the active site and the residues which contribute to catalysis (Section 1.5, [38, 40, 73, 75]). In the studies reported in this thesis, single site saturation mutagenesis has been used which allows generation of small focussed libraries [92]. This approach, however, relies heavily on knowledge of structure-function relationships.

In single site saturation libraries single residue positions are randomly mutated by utilisation of degenerate oligonucleotides to perform PCR-based replication of plasmid DNA incorporating the desired mutated codons [92, 96]. One critical factor in the generation of these libraries is the types of nucleotides included in the

degenerate oligonucleotides [94, 95]. A range of options can be considered to restrict the number of degenerate combinations at a defined position. For example, one can incorporate only a single or subset of nucleotides at a defined position in the DNA sequence, this limits the number of combinations represented in the screen therefore limiting the required screening effort [95]. This leads to maximum representation of amino acids with minimal DNA variation [94]. If the variation in DNA sequence is constrained too much the library will not contain all amino acid residues types at the target position, and under some circumstances this is desirable [94]. There is clearly a trade off between variation in sequence and amino acid coverage in the library. The degeneracy option selected for this experiment was use the bases NNK for target residue codons in the mutagenesis oligonucleotides [95]. This enables coverage of all 20 amino acid residues at the defined location whilst limiting the number of individual clones to be screened [94]. Consequently, for 95 % coverage of the library one needs to screen 94 colonies (calculated from equation given in Section 1.5.2.4 [95, 102]), which can be achieved in a single deep well block.

This approach should allow for rapid saturation mutagenesis of a number of positions in the active site of PETNR whilst maintaining reasonable library sizes, thereby allowing rapid screening of a number of target compounds.

A further important factor for the generation and screening of single site saturation libraries is the residue position selected for mutagenesis. As described previously (Section 1.4), a wide range of structure-function data is available for PETNR and other OYE homologues and these structural data were used to inform the selection of target residues for mutagenesis experiments. These residues were selected on the basis of their involvement in catalysis (as defined by literature reports) [38, 73, 75], by analogy with residues involved in catalysis in other OYE homologue (references within Section 1.4 and [169]), or by their proximity to the active site in the PETNR crystal structure [40]. Residues such as H181 and H184 (Section 1.5.3.4, [38, 73]), Y351 (Section 1.5.3.6, [15, 40]) are known to have a clear role in substrate binding. Residue Y186 (Section 1.5.3.5) was targeted following identification as the proton donor in OYE [77], however having a more obscure role in the remainder of

the family [41, 73]. Residues T26 (Section 1.5.3.1, [66, 73, 78]), Y68 (Section 1.5.3.2, [40, 78]), W102 (Section 1.5.3.3, [38, 75, 78]) and Q241 (Section 1.5.3.7, [15, 40, 41]) were targeted following implication in influencing reactivity from a combination of structural data and past studies.

4.1 Single site saturation mutagenesis library generation

Single site saturation libraries were generated by PCR amplification of $PETNR_{His}$ using single site degenerate oligonucleotides at a range of positions using Kod DNA polymerase according to the standard PCR protocol (Section 2.2.1.8). Oligonucleotide sequences are given in Appendix 1 (Page 236). Following *DpnI* digestion (Section 2.2.1.9), PCR products were analysed by agarose gel electrophoresis (Section 2.2.1.1) and gel purified (Section 2.2.1.3).

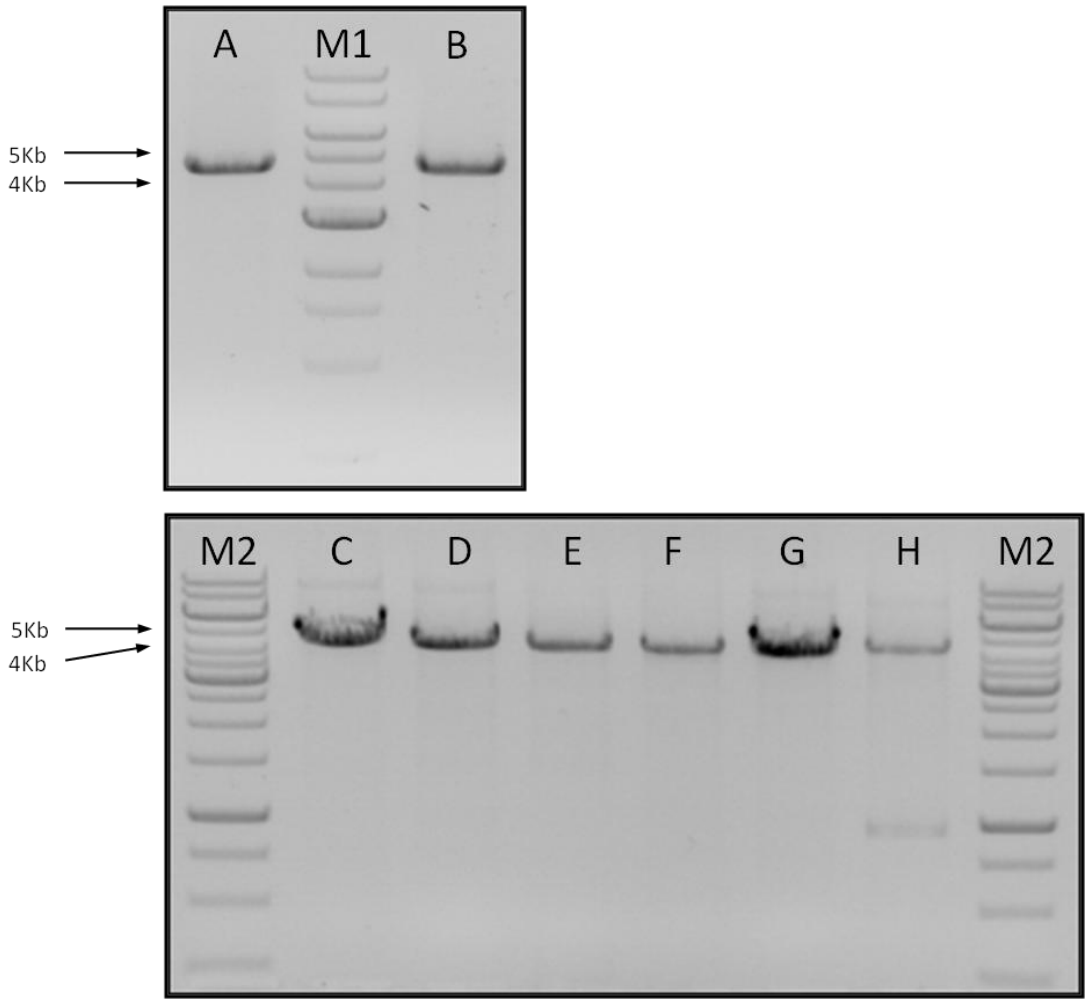


Figure 4.1. Agarose gels showing the results from PCRs to generate single site NNK libraries. Lane M1, 1 kbp ladder (NEB; 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10 kbp); Lanes M2, Generuler 1 kbp ladder (Fermentas; 250, 500, 750 bp, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 8, 10 kbp); Lane A, H181X; Lane B, H184X; Lane C, T26X; Lane D, Y68X; Lane E, W102X; Lane F, Y186X; Lane G, Q241X; Lane H, Y351X. The band marked on the gel is consistent with the 4 kb plasmid (pBluescript SK+) containing the PETNR insert. Electrophoresis was performed in a 0.6 % agarose gel at 100 V for 45 min.

All PCRs generated the expected ~4 kbp plasmid product in sufficient quantities (Figure 4.1). These products were purified from the agarose gel and transformed into *E. coli* strain JM109 (Section 2.2.1.8) for library generation. The transformations were plated onto selective LB agar plates following serial dilution. Libraries of mutants were generated by selecting individual colonies using the robotic colony picker, transferred to a 96 well plate containing TB medium, grown (overnight at 37 °C, according to the standard protocol, Section 2.2.4.10.2) and stored at -80 ˚C in 96 well PCR plates plate following addition of glycerol (Section 2.2.4.10.1). These glycerol stocks were used as the inocula for all library screening experiments (Section 2.2.4.10.2-3).

4.2 Random sequencing of colonies

In order to determine the degree of randomisation of the amino acid residue at each targeted position, a number of individual clones were selected for growth amplification, these cells were used to prepare DNA for sequencing. Consequently, a sample of 10 clones from each library (selected randomly) were grown and used for DNA sequence analysis (Section 2.2.1.2; Section 2.2.1.5); replicate cultures were also grown in 96 well format for quantification of protein expression, enzyme purification and assay. The randomly selected wells were replicated for all libraries and the same experiments were performed.

In the results described below, where large insertions/deletions were observed, these insertions/deletions are omitted from the alignment figures (Figure 4.2-9) for the sake of clarity. However, all results are detailed in the tabulated results (Tables 4.1-8). Labels adopt a standard format maintained throughout where the first block of numerals indicates the mutant library and the second alpha-numeric entry indicates the well sampled from that library (e.g. 26 A3 indicates well A3 from the T26X library).

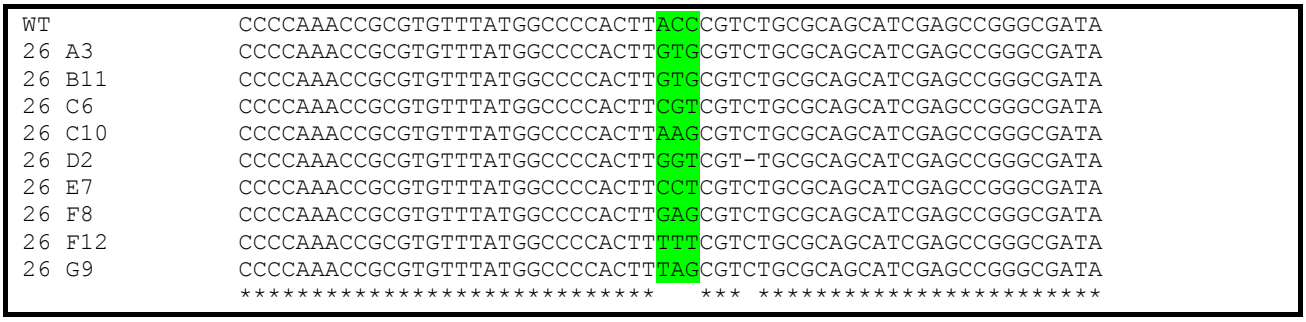


Figure 4.2. Alignment of 10 random selected colonies from the PETNR T26X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	ACC	Thr
A3	GTG	Val
B11	GTG	Val
C6	CGT	Arg
C10	AAG	Lys
D2	GGT	Gly
E7	CCT	Pro
F8	GAG	Glu
F12	TTT	Phe
G9	TAG	**Stop**
H4	*Insertion*	

Table 4.1. Identification of residue 26 in randomly sequenced clones from the T26X library.

```
WT              AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCTACGCCGGTGCACCGGGTCTGCACAGCC
68 A3           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCGTTGCCGGTGCACCGGGTCTGCACAGCC
68 B11          AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCAATGCCGGTGCACCGGGTCTGCACAGCC
68 C6           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCGGGGCCGGTGCACCGGGTCTGCACAGCC
68 C10          AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCCTGGCCGGTGCACCGGGTCTGCACAGCC
68 D2           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCAAGGCCGGTGCACCGGGTCTGCACAGCC
68 E7           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCAGGGCCGGTGCACCGGGTCTGCACAGCC
68 F8           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCCAGGCCGGTGCACCGGGTCTGCACAGCC
68 F12          AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCGTGGCCGGTGCACCGGGTCTGCACAGCC
68 G9           AAGCCACGCAGATTTCTGCTCAGGCAAAAGGCCATGCCGGTGCACCGGGTCTGCACAGCC
                ********************************  ************************
```

Figure 4.3. Alignment of 10 random selected colonies from the PETNR T68X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	TAC	Tyr
A3	GTT	Val
B11	AAT	Asn
C6	GGG	Gly
C10	CTG	Leu
D2	AAG	Lys
E7	AGG	Arg
F8	CAG	Gln
F12	GTG	Val
G9	CAT	His
H4	*Deletion*	

Table 4.2. Identification of residue 68 in randomly sequenced clones from the Y68X library.

```
WT              TTCAGCTGTGGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 A3          TTCAGCTGTAGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 B11         TTCAGCTGCTGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 C6          TTCAGCTGTCTCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 C10         TTCAGCTGTCTCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 D2          TTCAGCTGTTTCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 E7          TTCAGCTGAGGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 F8          TTCAGCTGGGGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 F12         TTCAGCTGCCGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 G9          TTCAGCTGATTCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
102 H4          TTCAGCTGGGGCACACCGGTCGTATCTCACACAGCAGCATCCAGCCTGGCGGTCAGGCGC
                ********  **********************************************
```

Figure 4.4. Alignment of 10 random selected colonies from the PETNR W102X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	TGG	Trp
A3	TAG	**Stop**
B11	CTG	Leu
C6	TCT	Ser
C10	TCT	Ser
D2	TTT	Phe
E7	AGG	Arg
F8	GGG	Gly
F12	CCG	Pro
G9	ATT	Ile
H4	GGG	Gly

Table 4.3. Identification of residue 102 in randomly sequenced clones from the W102X library.

Figure 4.5. Alignment of 10 random selected colonies from the PETNR H181X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	CAC	His
A3	AGG	Arg
B6	GCT	Ala
C10	*Insertion*	
D1	GAT	Asp
E7	TCT	Ser
F2	CGG	Arg
F10	AGT	Ser
G11	CTT	Leu
H2	TCG	Ser
H5	ACG	Thr

Table 4.4. Identification of residue 181 in randomly sequenced clones from the H181X library.

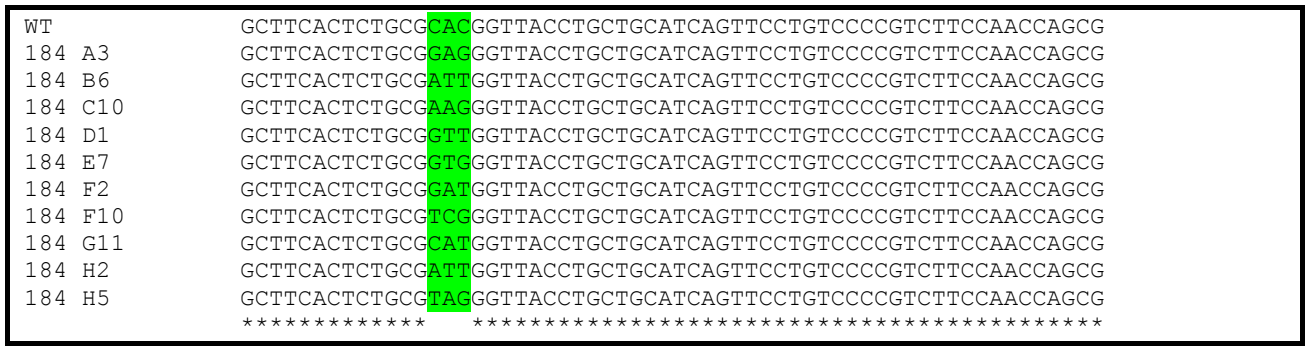


Figure 4.6. Alignment of 10 random selected colonies from the PETNR H184X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	CAC	His
A3	GAG	Glu
B6	ATT	Ile
C10	AAG	Lys
D1	GTT	Val
E7	GTG	Val
F2	GAT	Asp
F10	TCG	Ser
G11	CAT	His
H2	ATT	Ile
H5	TAG	**STOP**

Table 4.5. Identification of residue 184 in randomly sequenced clones from the H184X library.

```
WT             GGTACCCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 A3         GGTCGTCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 B11        GGTGGTCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 C6         GGTCGGCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 F12        GGTAATCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 G9         GGTGATCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
186 H4         GGTAGGCTGCTGCATCAGTTCCTGTCCCCGTCTTCCAACCAGCGTACCGACCAGTACGGC
               ***   **************************************************
```

Figure 4.7. Alignment of 10 random selected colonies from the PETNR Y186X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	TAC	Tyr
A3	CGT	Arg
B11	GGT	Gly
C6	CGG	Arg
C10	*Insertion*	
D2	*Insertion*	
E7	*Insertion*	
F8	*Insertion*	
F12	AAT	Asn
G9	GAT	Asp
H4	AGG	Arg

Table 4.6. Identification of residue 186 in randomly sequenced clones from the Y186X library.

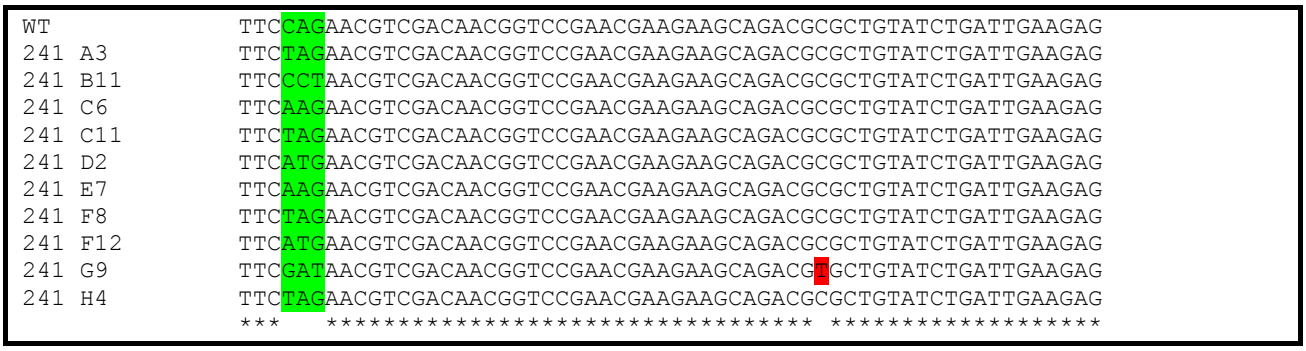


Figure 4.8. Alignment of 10 random selected colonies from the PETNR Q241X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	CAG	Gln
A7	TAG	**Stop**
B11	CCT	Pro
C6	AAG	Lys
C10	TAG	**Stop**
D2	ATG	Met
E7	AAG	Lys
F8	TAG	**Stop**
F12	ATG	Met
G9	GAT	Asp
H4	TAG	**Stop**

Table 4.7. Identification of residue 241 in randomly sequenced clones from the Q241X library.

Figure 4.9. Alignment of 10 random selected colonies from the PETNR Y351X library. Plasmid DNA was extracted from 5 ml cultures and sequenced in the region of the randomised position. The alignment was performed using the Clustal online tool [164].

Clone	Codon	Amino Acid
WT	TAT	Tyr
A7	GCT	Ala
B11	TCG	Ser
C6	AAT	Asn
C10	CTG	Leu
D2	TAT	Tyr
E7	TGG	Trp
F8	ACT	Thr
F12	CTT	Leu
G9	TTT	Phe
H4	TTG	Leu

Table 4.8. Identification of residue 351 in randomly sequenced clones from the Y351X library.

Library	Unique mutants	Duplicate amino acids[a]	Ins/Del	Wild-type	Stop codons	Usable clones[b] (%)	Calculated actual screening effort (clones)
T26X	6	1	1	0	1	80	76
Y68X	8	1	1	0	0	90	86
W102X	7	2	0	0	1	90	86
H181X	6	2	1	0	0	80	76
H184X	6	2	0	1	1	100	96
Y186X	4	2	4	0	0	60	57
Q241X	3	2	0	0	4	90	86[c]
Y351X	8	1	0	1	0	100	96

[a] Number of mutant amino acids present at least twice within the 10 library members sequenced.
[b] Percentage of clones that did not contain an insertion/deletion/additional mutation. Wild-type/STOP codons were included in the calculation.
[c] Likely to be an overestimate due to over-representation of STOP codons

Table 4.9. Statistics for the random sequencing of PETNR libraries.

Statistics summarising the library sampling are given in Table 4.9. It is important to point out that 10 out of 96 colonies were sampled representing ~10 % of wells. Statistical analysis would be improved if a wider sample number was chosen. However, for the purpose of confirming randomisation at the target codon the sample size was considered to be sufficient.

There are several important points to note when analysing the random sequencing results from the constructed libraries. First, there is an occasional occurrence of internal STOP codons at the randomised position, as seen in the libraries T26X, W102X and H184X. This is expected as the degeneracy used for the construction of the libraries (NNK) includes the termination codon TAG. In some cases, such as the Q241X library, the frequency of an internal stop codon seems to be greatly enhanced compared with that expected by probability alone. This might indicate some bias towards clones containing the STOP codons at this position (although one needs to consider the low sample number used in the analysis). An interesting point to note is the similarity between the wild type CAG codon of this library and the STOP TAG codon. This suggests that the STOP codon might bind preferentially

to the template DNA compared with other oligonucleotides. However, the same over-representation should also be present for wild type, but this does not appear to be the case. From the available evidence it is impossible to determine if this is mechanism based or a statistical artefact.

Other unexpected mutations were present in the libraries, such as, point mutations outside the targeted region and larger insertion and deletions of DNA sequence. This phenomenon is potentially problematic since clones with additional mutations could be misfolded, missing parts of the protein essential for activity, may be targeted for proteolysis *in vivo*, or have compromised activity. In some cases, large numbers of the same unintended mutant in a library can be traced back to a single mutation (e.g. insertions in Y186X). This change, thought to occur early in the PCR reaction, is then propagated further in subsequent PCR amplifications so that it is present in a large subset of the library. In the majority of cases, however, mutations outside the target region occur only once and are located close to the randomised codon. This often occurs in PCR-based mutagenesis protocols and might be explained by errors in the solid phase synthesis of oligonucleotide sequences. An alternate explanation where errors are further from the randomised codon is that polymerase enzymes used for PCR have an intrinsic error rate and as such naturally make occasional errors.

4.3 Expression trials of mutant libraries

Each site-saturated library was grown in TB AIM from the glycerol stocks (Section 2.2.4.10.2), harvested (2.2.4.10.3) and frozen at -20 ˚C. Samples of cell pellets were analysed by SDS PAGE to determine the levels and consistency of PETNR expression in each of the mutant libraries.

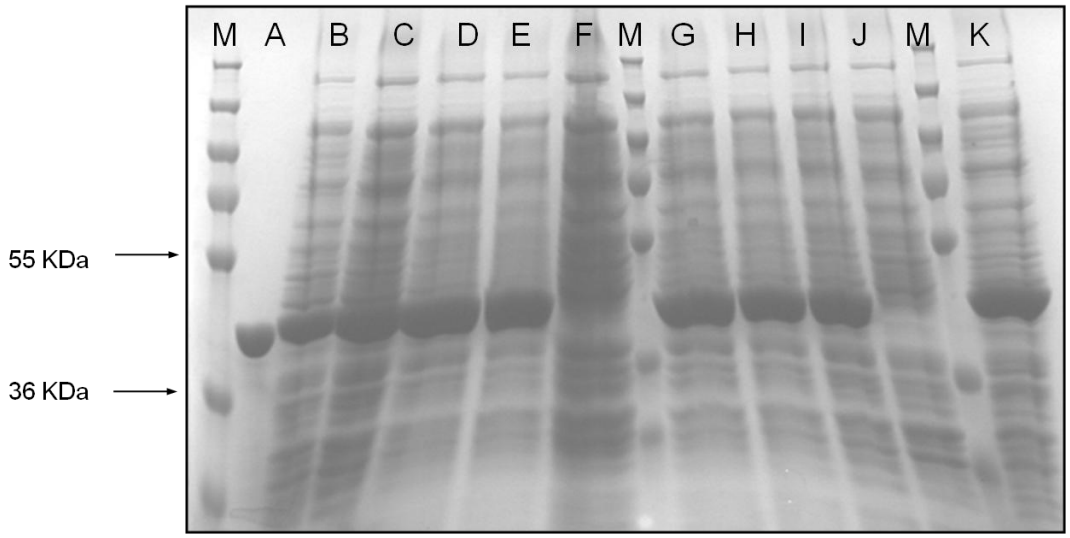


Figure 4.10. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the T26X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR standard; Lane B, 26 A3 – T26V-1; Lane C, 26 B11 – T26V-2; Lane D, 26 C6 – T26R; Lane E, 26 C10 – T26K; Lane F, 26 D2 – T26G; Lane G, 26 E7 – T26P; Lane H, 26 F8 – T26E; Lane I, T26 F12 – T26F; Lane J, 26 G9 – **STOP**; Lane K, *Insertion*. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

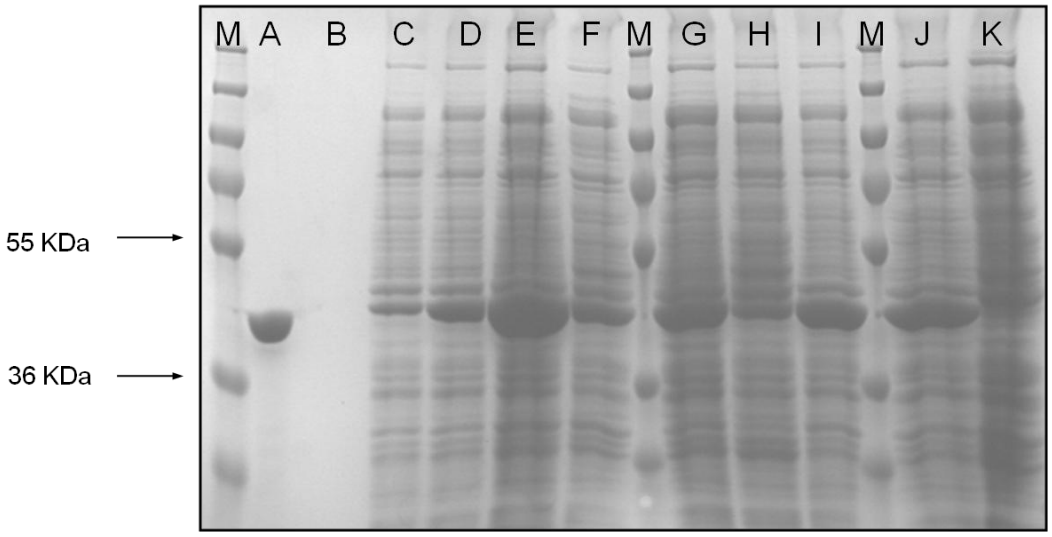


Figure 4.11. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the Y68X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A,PETNR standard; Lane B, 68 A3 – Y68V; Lane C, 68 B11 – Y68N; Lane D, 68 C6 – Y68G; Lane E, 68 C10 – Y68L; Lane F, 68 D2 – Y68K; Lane G, 68 E7 Y68R; Lane H, 68 F8 - Y68Q; Lane I, 68 F12 – Y68V; Lane J, 68 G9 – Y68H; Lane K, 68 H4 – *Deletion*. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

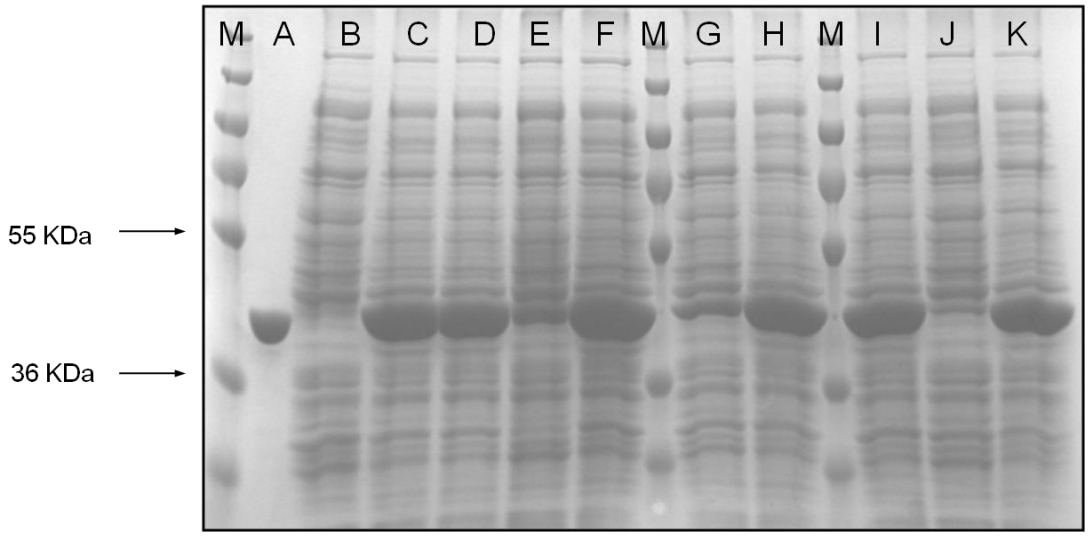


Figure 4.12. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the W102X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A,PETNR standard; Lane B, 102 A3 – **STOP**; Lane C, 102 B11 – W102L; Lane D, 102 C6 – W102S; Lane E, 102 C10 – W102S; Lane F, 102 D2 – W102F; Lane G, 102 E7 W102R; Lane H, 102 F8 – W102G; Lane I, 102 F12 – W102P; Lane J, 102 G9 – W102I; Lane K, 102 H4 – W102G. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

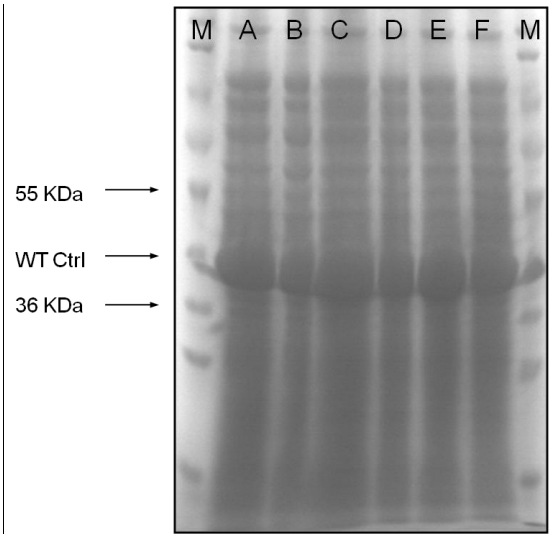


Figure 4.13. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the H181X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa) including PETNR standard (40 kDa); Lane A, 181 A11; Lane B,181 B2; Lane C, 181 C6; Lane D, 181 E9; Lane E, 181 G2; Lane F, 181 H7. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

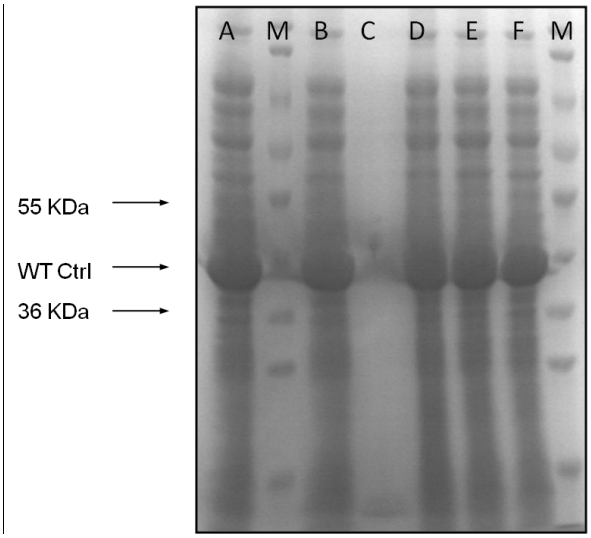


Figure 4.14. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the H184X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36 , 55, 72, 95, 130, 250 kDa) including PETNR standard (40 kDa); Lane A, 184 A11; Lane B, 184 B2; Lane C, 184 C6; Lane D, 184 E9; Lane E, 184 G2; Lane F, 184 H7. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

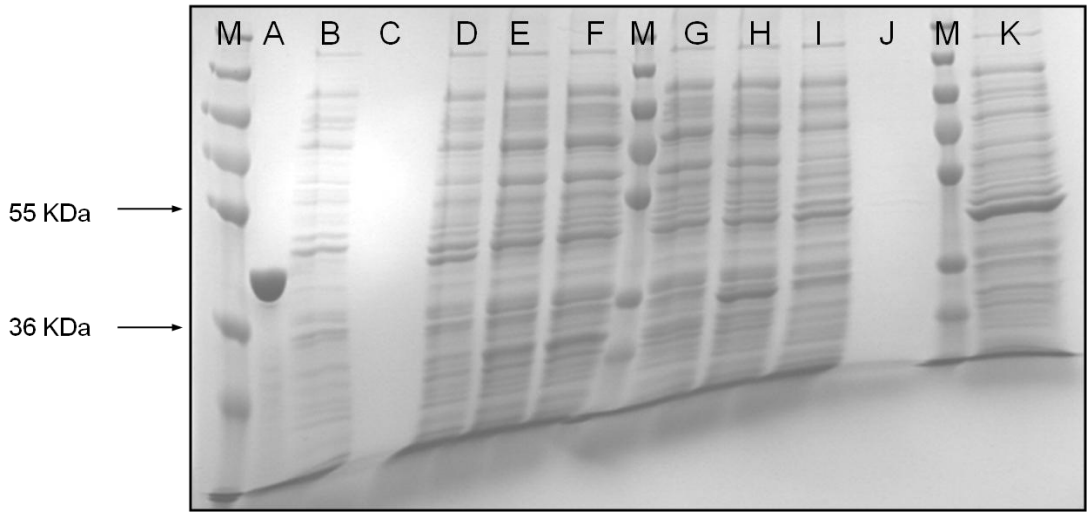


Figure 4.15. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the Y186X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR standard; Lane B, 186 A3 – Y186R-1; Lane C, 186 B11 – Y186G; Lane D, 186 C6 – Y186R-2; Lane E, 186 C10 – Y186M; Lane F, 186 D2 – Y186M; Lane G, 186 E7 – Y186I; Lane H, 186 F8 – *Deletion*; Lane I, 186 F12 – Y186N; Lane J, 186 G9 – Y186D; Lane K, 186 H4 – Y186R-3.Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

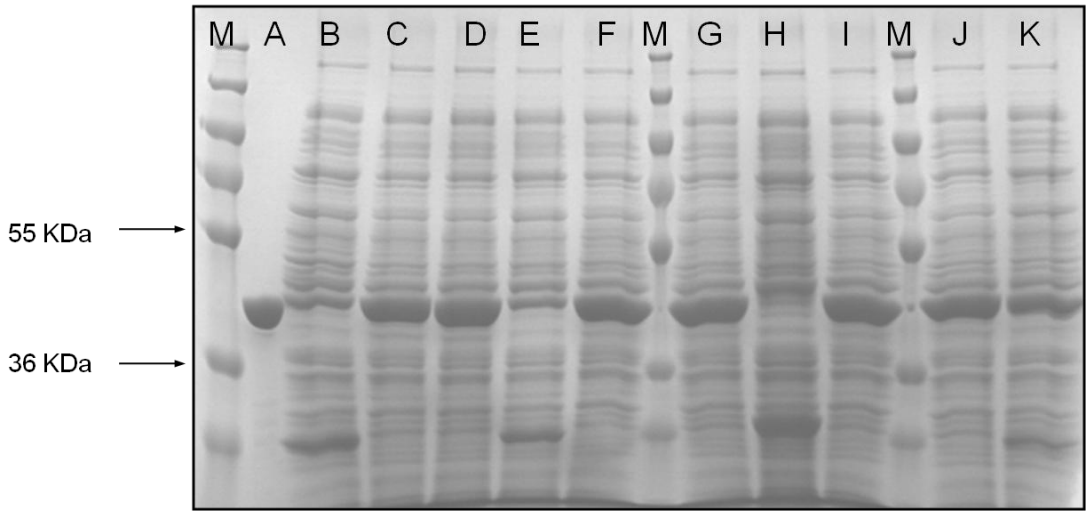


Figure 4.16. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the Q241X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR standard; Lane B, 241 A3 – **STOP**; Lane C, 241 B11 – Q241P; Lane D, 241 C6 – Q241K-1; Lane E, 241 C10 – **STOP**; Lane F, 241 D2 – Q241M-1; Lane G, 241 E7 – Q241K-2; Lane H, 241 F8 – **STOP**; Lane I, 241 F12 – Q241M-2; Lane J, 241 G9 – Q241D; Lane K, 241 H4 – **STOP**. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

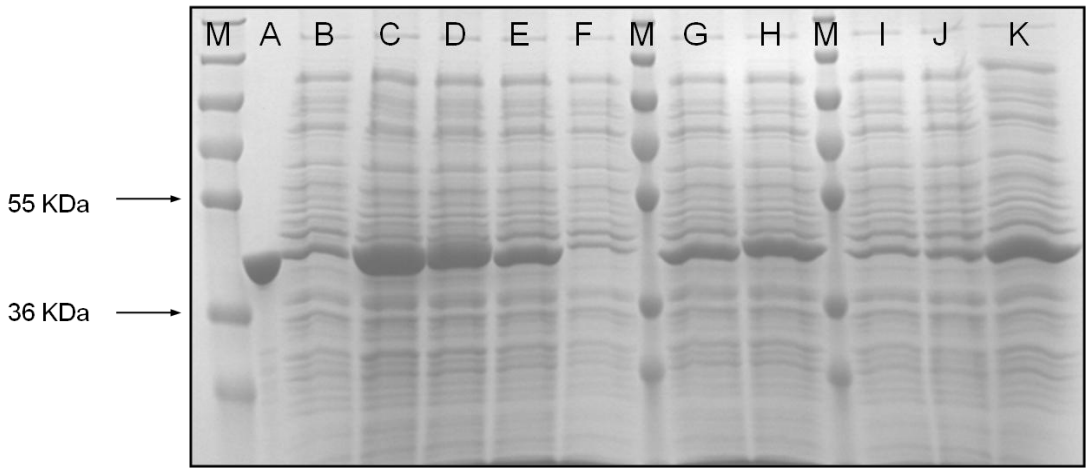


Figure 4.17. SDS PAGE analysis of expression of PETNR mutants in whole cell pellets of the Y351X library. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17 (not visible), 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR Standard; Lane B, 351 A3 – Y351A; Lane C, 351 B11 – Y351S; Lane D, 351 C6 – Y351N; Lane E, 351 C10 – Y351L-1; Lane F, 351 D2 – WT; Lane G, 351 E7 – Y351W;Lane H, 351 F8 – Y351T; Lane I, 351 F12 – Y351L-2; Lane J, 351 G9 – Y351F; Lane K, 351 H4 – Y351L-3. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

The majority of clones from the libraries are capable of expressing PETNR, although not all at the same level. However, the Y186X library failed to over express PETNR mutant protein (Figure 4.15). A possibility to be considered is that mutations at this residue lead to misfolded proteins that are then targeted for proteolysis by the cell. However, this is inconsistent with published reports that have demonstrated expression and purification of folded enzyme following directed mutagenesis at this position (Section 1.5.3.5, [73]). Localised sequence analysis around the site of mutagenesis indicated randomisation of the target codon (Figure 4.15 legend). However, sequence analysis of the whole gene was not performed and the possibility remains that errors in sequence distant from the mutagenesis site were generated in the PCR protocol. Analysis of mutants at this position in PETNR will likely require re-construction of the library. The expression problem in the Y186X library was not characterised any further.

Some libraries (e.g. T26X, Figure 4.10) demonstrated remarkable consistency in the amount of over-expression in the sample clones, suggesting mutations at this position had little effect on protein expression. However, libraries such as Y68X and Y351X (Figure 4.11 and 4.17 respectively) show greater variability in the levels of expression seen.

SDS PAGE analysis of clones containing internal STOP codons showed that in the majority of cases (wells 241 A3, C10, F8; Figure 4.16), the inferred smaller over-expression product is observed. However, the identity of these products was not confirmed. As the truncated products do not contain the C terminal His tag, these clones should not be purified by the robotics protocol to be used in library screening. Well H4 also coding for STOP has a substantial amount of product of the expected size for full length PETNR as well as the truncated product. The most likely explanation for this is that multiple colony forming units (cfu) were closely positioned on the plate. The robotic colony picker would see this as a single colony. In order to limit this occurrence in subsequent screening, the colony picking programme was modified so that only circular colonies were picked, thereby minimising recovery of ‘mixed’ colonies from the plate.

Given the large degree of variation in protein expression, lysed cell assays performed with these libraries are not comparable because the enzyme concentration in each sample is unknown. To overcome this problem, a micro-scale purification of enzyme from the library clones was designed. The aim was to recover similar quantities of protein from each culture sample amplified from a picked colony. Since recovery of purified protein is similar in each well of the 96 well plate, normalisation is achieved by using the same protein concentration in each assay well. Consequently, only very poorly expressing clones will be problematic in achieving normalisation in the automated activity assays.

4.4 Robotic purification of mutant libraries

As well as normalising enzyme concentrations, robotics purification protocols were designed to eliminate competing activity from other cellular NADPH utilising enzymes. Another complicating factor is that PETNR also has an intrinsic NADPH oxidase activity in the presence of oxygen. Consequently, experiments were performed in an anaerobic environment (Section 2.2.3.1). Following harvesting of cells from the bacterial cultures (Section 2.2.4.10.3), enzyme was purified from cultures contained in the 96 well plates using the purification protocol (Section 2.2.4.10.4). Eluted protein samples were analysed by SDS PAGE and quantified using the protein qubit concentration determination method (Section 2.2.1.4).

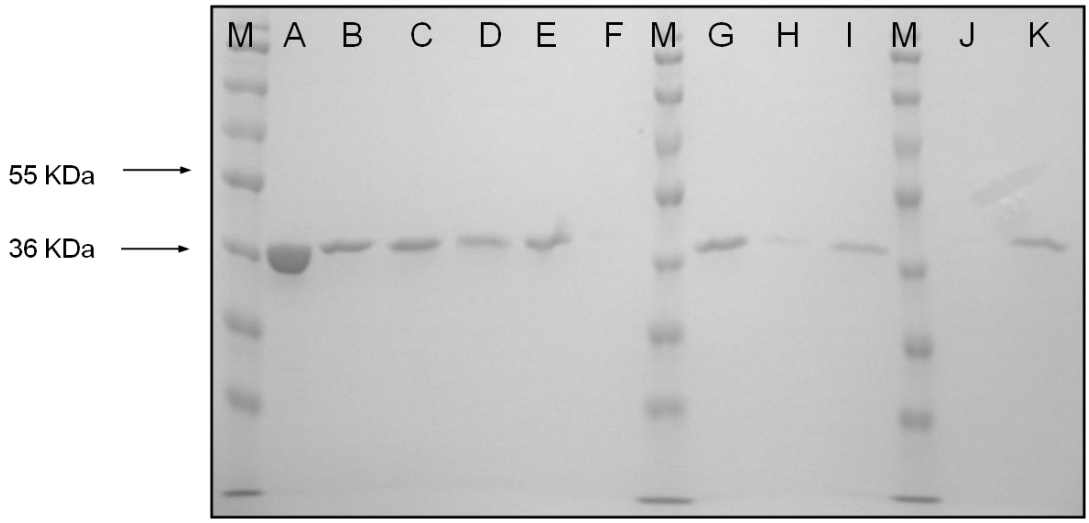


Figure 4.18. SDS PAGE analysis of robotically purified T26X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR standard; Lane B, 26 A3 – T26V-1; Lane C, 26 B11 – T26V-2; Lane D, 26 C6 – T26R; Lane E, 26 C10 – T26K; Lane F, 26 D2 – T26G; Lane G, 26 E7 – T26P; Lane H, 26 F8 – T26E; Lane I, T26 F12 – T26F; Lane J, 26 G9 – **STOP**; Lane K, *Insertion*. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

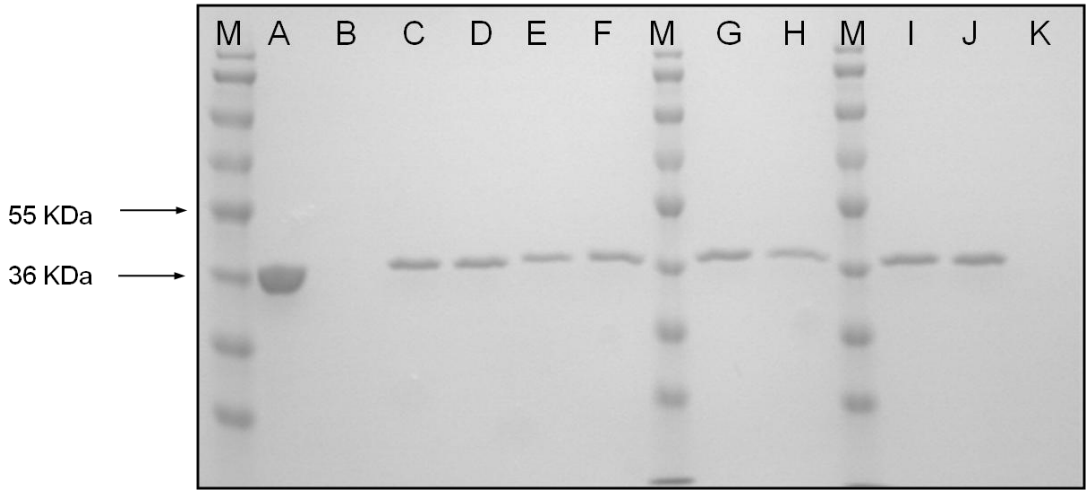


Figure 4.19. SDS PAGE analysis of robotically purified Y68X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa); Lane A,PETNR standard; Lane B, 68 A3 – Y68V; Lane C, 68 B11 – Y68N; Lane D, 68 C6 – Y68G; Lane E, 68 C10 – Y68L; Lane F, 68 D2 – Y68K; Lane G, 68 E7 Y68R; Lane H, 68 F8 - Y68Q; Lane I, 68 F12 – Y68V; Lane J, 68 G9 – Y68H; Lane K, 68 H4 – *Deletion*. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

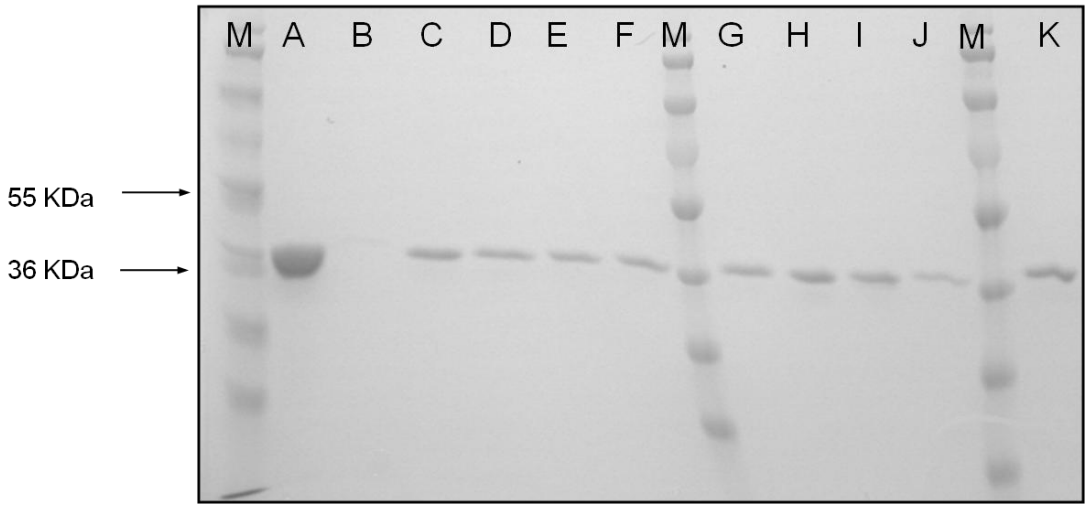


Figure 4.20. SDS PAGE analysis of robotically purified W102X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa); Lane A,PETNR standard; Lane B, 102 A3 – **STOP**; Lane C, 102 B11 – W102L; Lane D, 102 C6 – W102S; Lane E, 102 C10 – W102S; Lane F, 102 D2 – W102F; Lane G, 102 E7 W102R; Lane H, 102 F8 – W102G; Lane I, 102 F12 – W102P; Lane J, 102 G9 – W102I; Lane K, 102 H4 – W102G. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

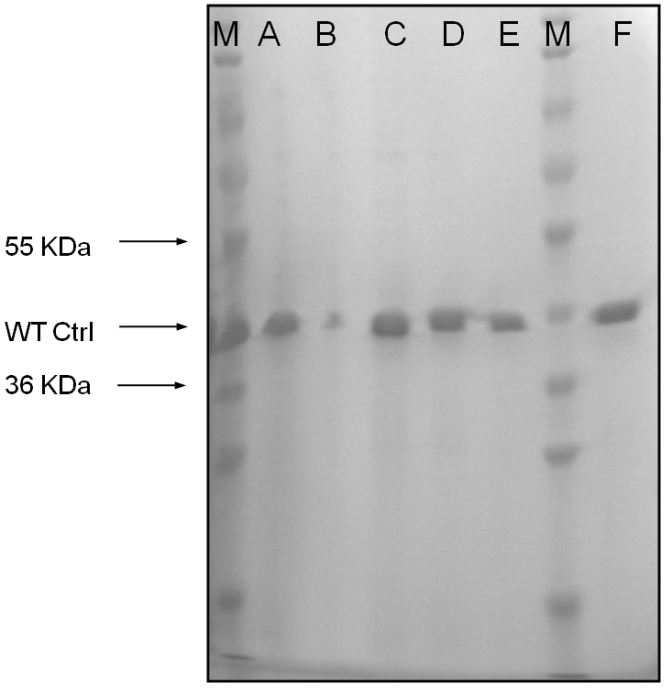


Figure 4.21. SDS PAGE analysis of robotically purified H181X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa) including PETNR standard (40 kDa); Lane A, 181 A11; Lane B,181 B2; Lane C, 181 C6; Lane D, 181 E9; Lane E, 181 G2; Lane F, 181 H7. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

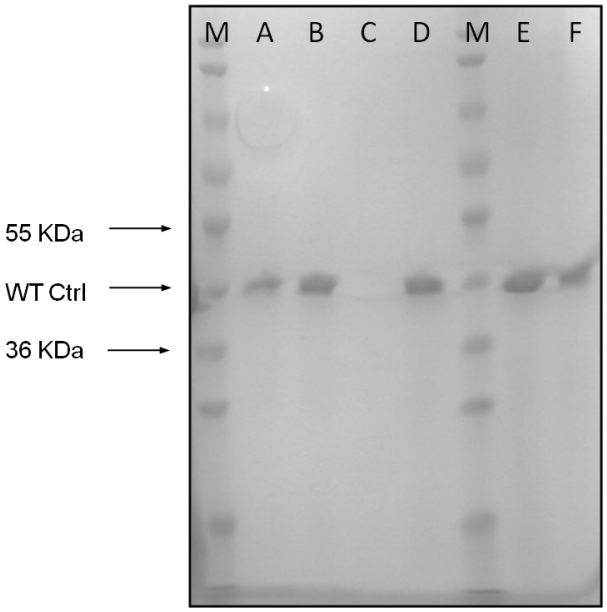


Figure 4.22. SDS PAGE analysis of robotically purified H184X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa) including PETNR standard (40 kDa); Lane A, 184 A11; Lane B, 184 B2; Lane C, 184 C6; Lane D, 184 E9; Lane E, 184 G2; Lane F, 184 H7. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

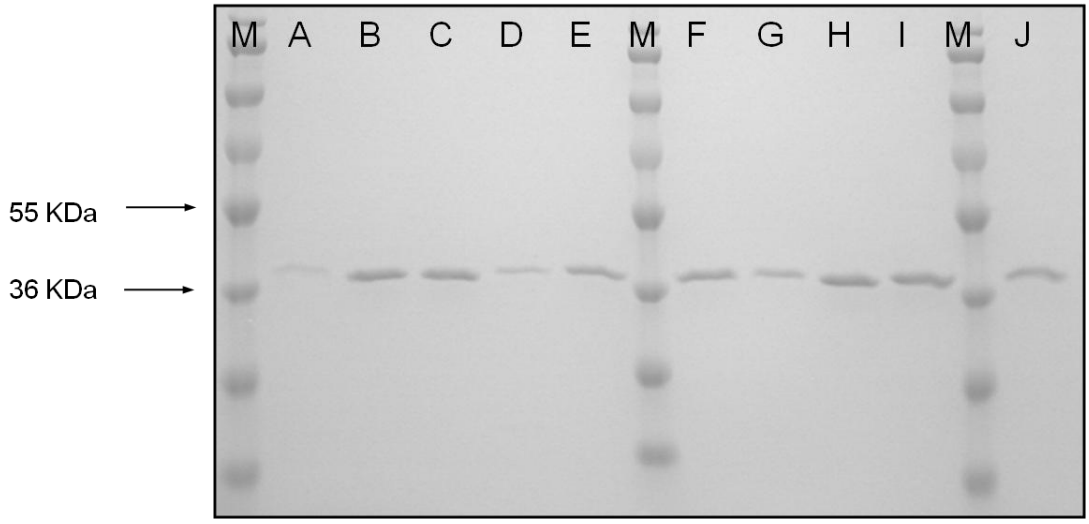


Figure 4.23. SDS PAGE analysis of robotically purified Q241X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, 241 A3 – **STOP**; Lane B, 241 B11 – Q241P; Lane C, 241 C6 – Q241K-1; Lane D, 241 C10 – **STOP**; Lane E, 241 D2 – Q241M-1; Lane F, 241 E7 – Q241K-2; Lane G, 241 F8 – **STOP**; Lane H, 241 F12 – Q241M-2; Lane I, 241 G9 – Q241D; Lane J, 241 H4 – **STOP**. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

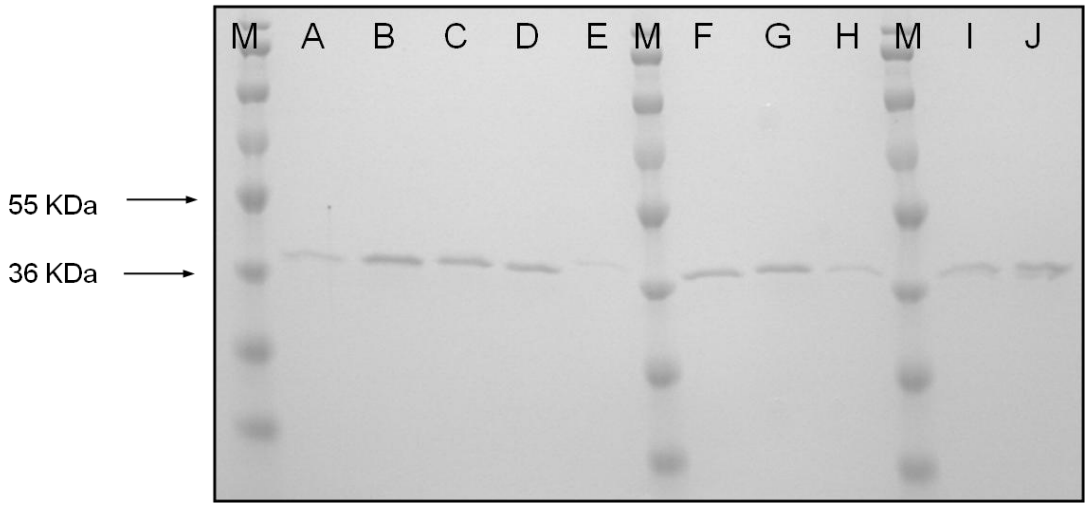


Figure 4.24. SDS PAGE analysis of robotically purified Y351X library clones. Lanes M Pageruler plus prestained protein ladder (Fermentas; 10 (not visible), 17, 28, 36, 55, 72, 95, 130, 250 kDa); Lane A, PETNR Standard; Lane B, 351 A3 – Y351A; Lane C, 351 B11 – Y351S; Lane D, 351 C6 – Y351N; Lane E, 351 C10 – Y351L-1; Lane F, 351 D2 – WT; Lane G, 351 E7 – Y351W;Lane H, 351 F8 – Y351T; Lane I, 351 F12 – Y351L-2; Lane J, 351 G9 – Y351F; Lane K, 351 H4 – Y351L-3. Analysis was performed on a 12 % SDS PAGE gel at 180 V for 45 min, gels were stained with Coomassie Brilliant Blue R250.

Alongside evidence obtained from SDS PAGE, which provides a visual assessment of protein recovery from the purification protocol, protein concentration was estimated for selected clones using the protein qubit quantification method (Table 4.10).

Library	Average Protein Concentration ($\mu g\ ml^{-1}$)	Standard deviation ($\mu g\ ml^{-1}$)
T26X	127	20
Y68X	144	18
W102X	146	15
H181X	*nd*	*nd*
H184X	*nd*	*nd*
Y186X	*n/a*	*n/a*
Q241X	122	13
Y351X	126	42

Table 4.10. Protein concentration determination for PETNR samples taken from wells of the 96 well plate for selected libraries.

Both SDS PAGE analysis and protein quantitation methods showed a high level of consistency in the quantity of purified protein obtained from robotic purification protocols (Figures 4.18 - 24, Table 4.10). This was the case even within libraries that showed a large degree of variation in PETNR expression (e.g Y351X, Figure 4.17). This finding is important for future screening experiments as it ensures that any differences in activity in the library eluates are not influenced substantially by variation in enzyme concentration. Excellent examples of this can be seen by comparing the growth and purification gels for libraries 68 and 102 (particularly 68 B11 vs 68 C10, and 102 C10 vs. 102 D2, See Figure 4.11/4.19, 4.12/4.20). These findings eliminate the need to determine the protein concentration of each eluate following protein purification since each well contains similar quantities of expressed protein. This is a significant advantage since incorporation of protein assays alongside activity assays would severely limit the high throughput nature of the procedure.

SDS PAGE analysis indicated that PETNR eluates were remarkably pure following a single purification step (Figures 4.18 - 24). This reduces considerably the likelihood of detecting activity due to the presence of contaminating *E. coli* NADPH oxidising enzymes. Cross contamination has been observed (e.g. with *E. coli* keto-reductase) in other reported experiments with OYE homologues [61].

There were no examples where an over-expressed protein detected by SDS PAGE in expression trials, resulted in no enzyme recovery following robotic purification. This indicates that the purification steps are reliable and reproducible, again a key requirement for the automated activity assays being developed.

Occasionally, some clones with premature STOP codons (e.g. 241 A3, C10, F8, H4), yielded small quantities of purified protein product (Figure 4.23). This is unexpected as, when expressed, these mutants do not contain a His tag and so should not be purified. A likely explanation is that these clones represent a mixture of two mutants, with one minor species capable of producing full length protein. However, no such erroneous purification was seen in other clones with premature STOP

codons (e.g. T26X G9, Y68X H8 and W102X A3, see Figures 4.18 - 20) suggesting this observation is limited to this specific library.

4.5 Library screening assays using the model substrate couple NADPH/2CH

Trial experiments for robotics-based purification and enzyme assays were first conducted with PETNR libraries using the known substrates NADPH and CH. The purpose of this experiment was to measure the range of activity in the libraries following mutagenesis using a series of substrates (NADPH/CH) that are known to operate well with $PETNR_{WT}$. The second aim of these assays was to identify wild-type PETNR clones from the libraries to use as an internal standard/control (internal wild type control) for activity screening within a single 96 plate. Controls for this experiment were performed by separately assaying known wild-type clones propagated and purified within the 96 well format (external wild type controls) for comparison with activity measured from the library plates. For the external controls, data for activity assays are presented as the average, standard deviation and maximum activity recorded from a 96 well $PETNR_{WT}$ plate (as discussed in Section 3.5.6). This is shown as the first data column in Figures (4.25 - 31). Activity assays from the library plates are single measurements and are shown ranked according to activity, determined spectroscopically as $\delta C/\delta t$ (nM s^{-1}). The 5 results with the highest activity were sequenced to determine which mutation was present.

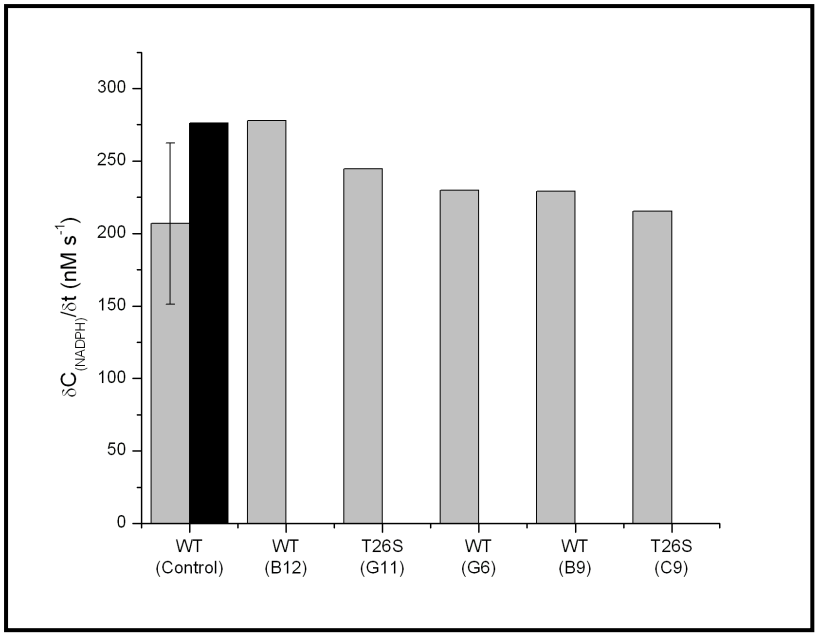


Figure 4.25. Activity of T26X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

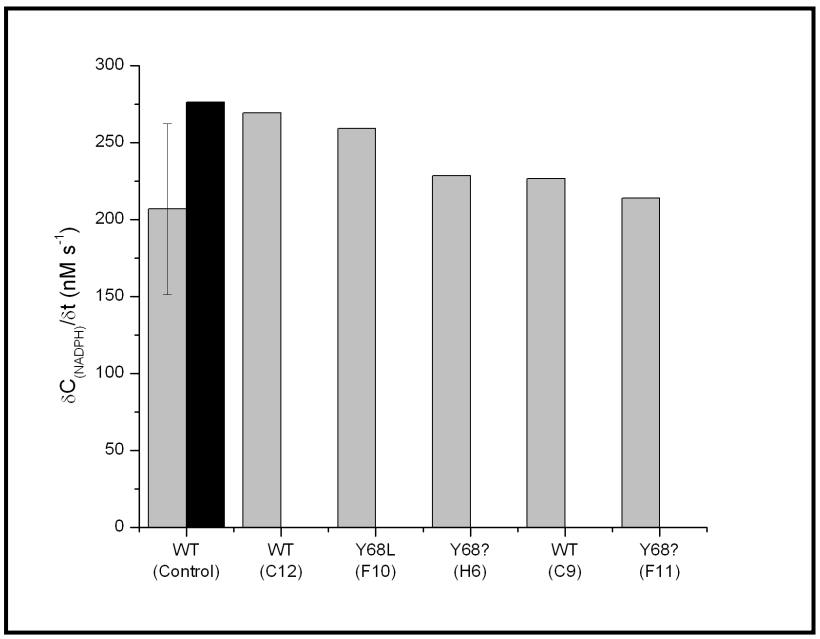


Figure 4.26. Activity of Y68X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

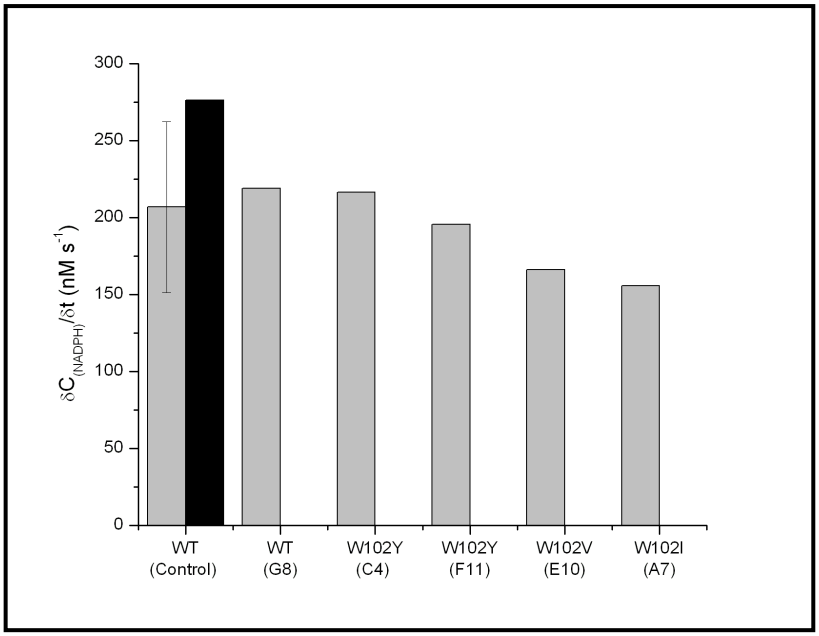


Figure 4.27. Activity of W102X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

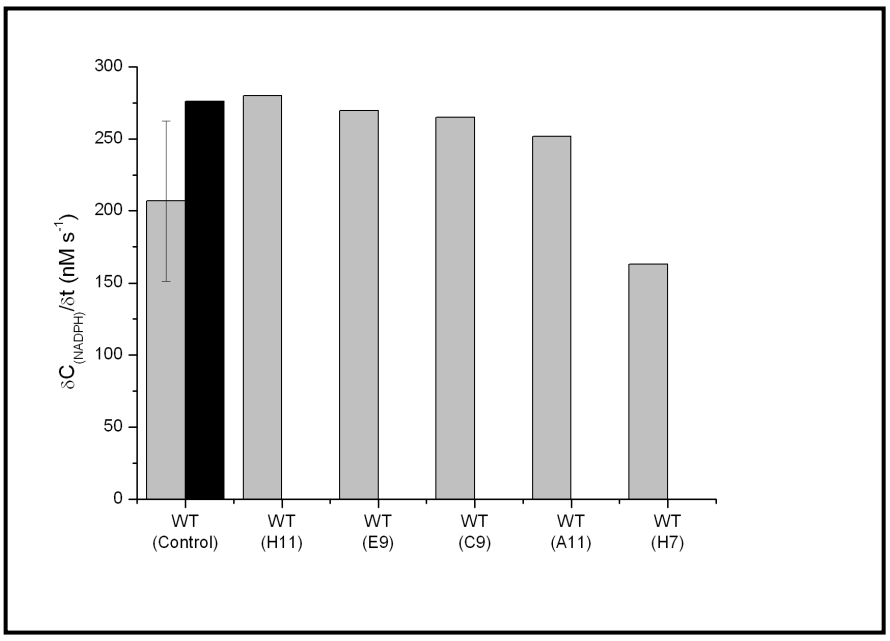


Figure 4.28. Activity of H181X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

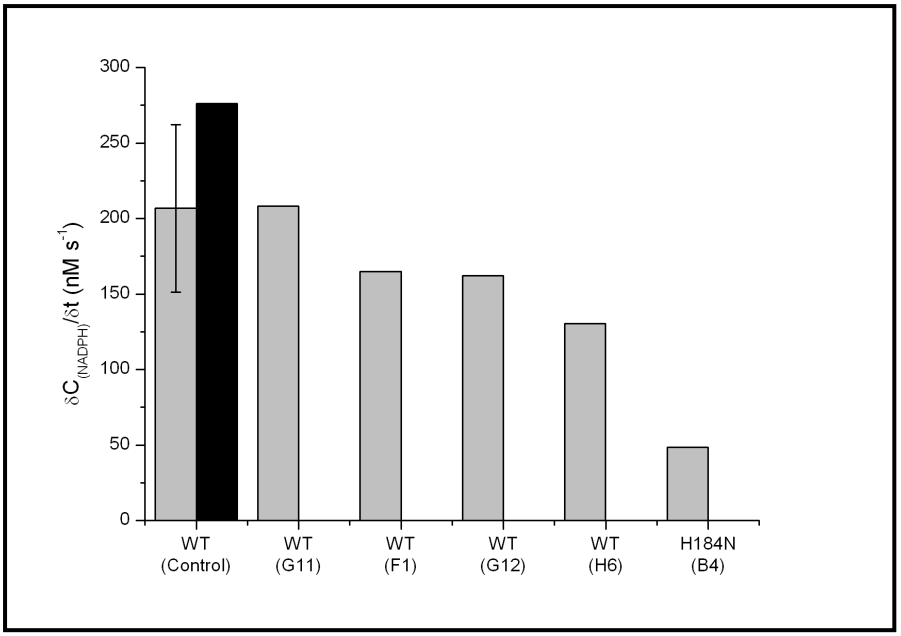


Figure 4.29. Activity of H184X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

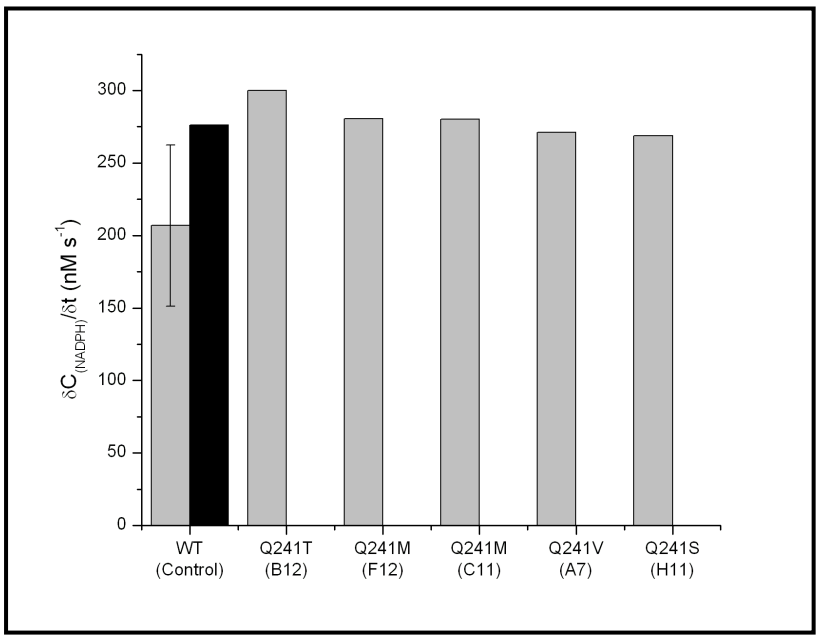


Figure 4.30. Activity of Q241X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

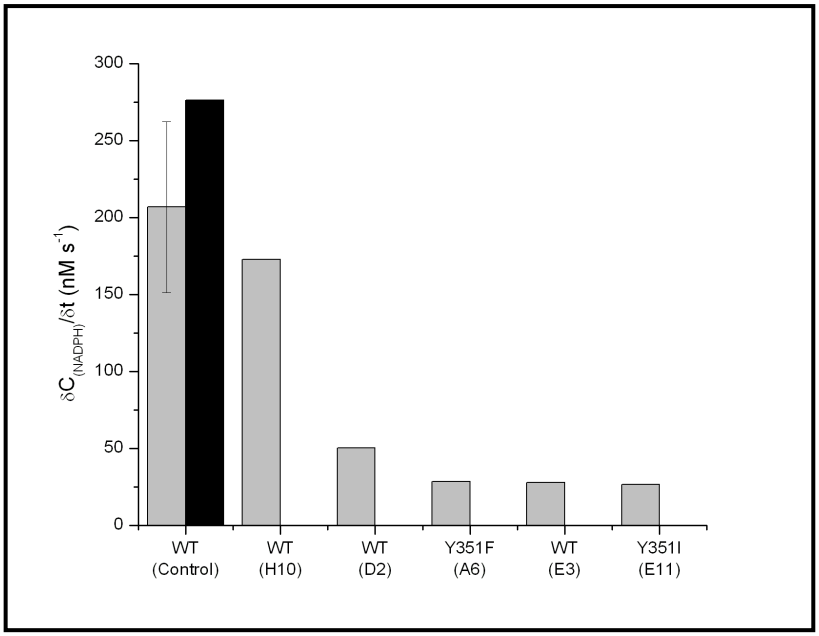


Figure 4.31. Activity of Y351X library eluates. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of a separate $PETNR_{WT}$ block.

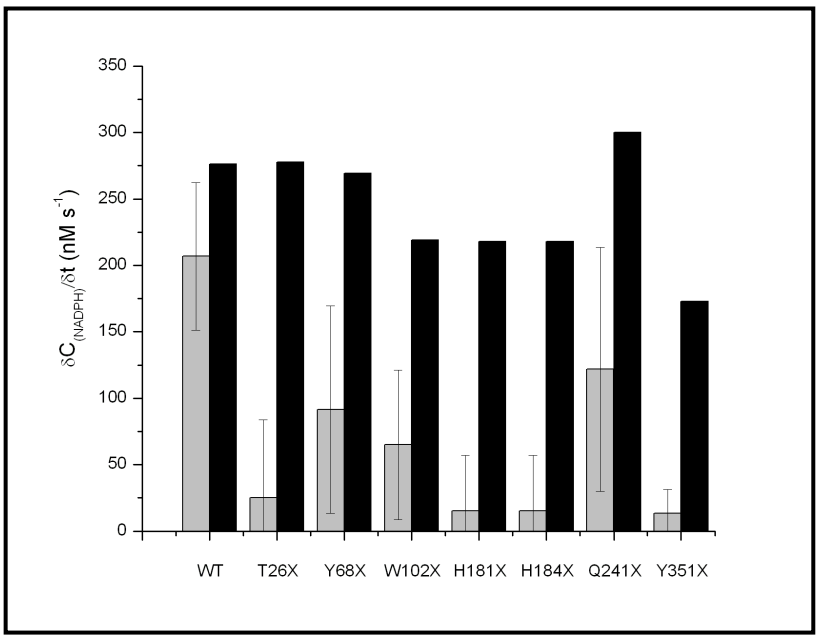


Figure 4.32. Comparative activities of 96 well purified enzyme blocks. Assays were performed according to the robotic protocol with 0.1 mM NADPH, 1 mM CH, 50 µl purified enzyme extract; activities were calculated over the first 126 s following insertion of the plate into the reader. Average of all wells on the plate (grey), standard deviation (error bars) and maximal rate (black) are provided.

An important point to note is that the majority of mutants identified from libraries had activities far lower than that obtained for $PETNR_{WT}$, whilst the maximal rate was generally comparable to that of wild type enzyme. The majority of clones in the library are seriously catalytically impaired (Figure 4.32).

Those clones which maintain high levels of activity in the individual screens (Figures 4.25 - 4.31) were sequenced and the majority were identified as $PETNR_{WT}$ clones. These clones are not thought to represent the entire wild type population of the block due to the fact that only the top 5 ranked clones (i.e. those clones which are most active with the NADPH/CH couple as monitored by the $\delta A_{340}/\delta t$ (s^{-1})) were sequenced. This observation supports the hypothesis that the majority of mutations compromise rather than improve activity.

The extent to which the activity is compromised, however, is not uniform. A number of libraries such as Y68X, W102X and Q241X (Figures 4.26/27/30/31) have retained a large fraction of the wild type levels of activity. However libraries such as H181X, H184X and Y351X (Figures 4.28/29/31/32) have poor residual activity. In particular, for libraries H181X and H184X (Figures 4.28/28/32) an explanation can be advanced that these residues, which are conserved in OYE homologues, have been implicated in binding the carbonyl group of the cyclohexenone substrate (Section 1.4.3.4, [40]). Mutation of these residues is therefore likely to affect substrate binding. Sequencing confirmed that after $PETNR_{WT}$ the next most active mutant was H184N. This substitution of histidine to asparagine is seen in some natural OYE homologue enzymes (e.g. OYE1 and 2 and MR [14, 16, 18]).

In the case of the Y351X library clones were characterised by a far lower level of activity than the other libraries (Figures 4.31/32), including the wild type clones present. This could potentially indicate a reduced concentration in this individual 96 well purification experiment, and may not necessarily represent a library with a lower level of wild type activity.

The library altered in the identity of Q241 ($PETNR_{WT}$) was found to have variable activity and was the only library for which wild type was not included in the most active enzymes (Figures 4.30/32). Also, within the library most enzymes retain high

levels of activity compared to $PETNR_{WT}$. This suggests many of the generated mutants are viable proteins and retain catalytic competence, this implies that this library could yield improved enzyme variants.

The top ranked enzyme from each library that was not $PETNR_{WT}$ was carried forward for further characterisation in Chapter 6. The aim of this characterisation was confirm these mutant enzymes were impaired compared to $PETNR_{WT}$ by determining specific activity and biotransformation profiles.

4.6 Summary

The results from this chapter have described the generation of library DNA by mutagenic PCR methods at a range of positions which are implicated as being important in catalysis and substrate binding on the basis of structural and mechanistic studies. These libraries were archived as glycerol stocks within a 96 well format and analysed using a robotics platform. Libraries were sequenced to determine the degree of randomisation. The libraries were then subjected to inductive growth and characterised with respect to protein expression. Libraries were then micro-scale purified in a 96 well format. A method was developed that enabled approximately equal recovery of target protein in each sample well of the plate and comparative assay across wells without the need for protein quantification. Purified protein libraries were demonstrated to display a variety of activities defined by the function of the targeted residue or experimental parameters. Where possible, wild type clones were identified from the reactions to act as internal controls for future screening experiments.

In summary this chapter has established a number of DNA libraries for use in the rapid screening of variants using the robust high throughput robotics methods described in chapter 3 under anaerobic conditions. These libraries were generated and their randomisation was validated by random sequencing of clones. Libraries were then expressed and purified according to the protocols derived from work in chapter 3. Libraries were subjected to initial characterisation assays with the model substrates NADPH/CH and demonstrated a variety of activities. Following this in-depth assessment of library quality these libraries can be used for screening with a panel of novel substrates, which will be discussed in Chapter 6, with the intention of isolating novel variants capable of accommodating non-natural substrates.

5 Screening of $PETNR_{His}$ libraries against a panel of novel substrates

5.1 Screening of H181X and H184X libraries to accommodate novel activating groups

Initial experiments focussed on trying to modify the PETNR active site to accept novel functionalities in the position occupied by the carbonyl in the model substrate cyclohexenone. A range of functionalities have been shown to be accommodated in this position in PETNR including keto, nitro, di-keto and di-acid (Section 1.3.2; [54, 59, 63]).

While some enone derivatives are reduced by PETNR and other OYE homologues others are not, for example nitrocyclohex-1-ene (NCH) has been demonstrated to act as a substrate whilst the respective acid and nitrile compounds were trialled and demonstrated no activity (Section 1.3.2; [54, 59, 63]). Similarly (E)-1-nitro-2-phenylethene has been demonstrated to be a substrate for PETNR, however the acid and nitrile derivatives are not (Section 1.3.2). Interestingly, di-acid compounds have been demonstrated to be substrates for PETNR suggesting the acid functionality can be accommodated (Section 1.3.2), however a literature search has found no evidence of nitrile reduction by OYE's.

Structural studies of OYE homologues complexed with oxidative substrates and inhibitors (Section 1.5.3.4; [40]) have highlighted a key role for residues H181 and H184 in binding the electron withdrawing or activating groups of these substrates. Thus, these libraries were the first target for mutagenesis to see if different residues at positions 181 / 184 could accommodate different activating groups. However, it is important to note that these residues are also implicated in FMN binding and NADPH reduction (Section 1.4.1.3 and 1.4.2; [69] and [38]) suggesting a subset of the library could be catalytically compromised.

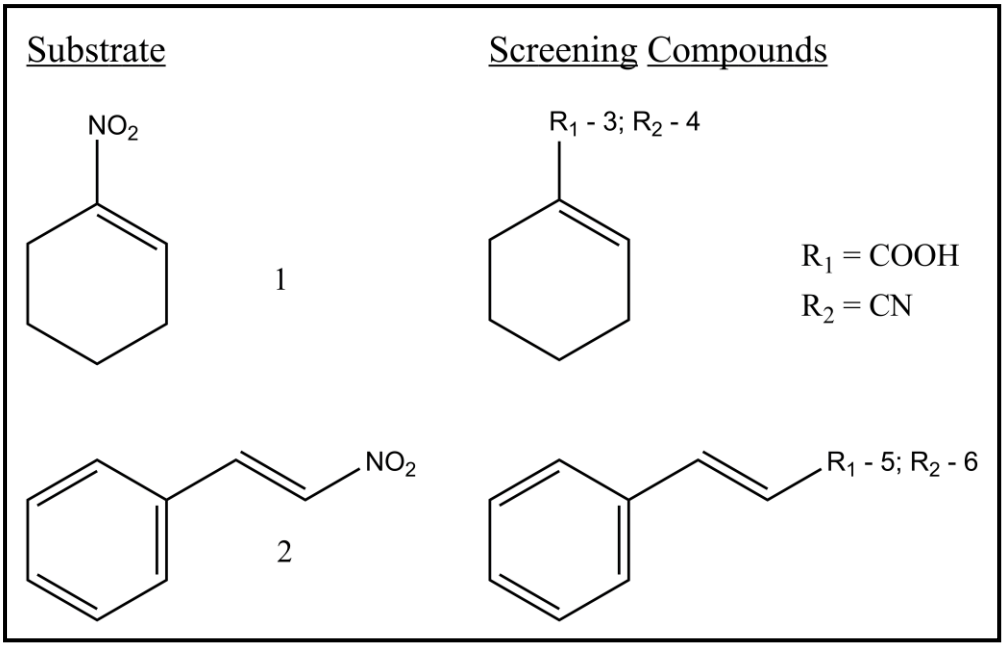


Figure 5.1. Schematic representation of screening substrates, compounds on the left are known substrates of PETNR whereas compounds on the right are non substrates with which to screen mutagenic libraries. Panel 1, nitrocyclohex-1-ene; Panel 2, (E)-1-Nitro-2-phenylethene; Panel 3, cyclohexene-1-carboxylic acid; Panel 4, cyclohexene-1-carbonitrile; Panel 5, *trans*-cinnamic acid; Panel 6, cinnamonitrile;

In order to attempt to widen the substrate specificity of PETNR, libraries were generated at H181 and H184 as discussed previously (Section 4.1). These libraries were expressed and purified (Section 4.3, 4.4) before being assayed with novel substrates (Section 2.2.4.10.6). Following assay, Z factors were calculated from screening results (Section 2.2.4.10.7). Hits were defined by a Z factor of 2 or above (up to a maximum of 2 per library per substrate). If no hits were present according to these criteria the top ranked clone was considered a hit. In these experiments, clones previously identified as $PETNR_{His}$ WT (Section 4.5), were used as internal controls in the assay reactions. DNA from positive hits was purified (Section 2.2.1.2) and sequenced (Section 2.2.1.5).

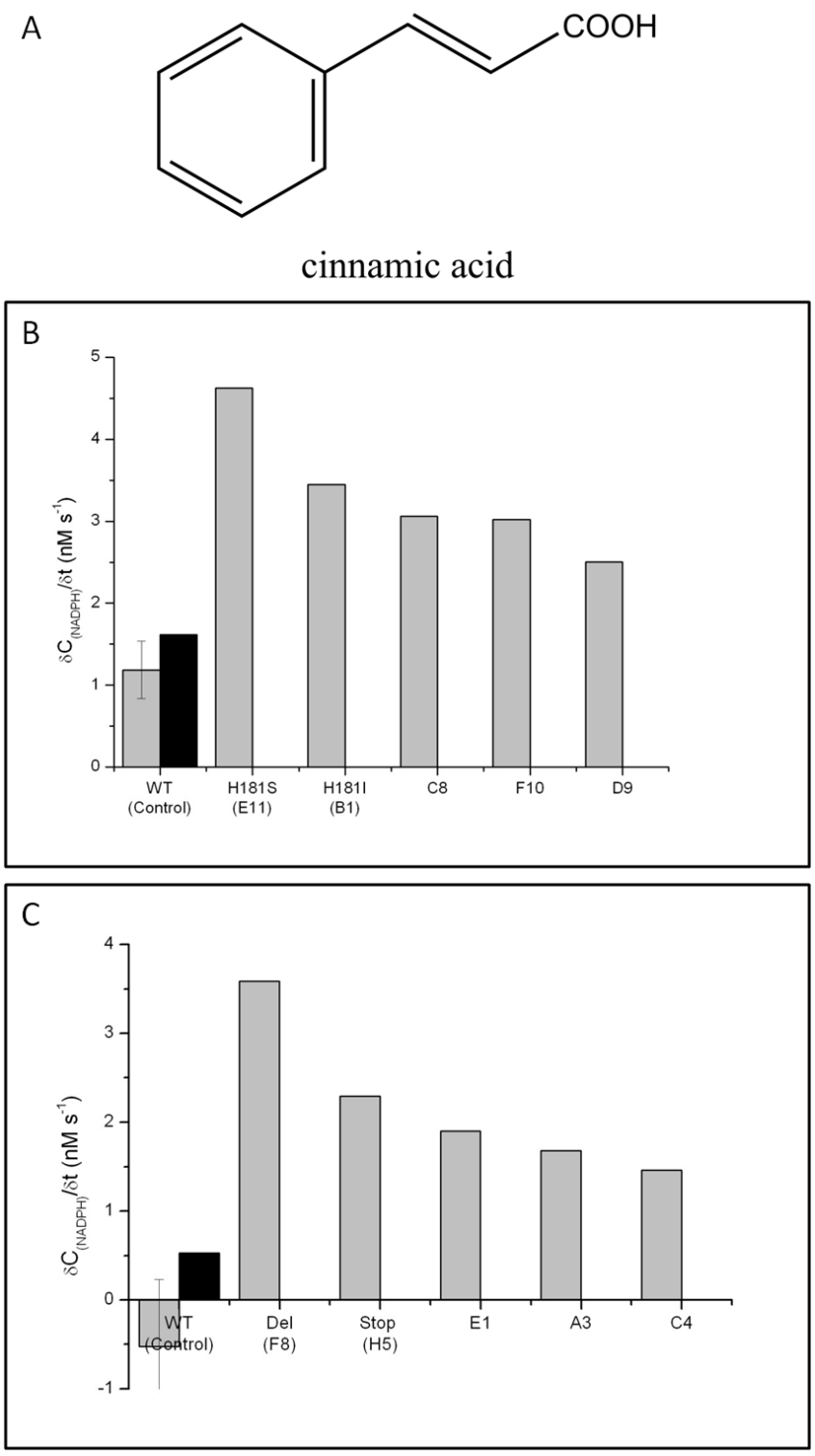


Figure 5.2. Screening of H181X and H184X library eluates against *trans*-cinnamic acid. Panel A, *trans*-cinnamic acid structure; Panel B, H181X screen; Panel C, H184X screen. Reactions comprised 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in 50 mM potassium phosphate buffer pH7 containing 5 % ethanol. Reaction rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of internal WT controls.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
H181S (E11)	4.6	4.2
H181I (B1)	3.4	2.9
H181- (C8)	3.1	2.4
H181- (F10)	3.0	2.4
H181- (D9)	2.5	1.8
Del (F8)	3.6	2.9
STOP (H5)	2.3	2.0
H184- (E1)	1.9	1.7
H184- (A3)	1.7	1.6
H184- (C4)	1.5	1.4

Table 5.1. Summary table showing data from H181X and H184X screening with *trans*-cinnamic acid. Potential hits are highlighted; where data were available well and clone details are given.

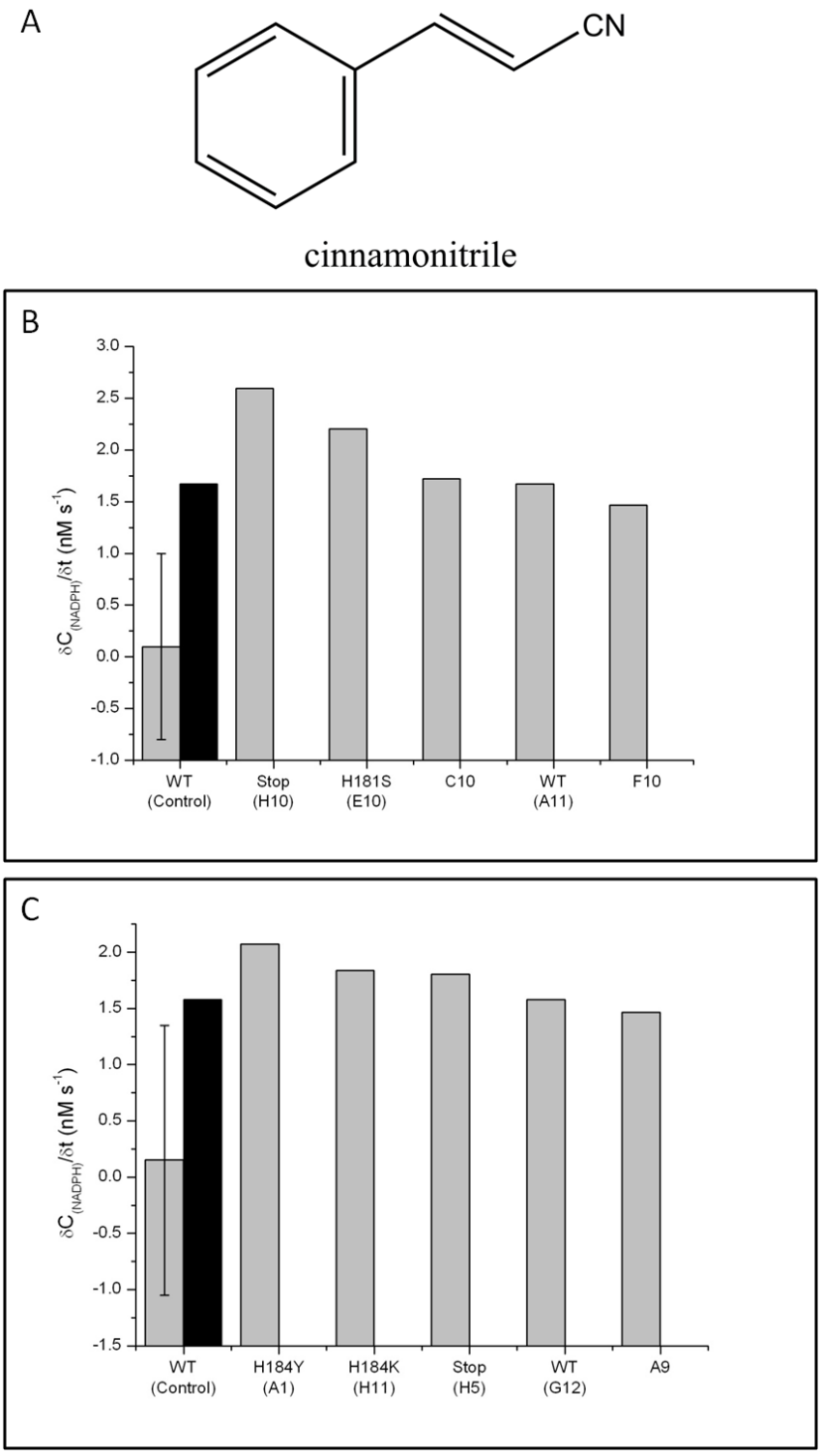


Figure 5.3. Screening of H181X and H184X library eluates against cinnamonitrile. Panel A, cinnamonitrile structure; Panel B, H181X screen; Panel C, H184X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction (δC/δt (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of internal WT controls.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
STOP (H10)	2.6	2.7
H181S (E10)	2.2	2.3
H181- (C10)	1.7	1.8
WT (A11)	1.7	1.8
H181- (F10)	1.5	1.5
H184Y (A1)	2.1	2.4
H184K (H11)	1.8	2.2
STOP (H5)	1.8	2.2
WT (G12)	1.6	2.0
H184- (A9)	1.5	1.9

Table 5.2. Summary table showing data from H181X and H184X screening with cinnamonitrile acid. Potential hits are highlighted, where data were available well and clone details are given.

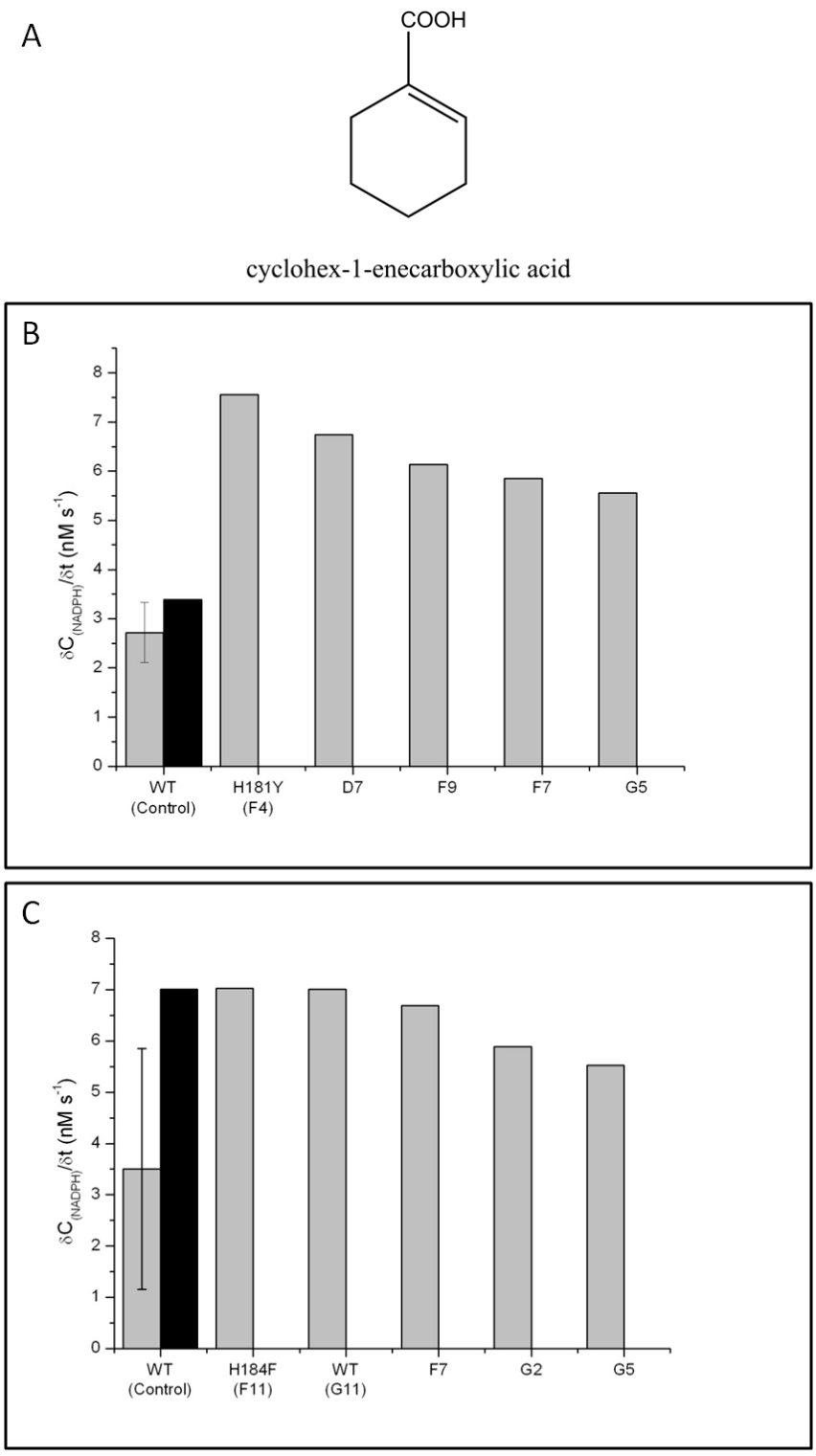


Figure 5.4. Screening of H181X and H184X library eluates against cyclohexene-1-carboxylic acid. Panel A, cyclohexene-1-carboxylic acid structure; Panel B, H181X screen; Panel C, H184X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of Internal WT controls.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
H181Y (F4)	7.6	2.1
H181- (D7)	6.7	1.8
H181- (F9)	6.1	1.5
H181- (F7)	5.8	1.4
H181- (G5)	5.6	1.2
H184F (F11)	7.0	2.5
WT (G11)	7.0	2.5
H184- (F7)	6.7	2.4
H184- (G2)	5.9	2.0
H184- (G5)	5.5	1.8

Table 5.3. Summary table showing data from H181X and H184X screening with 1-cyclohexene-1-carboxylic acid. Potential hits are highlighted, where data were available well and clone details are given.

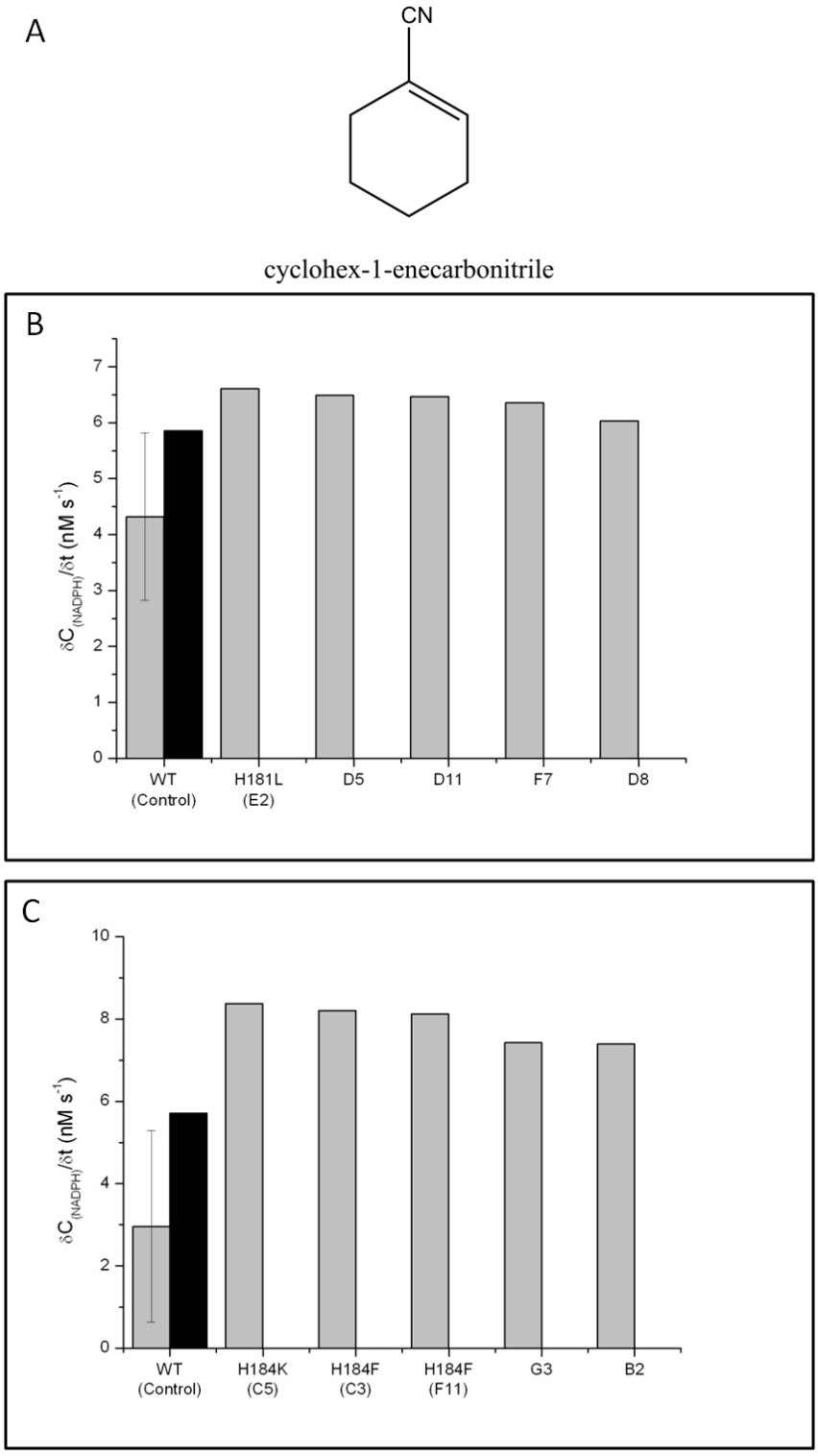


Figure 5.5. Screening of H181X and H184X library eluates against cyclohexene-1-carbonitrile. Panel A, cyclohexene-1-carbonitrile structure; Panel B, H181X screen; Panel C, H184X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM CH in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of Internal WT controls.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
H181L (E2)	6.6	1.9
H181- (D5)	6.5	1.9
H181- (D11)	6.5	1.8
H181- (F7)	6.4	1.8
H181- (D8)	6.0	1.6
H184K (C5)	8.4	1.9
H184F (C3)	8.2	1.9
H184F (F11)	8.1	1.8
H184- (G3)	7.4	1.6
H184- (B2)	7.4	1.5

Table 5.4. Summary table showing data from H181X and H184X screening with cyclohexene-1-carbonitrile. Potential hits are highlighted, where data were available well and clone details are given

The results indicate that all mutants and WT clones showed poor activity against these substrates. In the case of *trans*-cinnamic acid potential hits were obtained for the H181X library, however the same screen for the H184X library highlighted clones with significant mutations not conducive with improved catalysis (internal STOP and deletion).

These data suggest that the rates highlighted as hits in the experiment may not be genuine hits but outliers of a normal distribution around a low mean. As a consequence of this the errors inherent in the purification and assay are far larger than variation attributable to any activity improved or not so. This can be seen frequently throughout the screening data for the first experiment.

Despite the low likelihood of these results being genuine hits any clones matching the selection criteria discussed previously were purified on a small scale and taken forward to the next stage of experimentation. Hits taken forward for further experimentation are outlined in Table 5.5.

Substrate	H181X	H184X
Trans-cinnamic acid	H181S, H181I	n/a
Cinnamonitrile	H181S,	H184Y, H184L
1-Cyclohexene-1-carboxylic acid	H181Y	H184F
Cyclohexene-1-carbonitrile	H181L	H184L

Table 5.5. Hits taken forward for further characterisation from screening of H181X, H184X single site saturation libraries for novel functionality.

Also, the use of internal control reactions has proved problematic in these cases often providing large error values because variation due to noise represents a far higher fraction of total activity in smaller rates, hence experimental 'noise' makes a more significant contribution to error. For future screening reactions an external control block of $PETNR_{His}$ WT will be recorded for all reactions.

5.2 Screening of T26X, Y68X, W102X, Q241X and Y351X libraries for expanded substrate acceptance.

In addition to studies attempting to modify binding interactions a number of variations of known PETNR substrates with much lower activity were identified as outlined in figure 5.6. These substrates involve the substitution of methyl groups at various positions on the known substrates cyclohexenone, cyclopentanone and in the case of OYE cinnamaldehyde (Section 1.3.2.4, [59, 62, 63]) and the methyl substituents have largely been shown to be poor substrates. In these experiments WT blocks were used in preference to internal controls in order to improve the accuracy of the estimate used for $PETNR_{WT}$ by minimising the effect of outliers.

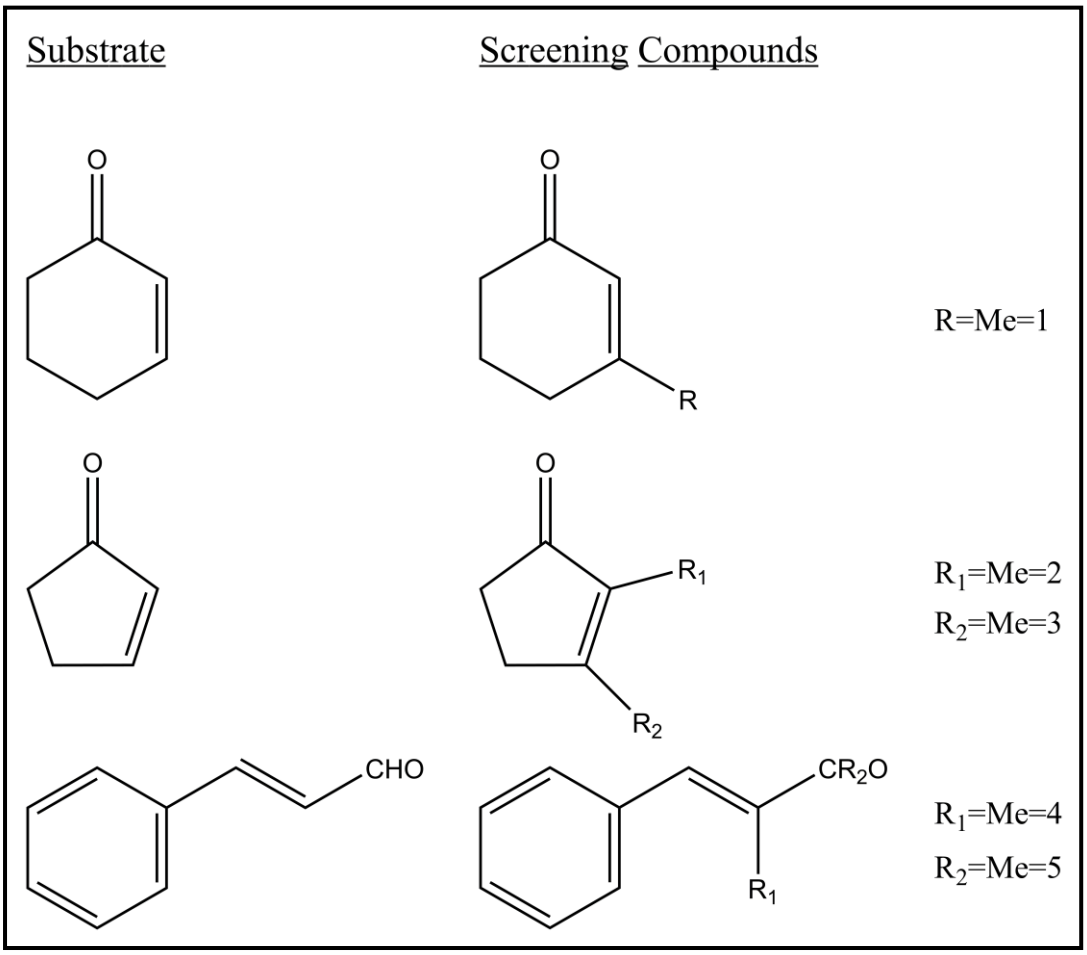


Figure 5.6. Table showing the screening substrates used to widen the substrate acceptance of PETNR along with the respective substrates which are active with WT. Panel 1, 3-methylcyclohex-2-enone; Panel 2, 2-methylcyclopent-2-enone; Panel 3, 3-methylcyclopent-2-enone; Panel 4, α-Methyl-*trans*-cinnamaldehyde; Panel 5, methyl-*trans*-cinnamate.

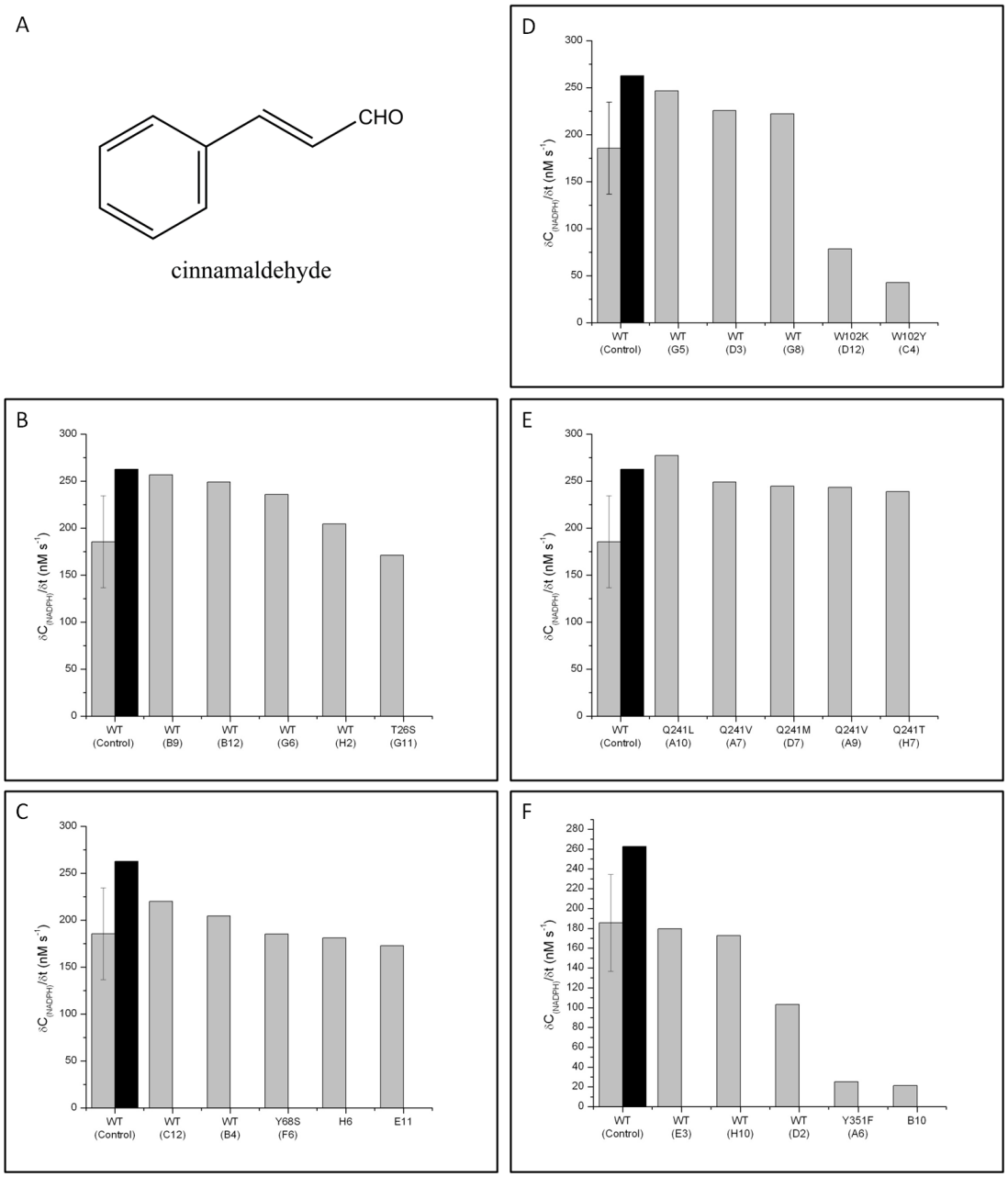


Figure 5.7. Screening of a range of $PETNR_{His}$ libraries against *trans*-cinnamaldehyde. Panel A, *trans*-cinnamaldehyde structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
WT (B9)	257	4.4
WT (B12)	249	4.3
WT (G6)	236	4.0
WT (H2)	204	3.4
T26S (G11)	171	2.8
WT (C12)	220	2.5
WT (B4)	204	2.2
Y68S (F6)	185	1.9
Y68- (H6)	181	1.9
Y68- (E11)	173	1.7
WT (G5)	247	5.5
WT (D3)	226	5.0
WT (G8)	222	4.9
W102L (D12)	78	1.4
W102Y (C4)	43	0.5
Q241L (A10)	277	2.0
Q241V (A7)	249	1.6
Q241M (D7)	245	1.6
Q241V (A9)	243	1.6
Q241T (H7)	239	1.5
WT (E3)	180	6.3
WT (H10)	173	6.1
WT (D2)	103	3.5
Y351F (A6)	25	0.6
Y351- (B10)	21	0.4

Table 5.6. Summary table showing data from a range of libraries screened with *trans*-cinnamaldehyde. Potential hits are highlighted, where data were available well and clone details are given.

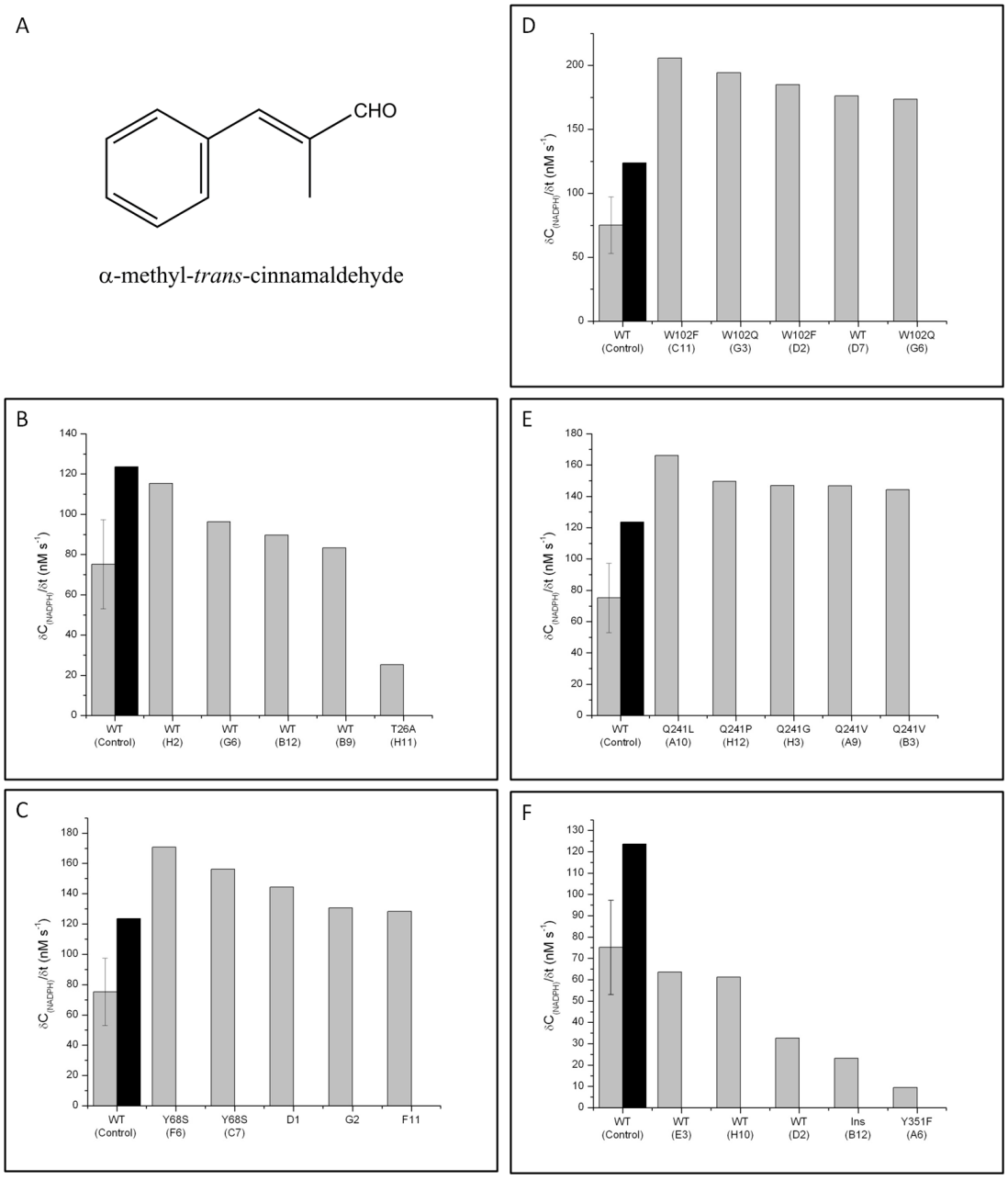


Figure 5.8. Screening of a range of $PETNR_{His}$ libraries against α-methyl-*trans*-cinnamaldehyde. Panel A, α-methyl-*trans*-cinnamaldehyde structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction (δC/δt (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
WT (H2)	115	5.4
WT (G6)	96	4.4
WT (B12)	90	4.1
WT (B9)	83	3.8
T26A (H11)	25	0.9
Y68S (F6)	171	2.8
Y68S (C7)	156	2.5
Y68- (D1)	144	2.2
Y68- (G2)	131	1.9
Y68- (F11)	128	1.9
W102F (C11)	206	2.3
W102Q (G3)	194	2.1
W102F (D2)	185	1.9
WT (D7)	176	1.8
W102Q (G6)	174	1.8
Q241L (A10)	166	2.0
Q241P (H12)	150	1.6
Q241G (H3)	147	1.6
Q241V (A9)	147	1.6
Q241V (B3)	144	1.5
WT (E3)	64	5.8
WT (H10)	61	5.5
WT (D2)	33	2.9
Ins (B12)	23	2.0
Y351F (A6)	9	0.7

Table 5.7. Summary table showing data from a range of libraries screened with α-methyl-*trans*-cinnamaldehyde. Potential hits are highlighted, where data were available well and clone details are given.

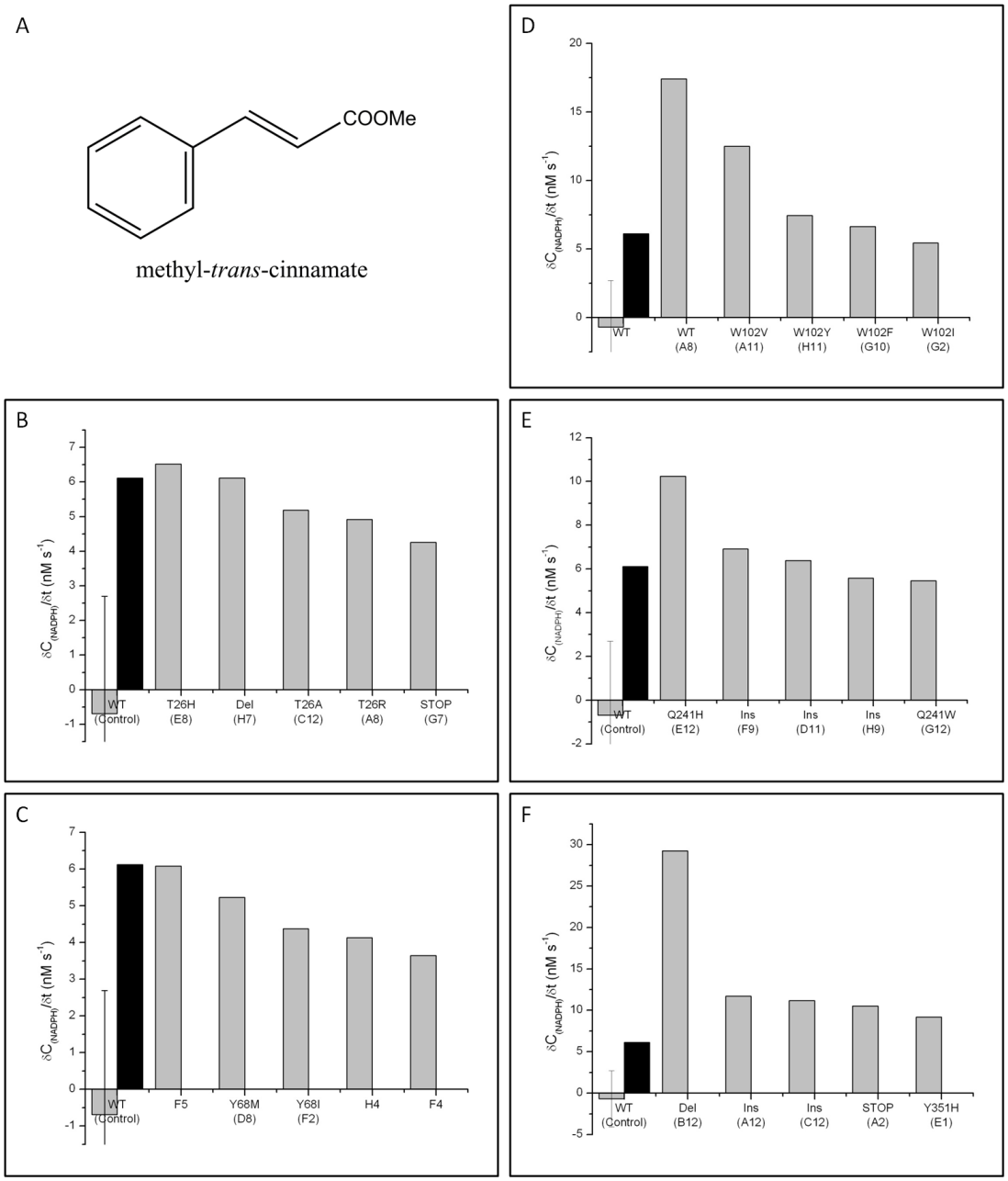


Figure 5.9. Screening of a range of $PETNR_{His}$ libraries against methyl-*trans*-cinnamate. Panel A, α-methyl-*trans*-cinnamate structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction (δC/δt (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
T26H (E8)	6.5	1.8
Del (H7)	6.1	1.7
T26A (C12)	5.2	1.5
T26R (A8)	4.9	1.4
STOP (G7)	4.3	1.2
Y68- (F5)	6.1	2.1
Y68M (D8)	5.2	1.8
Y68I (F2)	4.4	1.6
Y68- (H4)	4.1	1.5
Y68- (F4)	3.6	1.4
WT (A8)	17.4	2.8
W102V (A11)	12.5	2.0
W102Y (H11)	7.4	1.3
W102F (G10)	6.6	1.1
W102I (G2)	5.5	1.0
Q241H (E12)	10.2	3.1
Ins (F9)	6.9	2.2
Ins (D11)	6.4	2.0
Ins (H9)	5.6	1.8
Q241W (G12)	5.5	1.8
Del (B12)	29	3.0
Ins (A12)	12	1.3
Ins (C12)	11	1.3
STOP (A2)	11	1.2
Y351H (E1)	9	1.1

Table 5.8. Summary table showing data from a range of libraries screened with methyl-*trans*-cinnamate. Potential hits are highlighted, where data were available well and clone details are given.

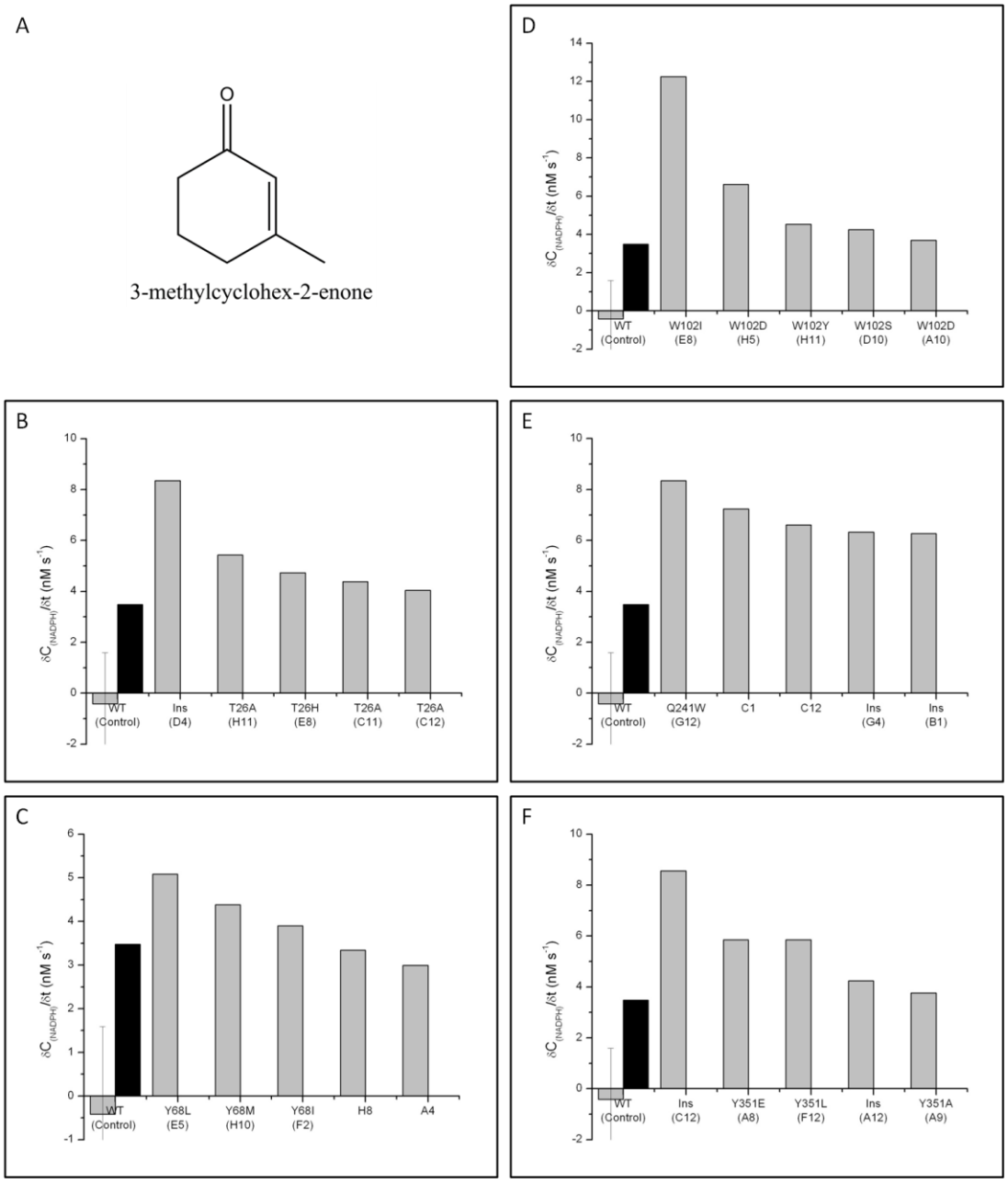


Figure 5.10. Screening of a range of $PETNR_{His}$ libraries against 3-Methyl-2-cyclohexen-1-one. Panel A, 3-methyl-2-cyclohexen-1-one structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block..

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
Ins (D4)	8.4	3.5
T26A (H11)	5.4	2.1
T26H (E8)	4.7	1.7
T26A (C11)	4.4	1.5
T26A (C12)	4.0	1.4
Y68L (E5)	5.1	0.8
Y68M (H10)	4.4	0.7
Y68I (F2)	3.9	0.6
Y68- (H8)	3.3	0.5
Y68- (A4)	3.0	0.4
W102I (E8)	12	4.0
W102D (H5)	7	2.0
W102Y (H11)	5	1.3
W102S (D10)	4	1.2
W102D (A10)	4	1.0
Q241W (G12)	8.4	2.6
Q241- (C1)	7.2	2.2
Q241- (C12)	6.6	2.0
Ins (G4)	6.3	1.9
Ins (B1)	6.3	1.9
Ins (C12)	8.6	2.2
Y351E (A8)	5.8	1.5
Y351L (F12)	5.8	1.5
Ins (A12)	4.2	1.1
Y351A (A9)	3.8	0.9

Table 5.9. Summary table showing data from a range of libraries screened with 3-methyl-2-cyclohexenone. Potential hits are highlighted, where data were available well and clone details are given.

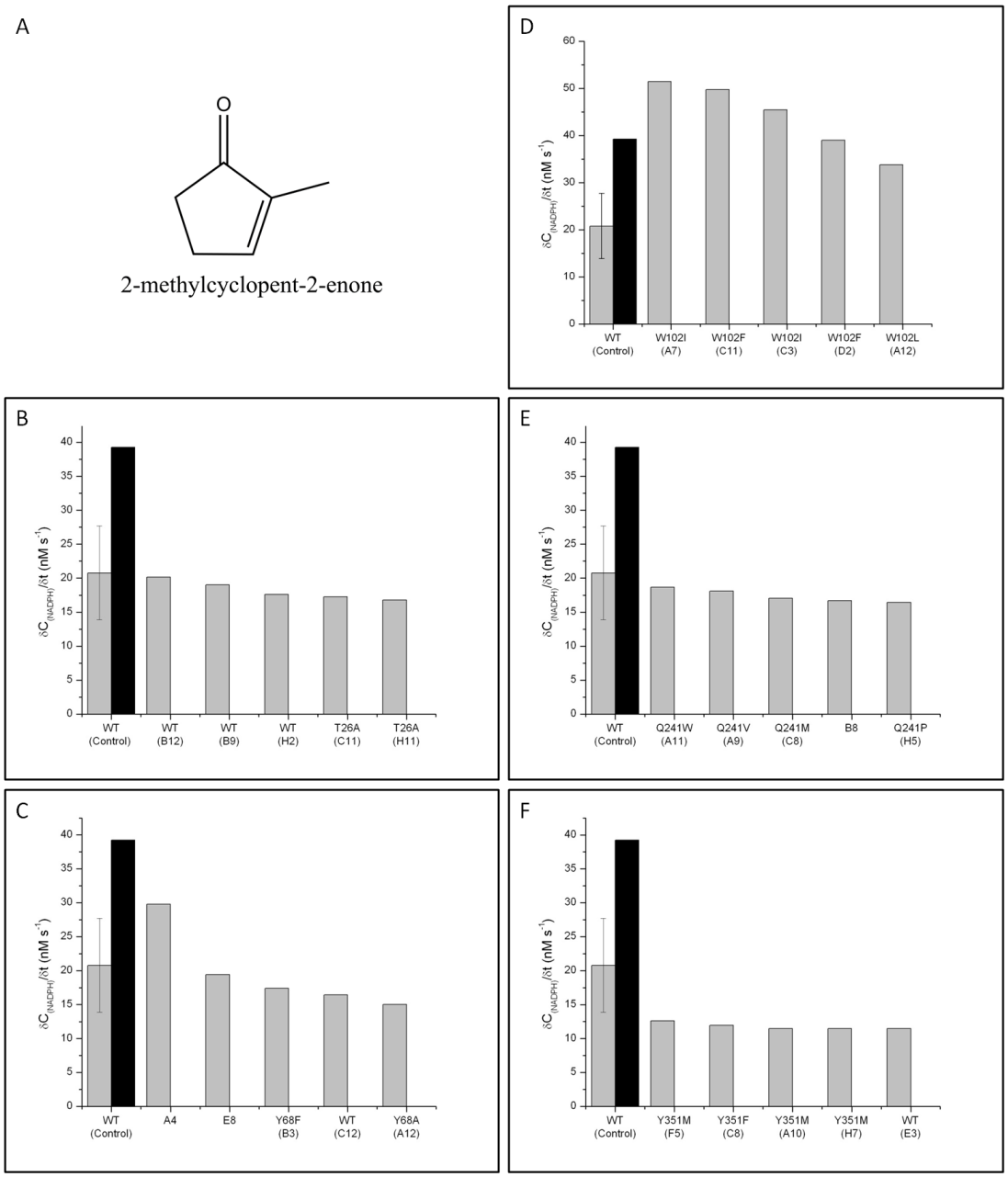


Figure 5.11. Screening of a range of $PETNR_{His}$ libraries against 2-methyl-2-cyclopenten-1-one. Panel A, 2-methyl-2-cyclopenten-1-one structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction (δC/δt (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
WT (B12)	20	3.5
WT (B9)	19	3.3
WT (H2)	18	3.0
T26A (C11)	17	2.9
T26A (H11)	17	2.9
Y68- (A4)	30	4.0
Y68- (E8)	19	2.2
Y68F (B3)	17	1.9
Y68Y (C12)	16	1.7
Y68A (A12)	15	1.5
W102I (A7)	51	3.3
W102F (C11)	50	3.2
W102I (C3)	45	2.8
W102F (D2)	39	2.3
W102L (A12)	34	1.9
Q241W (A11)	19	2.7
Q241V (A9)	18	2.6
Q241M (C8)	17	2.4
Q241- (B8)	17	2.3
Q241P (H5)	16	2.2
Y351M (F5)	13	2.6
Y351P (C8)	12	2.4
Y351M (A10)	11	2.3
Y351M (H7)	11	2.3
Y351T (E3)	11	2.3

Table 5.10. Summary table showing data from a range of libraries screened with 2-methyl-2-cyclopentenone. Potential hits are highlighted, where data were available well and clone details are given.

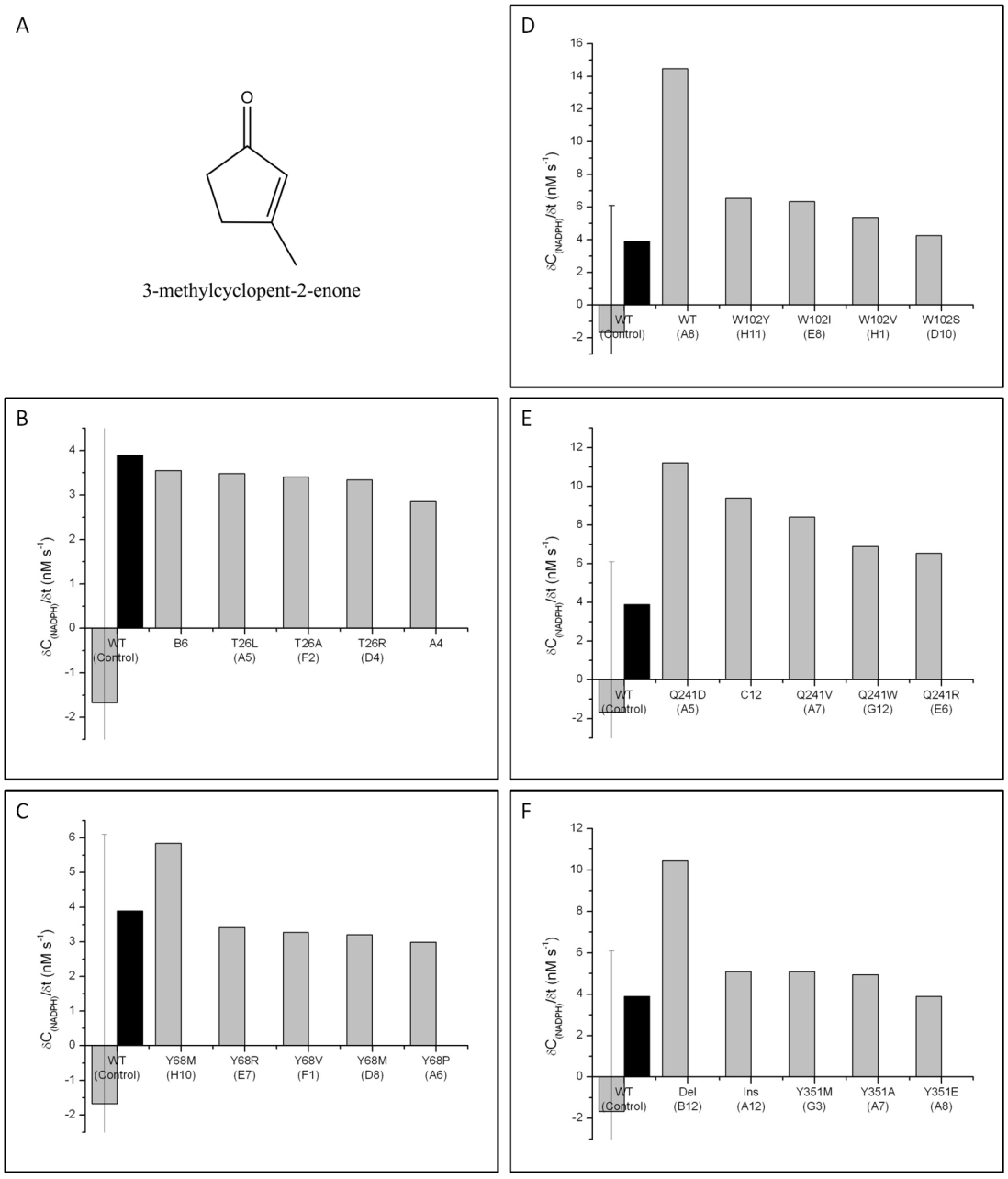


Figure 5.12. Screening of a range of $PETNR_{His}$ libraries against 3-methyl-2-cyclopenten-1-one. Panel A, 3-methyl-2-cyclopenten-1-one structure; Panel B, T26X screen; Panel C, Y68X screen; Panel D, W102X screen; Panel E, Q241X screen; Panel F, Y351X screen. Reactions were composed of 50 µl PETNR library eluate, 0.1 mM NADPH, 1 mM substrate in potassium phosphate buffer containing 5 % ethanol. Rates were calculated by monitoring A_{340} over the first 126 s of the reaction ($\delta C/\delta t$ (nM s^{-1})). WT represents the average (grey), standard deviation (error bars) and maximum (black) of an external $PETNR_{WT}$ control block.

Clone	$\delta C_{(NADPH)}/\delta t$ (nM s^{-1})	Z score
T26- (B6)	3.5	1.3
T26L (A5)	3.5	1.3
T26A (F2)	3.4	1.2
T26R (D4)	3.3	1.2
T26- (A4)	2.9	1.0
Y68M (H10)	5.8	1.9
Y68A (E7)	3.4	1.0
Y68V (F1)	3.3	1.0
Y68M (D8)	3.2	0.9
Y68P (A6)	3.0	0.9
WT (A8)	14	4.7
W102Y (H11)	7	1.9
W102I (E8)	6	1.9
W102V (H1)	5	1.5
W102S (D10)	4	1.1
Q241D (A5)	11.2	3.2
Q241- (C12)	9.4	2.7
Q241V (A7)	8.4	2.4
Q241T (G12)	6.9	2.0
Q241A (E6)	6.5	1.9
Del (B12)	10.4	2.9
Ins (A12)	5.1	1.4
Y351M (G3)	5.1	1.4
Y351A (A7)	4.9	1.4
Y351E (A8)	3.9	1.1

Table 5.11. Summary table showing data from a range of libraries screened with 3-methyl-2-cyclopentenone. Potential hits are highlighted, where data were available well and clone details are given.

A large number of mutants with possibly improved activity were identified in the screening protocols and were selected for further characterisation. A wide range of activities were observed from the good substrate *trans*-cinnamaldehyde to the poorer substrates including the methyl substituted cyclic derivatives. These potential hits are detailed in the table below (Table 5.12).

Substrate	T26X	Y68X	W102X	Q241X	Y351X
Trans-cinnamaldehyde	T26S	Y68S	W102L	Q241L	Y351F
α-Methyl-*trans*-cinnamaldehyde	T26A	Y68S	W102F, W102Q	Q241L	Y351F
methyl *trans*-cinnamate	T26H	Y68M	W102V	Q241H	Y351H
3-Methyl-2-cyclohexenone	T26A	Y68L	W102I, W102D	Q241W	Y351E
2-Methyl-2-cyclopentenone	T26A	Y68F	W102I, W102F	Q241W, Q241V	Y351M, Y351P
3-Methyl-2-cyclopentenone	T26L	Y68M	W102Y	Q241D, Q241V	Y351M

Table 5.12. Potential hits taken forward for further characterisation from screening of T26X, Y68X, W102X, Q241X and Y351X single site saturation libraries for expanded substrate acceptance.

5.3 Discussion

Overall 38 deep well blocks were screened against 10 novel substrates, corresponding to over 3,600 individual screening reactions. The application of strict selection criteria yielded a total of 45 possible hits representing ~1% of the total number of reactions screened. Without further characterisation of library hits (Chapter 6) discussion of potential hits at this juncture is difficult. However a number of points of interest can be raised from the screening data alone.

The first conclusion from this experiment is an understanding of the relative activities of the substrates which are screened with WT enzyme as derived from control reactions. This indicates the range of native activities to be improved upon ranging from effectively no activity, including *trans*-cinnamic acid, to much more active substrates, including *trans*-cinnamaldehyde. Few of these activities however come close to the activity achieved with the NADPH / CH couple (Table 5.13).

Substrate	Activity (nM s^{-1})	Standard deviation (nM s^{-1})
2-Cyclohexen-1-one	207	55
Trans-cinnamic acid	*nd*	*n/a*
Cinnamonitrile	*nd*	*n/a*
1-Cyclohexene-1-carboxylic acid	*nd*	*n/a*
Cyclohexene-1-carbonitrile	*nd*	*n/a*
Trans-cinnamaldehyde	186	49
α-Methyl-trans-cinnamaldehyde	75	22
Methyl trans-cinnamate	0.688	3.38
3-Methyl-2-cyclohexenone	20.8	6.9
2-Methyl-2-cyclopentenone	1.67	7.77
3-Methyl-2-cyclopentenone	0.419	2.01

Figure 5.13. Summary of WT activity derived from single $PETNR_{His}$ block purifications, used as control for screening reactions; *nd*, none detected.

A number of false positives were also identified at this point, which do not encode functional proteins. In this case, false positives represent a population of hits which

when sequenced reveal serious mutations such as large deletions / insertions and internal STOP codons. Notably, these false positives occur mainly in reactions with a lower activity. It is likely that these hits are due either to outliers around a low mean (Section 5.1.5) or mixing artefacts affecting a very low mean since the fitting method is insensitive to such phenomena. Under such conditions in order to be confident of a hit being genuine a far larger increase in activity is required.

A number of assays above have highlighted a number of WT codons, particularly with substrates with a reasonable level of WT activity. This phenomenon is due to the fact that the majority of generated mutants have compromised activity, hence WT clones are the only ones with activity remaining. This also serves to demonstrate the efficacy of the assay, for example if wild type clones were not highlighted from experiments with known WT active substrate couples (e.g. NADPH / CH) it would suggest problems were present with the assay. In these cases “the next best” clone was analysed.

6 Characterisation of identified variants from library screening

The screening of 3600 clones from 8 libraries with 11 substrates generated a large number of potentially improved variants (as described in Chapter 5). However, because of the high incidence of false positives [e.g. i) outliers surrounding a low mean or ii) mixing artefacts as described in Chapter 5], it was necessary to confirm that positive hits were actually improved catalysts. Each mutant was characterised by investigating its steady state reaction with CH and carrying out biotransformation reactions to positively identify the product(s).

Variants of interest were grown and purified on an intermediate growth scale (1L media) according to the standard $PETNR_{His}$ purification protocol (Section 2.2.4.5). Variants were purified in order to minimise potential competing reactions either modifying substrate / product, or utilising NADPH. It is important to consider that previous attempts to engineer OYE homologues have been hampered by the presence of *E. coli* keto-reductase enzymes which were not eliminated by other workers [61].

Following purification, mutants were analysed by determination of the specific activity towards the substrate of interest (the substrate used in the screen which identified that particular mutant) with NADPH as reductant, using spectrophotometric methods (Section 2.2.3.2). The assay monitors the oxidation of NADPH associated with the turnover of the alkene substrate to the alkane product, however implicit in this method are a number of assumptions including that the NADPH is turned over only by PETNR and only in response to the desired oxidative half reaction. Therefore, the reaction of substrate and the formation of product was monitored directly by using GC to determine the concentration of substrate and product following 24 hr biotransformation reactions (Section 2.2.3.5).

6.1 Characterisation of identified variants from screens performed with cyclohexenone as substrate.

Screening with CH as the oxidative substrate was performed as part of the initial characterisation of the PETNR libraries (Section 4.5). These experiments were performed in order to assess the degree to which mutant libraries were compromised with respect to diagnostic wild type activity. The specific activity data for all the variant enzymes (Table 6.1) demonstrate a considerably lower steady state turnover rate when compared with the wild type control data. When the activity retained by variants isolated from CH screens is contrasted with the WT controls for other substrates used for screening it is clear that a substantial amount of activity is still retained (see Section 6.2 and 6.3). All library screening reactions with CH, with the exception of Q241X, showed that the majority of the top ranked hits were in fact WT clones. Interestingly, the only library that produced a clone with an elevated level of activity over other library variants was Q241X. This library had the most variety in the top ranked hits, suggesting that a number of variants from this library retain wild-type-like catalytic competence.

In contrast, all the biotransformation reactions were shown to have proceeded to near completion with the CH substrate, irrespective of their lower steady state rates (Table 6.1). This is most likely due to the length of the biotransformation reactions. The most active variants react to completion in a short period of time; however despite the less active variants reacting slower there is still sufficient time available to react to completion. Hence, at the endpoint of the reaction where data are recorded, differences in the rate of product formation are impossible to determine. This “catching up” phenomenon is also observed for other reactions discussed in this chapter.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)
WT	5.82	0.68	100	100
T26S	0.84	0.17	100	96
Y68L	1.21	0.11	100	82
W102Y	0.84	0.14	100	90
Q241T	1.81	0.03	100	88
Y351F	1.21	0.03	100	85

Table 6.1. Specific activity and GC analysis of reactions with $PETNR_{His}$ variants identified from screens with CH as substrate. Specific activity reactions were composed of 50 nM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM CH in 50mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC Reactions were composed of 5 mM CH, 6 mM NADPH and 2 µM enzyme in 50mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml.

6.2 Characterisation of identified variants from screens performed with substrates differing in their activating group located next to the unsaturated C=C bond.

Previous studies with $PETNR_{His}$ showed no activity towards substrates *trans*-cinnamic acid, cinnamonitrile, 1-cyclohexene-1-carboxylic acid or cyclohexene-1-carbonitrile [59]. The screening of libraries H181X and H184X generated a few possible hits with these substrates.

Specific activity data indicated no significantly improved variants, a conclusion which was supported by the absence of a detectable product in the analysis of the biotransformation reactions. Data from the screening reactions are thought to represent false positives [e.g. i) outliers surrounding a low mean or ii) mixing artefacts; outlined in Chapter 5]. This idea is supported by the observation of low rates in the specific activity assays for the variant enzymes (Table 6.2). Due to the low levels of activity consistent experimental rates were difficult to obtain.

GC analysis of biotransformation reactions indicated the complete lack of conversion of substrate to product consistent with the specific activity data. Therefore, no mutants isolated from the H181X and H184X libraries showed significant activity towards these substrates. This is not unexpected since H181 and H184 have previously been demonstrated to be critical to substrate binding in OYEs [38, 69].

Substrate	Enzyme	Specific Activity (μmol min^{-1} mg^{-1})	Standard deviation (μmol min^{-1} mg^{-1})
Trans-cinnamic acid	WT	0.0047	0.00042
	H181I	0.0085	0.0049
	H181S	0.01	0.007
Cinnamonitrile	WT	0.016	0.0035
	H181S	0.020	0.0036
	H181L	0.016	0.0045
	H184K	0.022	0.0077
	H184Y	0.021	0.0073
1-Cyclohexene-1-carboxylic acid	WT	0.0074	0.0030
	H181Y	0.0060	0.0021
	H184F	0.0096	0.0043
Cyclohexene-1-carbonitrile	WT	0.0096	0.0030
	H181L	0.0055	0.0018
	H184L	0.0033	0.0007
	H184F	0.0055	0.0021
	H184K	0.0085	0.0023

Table 6.2. Specific activity of $PETNR_{His}$ variants assayed with a range of substrates differing in their activating group located next to the unsaturated C=C bond. Reactions were composed of 2 μM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM substrate in 50mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. WT represents the average and standard deviation of $PETNR_{His}$ control reactions.

It is important to note that there are two possible conclusions which explain the experimental data; i) there are no mutations at H181 and H184 which lead to improvements with respect to these substrates (i.e. the library is said to be empty) or ii) there are mutations at H181 and H184 which lead to improvements, but the conditions of the screening reaction are not suitable to identify them. In the case of this experiment the most likely conclusion is that the H181 and H184 libraries do contain improved variants but the screen was insufficient to identify them. Previously published data supports this conclusion. A number of mutants at position 181 and 184 have been demonstrated to yield improved catalysts with a

range of other substrates. These variants have been demonstrated to yield catalysts with improved stereochemical control [170]. Interestingly however the majority of these clones had compromised turnover [170] as has been observed in the initial library characterisation with CH (Section 4.5).

There are no known examples of PETNR reacting with alkenes containing a nitrile activating group, however a few rare examples of other OYEs reducing nitrile containing compounds are known [169]. This could be attributable to the poor activating properties of the nitrile group. In addition, a number of di-carboxylic acids are known substrates of PETNR [62], but there are no known examples of mono-acid substrates.

As described in Table 6.2 there were no improvements in the specific activity or GC from any of the variants isolated from screening reactions. Potential reasons for the lack of improved variants with nitrile and mono acid substrates include; i) the change in amino acid identity at these positions was not sufficient, at least in isolation, to improve activity against these substrates ii) the active site of PETNR may not be capable of catalysing reactions with these substrates due to the poor activating properties of the groups adjacent to the unsaturated C=C bond.

In addition to the reasons discussed above, a number of other factors may impact on the inability to detect improved mutants against these substrates including; i) poor expression leading to a reduced concentration of enzyme in the screening assay, as observed with the OYE1 N194 variant [69] ii) variation in FMN binding leading to a reduced subset of the purified PETNR being catalytically active iii) variation in the binding of NADPH as observed with the MR H186A and N189A variants [171].

Steps were included to hopefully compensate for some of these effects, such as the inclusion of FMN in the purification buffers to try and ensure complete flavination and the micro scale purification of variants ensuring protein concentration is limited by the binding capacity of the beads rather than the expression levels. Such steps are only capable of correcting some errors, for example additional FMN can occupy

an active site capable of accepting FMN, but not an active site unflavinated due to the mutation preventing binding.

Previous work has suggested that a large number of variants at residues H181 and H184 both maintain activity and in some cases provide improved biocatalysts [170] suggesting that these residues can be successfully varied. The studies reported in this section detected no variants with improved specific activity and no variants with improved biotransformation reactions. This could be explained either by no improved variants being present in the library or by the improved variants which are present in the library being undetectable using the current screening protocol. Due to the lack of success in detecting mutants in the H181X and H184X using the current assay these libraries were not screened against other substrate classes.

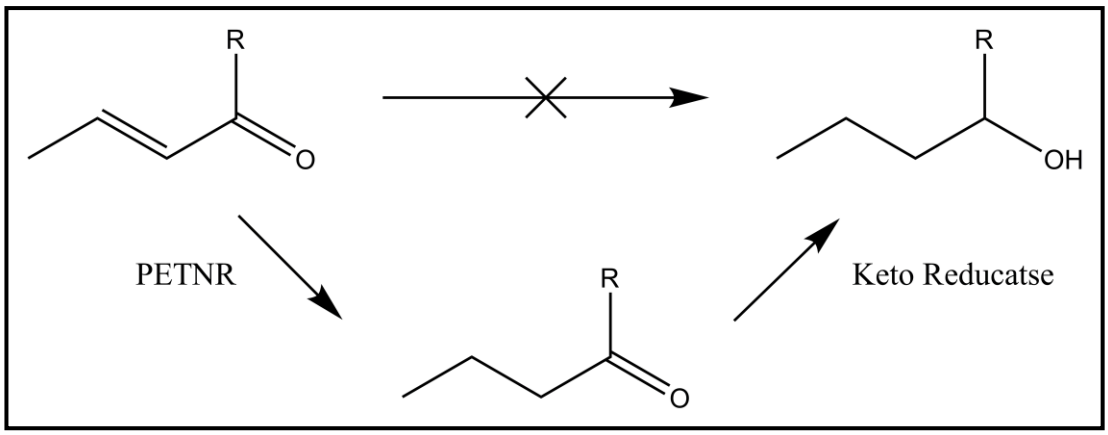
6.3 Characterisation of identified variants from screens performed with a range of methyl substituted substrates at both the α and β position of the alkene group.

6.3.1 *Trans*-cinnamaldehyde

WT control reactions demonstrated cinnamaldehyde to be an effective oxidative substrate of PETNR (Chapter 5). The activity of WT PETNR with cinnamaldehyde (2.8 ± 0.2 μmol min^{-1} mg^{-1}) is approximately half of the specific activity of model substrate CH (5.8 ± 0.7 μmol min^{-1} mg^{-1}). From these data it is clear that cinnamaldehyde is capable of binding PETNR in a conformation compatible with catalysis. However, the slight increases in activity in the variants isolated from the library screens suggest that there are more favourable binding geometries restricted in the wild type enzymes.

Specific activity data with *trans*-cinnamaldehyde as substrate (Table 6.2) indicate small increases in the rate of NADPH turnover for the T26S and Q241L enzymes (3 ± 0.2 and 3.3 ± 0.1 μmol min^{-1} mg^{-1} respectively). The remainder of the variants demonstrated compromised activity.

The analysis of GC results is complicated by the presence of an additional peak. This peak was identified as the alcohol product derived from the further reduction of the PETNR derived product aldehyde (Scheme 6.1), in this case 3-phenylpropan-1-ol. Mass spectrophotometric analysis of purified protein indicates the presence of a minor contaminant identified as YahK (an alcohol dehydrogenase like protein; results not shown). Similar side product formation has been detected previously in whole cell biotransformations with YqjM mutants [61]. This reaction is likely to proceed only from the reduced aldehyde (Scheme 6.1) [61] hence any alcohol present was most likely derived from the PETNR variant product and can be considered as a biotransformation product. In order to remove this impurity further purification steps could be incorporated.



Scheme 6.1. Generallised reaction scheme showing alkene reduction of PETNR followed by the action of an unidentified *E. coli* keto-reductase.

In the case of the W102L variant the enzyme appears kinetically compromised from the specific activity data. However this variant demonstrated superior biotransformation characteristics. This effect is likely due to the enzyme's inhibition characteristics at a range of substrate concentrations. In the case of $PETNR_{His}$ substrate inhibition is observed (Chapter 3). This substrate inhibition has been demonstrated previously to vary in response to mutation at a single amino acid position where the effect was practically abolished by the mutation of H184N [170]. It is possible that under some inhibition conditions relative reaction rates at 1 mM and 5 mM can vary.

There was a slight increase in both the specific activity and substrate conversion of T26S mutant with *trans*-cinnamaldehyde. However, the product yield was similar to wild-type enzyme (Table 6.3). The remaining 13 % of product(s) was not due to the formation of the alcohol by-product. Possible explanations for the unaccounted product(s) include i) substrate/product degradation and/or volatility over extended reaction times, ii) the action of enzymatic contaminants on substrate/product and iii) the conversion of substrate/product to a novel compound not detected by our analytical techniques. A similar effect is seen in the Y68S reaction. Like the T26S mutant, both the specific activity and substrate conversion was moderately higher for the Q241L mutant.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)	Alcohol product yield (%)
WT	2.83	0.21	72	72	*nd*
T26S	3.01	0.15	82	69	*nd*
Y68S	1.39	0.03	92	68	*nd*
W102L	0.94	0.02	88	87	*nd*
Q241L	3.29	0.11	100	57	43
Y351F	0.35	0.03	100	32	*nd*

Table 6.3. Specific activity and GC analysis of reactions with $PETNR_{His}$ variants identified from screens with *trans*-cinnamaldehyde as substrate. Specific activity reactions were composed of 1 - 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM *trans*-cinnamaldehyde in 50 mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC reactions were composed of 5 mM *trans*-cinnamaldehyde, 6 mM NADPH and 2 µM enzyme in 50mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml. *nd* signifies that no product was detected.

Position 241 was randomised due to its location, on a flexible loop close to the substrate in the active site, inferred from crystal structure data for PETNR [40] and other OYEs [15, 41]. Q241L represents only a minor change in size, but the mutation does represent the conversion from a hydrophilic to a hydrophobic side chain. The effect on activity may be mediated by movement in the flexible loop to accommodate the slightly bulkier leucine residue. This movement could have indirect effects on neighbouring residues. Any effect present has been demonstrated, from the specific activity data, to be minor and without further studies determination of the molecular mechanism underlying the change is impossible.

6.3.2 α-Methyl-*trans*-cinnamaldehyde

α-Methyl-*trans*-cinnamaldehyde has been demonstrated to be a quite poor substrate for PETNR (0.6 ± 0.001 μmol $min^{-1}mg^{-1}$). Analogy with the previously studied substrate cinnamaldehyde (2.8 ± 0.2 μmol min^{-1} mg^{-1}) indicates the methyl substituent located on the α carbon is responsible for the compromised activity with this substrate. The most likely reason for this is the methyl substituent sterically interfering with binding in an optimal conformation for catalysis, which are accessible in the case of *trans*-cinnamaldehyde.

The specific activity of a range of enzyme variants isolated from screening with α-methyl-*trans*-cinnamaldehyde (Table 6.3) suggests that a number of enzymes have shown a substantial improvement activity with α-methyl-*trans*-cinnamaldehyde as substrate. Mutants W102F and W102Q were shown to have dramatically improved specific activity when compared to $PETNR_{WT}$ (5.9 ± 0.2 and 1.9 ± 0.06 μmol min^{-1} mg^{-1}, respectively). Mutant Q241L also demonstrated a modest improvement (0.8 ± 0.02 μmol min^{-1} mg^{-1}), however the other variants screened demonstrate reduced specific activity.

Biotransformation data with α-methyl-*trans*-cinnamaldehyde (Table 6.4) support the observations derived from specific activity data, which indicated that WT, W102F, W102Q, and Q241L all demonstrated complete depletion of substrate with subsequent detection of the desired aldehyde product (2-methyl-3-phenylpropanal) or the alcohol side-product (2-methyl-3-phenylpropan-1-ol) discussed previously (Chapter 6.3.1). The magnitude of the activity increase is not quantified by GC due to all the active samples reacting to completion under the assay conditions.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)	Alcohol product yield (%)	e.e. (%)
WT	0.65	0.001	99	99	1	13 (*S*)
T26A	0.21	0.008	10	0	6	0 (*S*)
Y68S	0.10	0.004	81	77	8	6 (*R*)
W102F	5.93	0.22	100	97	3	62 (*R*)
W102Q	1.90	0.068	100	1	99	ND
Q241L	0.82	0.022	99	77	22	5 (*S*)
Y351F	0.18	0.002	99	98	1	38 (*S*)

Table 6.4. Specific activity and GC analysis of biotransformation reactions with $PETNR_{His}$ variants identified from screens with α-methyl-*trans*-cinnamaldehyde as substrate. Specific activity Reactions were composed of 0.1 - 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM α-methyl-*trans*-cinnamaldehyde in 50 mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC Reactions were composed of 5 mM α-methyl-*trans*-cinnamaldehyde, 6 mM NADPH and 2 µM enzyme in 50mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml for 4 hours.

The increased specific activity of W102F, W102Q and Q241L mutants towards α-methyl-*trans*-cinnamaldehyde suggests the substrate binds to the enzyme in a more optimal conformation for catalysis. W102 has already been demonstrated to move in response to substrate binding [75], has been demonstrated to have an effect in discrimination between oxidative substrate mechanisms [75] and the W102Y/F mutants have been demonstrated to relieve a steric clash with substrate (Section 1.4.3.3; [75]). The specific mutations above represent the removal of a bulky intermediate side-chain, and its replacement by either a slightly smaller polar side-chain (W102Q) or a slightly smaller hydrophobic side-chain (W102F). This suggests that size is the important factor.

Residue 241 was selected for mutagenesis studies due to the possibility of opening up the active site on the basis of the crystal structure (Section 1.4.3.7). The Q241L substitution involves the substitution of a long polar side-chain with the shorter non-polar side-chain, the mechanism of this marginal improvement is unclear.

In addition, the comparison of $PETNR_{His}$ with W102F highlights a reversal in stereochemistry. This effect implies that the angles of hydride transfer from the FMN prosthetic group (N_5) to the substrate (C_β) are changed. Such stereochemical variation is explained by the existence of multiple substrate binding modes each with a characteristic angle of hydride transfer resulting in the various resultant stereochemistries. These multiple binding modes are evident in any reaction which does not yield an enantiopure product. The suggestion in this case is that the distribution of binding mode in the W102F variant is different to that of $PETNR_{His}$.

The variation of activity level and stereochemistry in response to mutation of W102 has been demonstrated previously in other OYEs [61, 148]. In the case of OYE1 screening of W116X libraries with α and β substituted cyclohexenone identified a range of mutants at this position including the W116F mutant (analogous to the PETNR W102F mutant) [148]. In contrast with the data described above, all W116X variants were demonstrated to have inferior specific activates with the β substituted substrate [148], whereas both the W102F and Q variants in our study were shown to have enhanced kinetic properties (albeit with an α substituted substrate in this case). The W116X variants were trialled with α substituted carvone as substrate and W116I demonstrated a reversal of stereochemistry at the α carbon [148] similar to that described in this study for the W102F variant. Additional data suggest that the addition to the unsaturated bond still occurs as a *trans* addition mechanism. However the reversal in stereochemistry is attributed to binding in an alternate 'flipped' geometry [148]. Interestingly, in the case of carvone, the C_4 substituent was identified as being critical for this binding geometry [148], however, this substituent is replaced by a phenyl substituent in the β position in our substrate. The screening of additional substrates also demonstrated the reversal of stereochemistry associated with an improvement in catalytic efficacy with the aliphatic substrate neral by W102F [148] which is directly analogous to the reaction we have demonstrated. Interestingly, the variant equivalent to W102Q was not observed under the screening conditions of this experiment [148].

The respective residue in YqjM is A104 [27] and was studied by the generation of ISM libraries [61]. The exchange of a bulky tryptophan for the minimal alanine in

thermophillic-like OYEs does suggest that the residue may not perform the same function. A number of variants at this position were demonstrated to improve catalysis including an approximate 15 fold increase in the percent conversion associated with the A104W variant [61].

6.3.3 Methyl-*trans*-cinnamate

The data from the specific activity determination of a range of variants with methyl-*trans*-cinnamate (Figure 6.5) suggest a minimal improvement in activity with Y351H. However, as the specific activity is very low for all enzymes, these small variations in activity are within the experimental error and therefore may not be representative of a genuine improvement in activity with methyl-*trans*-cinnamate (Figure 6.5).

Analysis of the biotransformation demonstrated the complete absence of substrate conversion or product yield with all variants including $PETNR_{His}$ WT (data not shown). This indicates that the variants isolated from screening are all false positives, and the variation is due to error in conjunction to a low signal: noise ratio.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})
WT	0.013	0.002
T26H	ND	ND
Y68M	ND	ND
W102V	0.013	0.006
Q241H	0.005	0.0006
Y351H	0.016	0.002

Table 6.5. Specific activity of $PETNR_{His}$ variants assayed with methyl-*trans*-cinnamate as substrate. Reactions were composed of 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM methyl-*trans*-cinnamate in 50 mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. WT represents the average and standard deviation of $PETNR_{His}$ control reactions.

The lack of activity of PETNR variants with methyl-*trans*-cinnamate is not completely unexpected since mono-esters have been demonstrated to be substrates for OYEs rarely [2, 28, 59]. This observation suggests that either; i) there are no improved mutants in the libraries screened, representing a range of conditions as discussed previously in section 6.2; or ii) that any improved variants present in the library are not detected under the conditions by which the assay is limited (false negatives). This again represents a large subset of conditions which are detailed in Section 6.2.

6.3.4 3-Methyl-2-cyclohexen-1-one

The compound 3-methyl-2-cyclohexen-1-one has been demonstrated to be a poor substrate for PETNR (0.009 ± 0.0002 µmol min^{-1} mg^{-1}). This contrast is particularly stark when compared with CH (5.8 ± 0.7 µmol min^{-1} mg^{-1}) which is often used as a model substrate for OYEs. The addition of a methyl substituent on the β carbon results in an almost total loss of activity. This difference in activity is thought to be due to steric hindrance between the substrate methyl substituent and residue T26 CG2 atom, preventing access to catalytically competent conformations, as has been described previously for the inhibitor progesterone [73].

The specific activity with 3-methyl-2-cyclohexen-1-one as substrate (Table 6.5) suggests that T26A has acquired a substantial increase in activity (0.024 ± 0.0008 µmol min^{-1} mg^{-1}) and Q241W has improved to a far smaller extent (0.015 ± 0.002 µmol min^{-1} mg^{-1}). The remainder of the clones demonstrated no improvement in activity.

Biotransformation data with 3-methyl-2-cyclohexen-1-one as substrate (Table 6.6) demonstrated that only the T26A variant had a significant improvement in substrate conversion (%) and product yield (%) with no by-products detected. The remainder of the mutants demonstrated compromised biotransformation, including the Q241W variant which demonstrated improved specific activity. Such inconsistencies are probably attributable to either i) the mutation having an effect on protein stability, so that biotransformations carried over extended periods of time showed a reduction in overall product formation or ii) the difference in conditions in each experiment, particularly the increased concentration of oxidative substrate (1 mM in specific activity assays and 5 mM in biotransformation reactions). Previous experiments with $PETNR_{WT}$ have indicated the presence of substrate inhibition at elevated CH concentrations (Section 3.2.2). The K_i associated with this phenomenon has been demonstrated to vary in response to mutations in PETNR [170].

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)
WT	0.010	0.0002	21	21
T26A	0.024	0.0008	41	41
Y68L	0.009	0.0018	8	8
W102D	0.005	0.0008	0	0
W102I	0.007	0.0005	0	0
Q241W	0.015	0.0024	8	8
Y351E	0.006	0.0007	8	8

Table 6.6. Specific activity and GC analysis of biotransformation reactions with $PETNR_{His}$ variants identified from screens with 3-methyl-2-cyclohexen-1-one as substrate. Specific activity reactions were composed of 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM 3-methyl-2-cyclohexen-1-one in 50mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC Reactions were composed of 5 mM 3-methyl-2-cyclohexen-1-one, 6 mM NADPH and 2 µM enzyme in 50mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml.

Residue 26 has previously been demonstrated to play a role in preventing the inhibitor, progesterone, binding in a catalytically competent conformation (Section 1.4.3.1.; [73]). These interactions are possibly responsible for the decrease in activity associated with the methyl substitution at the C_β of CH. It is likely that the mutation of T26 to the smaller alanine residue removes a steric clash associated with either the methyl or hydroxyl group of the T26 residue. This reduction in size of residue 26 increases the size of the substrate binding pocket allowing the accommodation of the substrate in a conformation more conducive with catalysis. Residue 26 is conserved within the thermophilic-like OYEs subclass as a cysteine residue (C26). A range of variants at this position have been screened and demonstrated to improve activity and to affect stereochemistry with a wide range of β substituted substrates [61].

6.3.5 2-Methyl-2-cyclopenten-1-one

2-Methyl-2-cyclopenten-1-one (0.17 ± 0.02 µmol min^{-1} mg^{-1}) has been demonstrated to be a poor substrate for PETNR, particularly when compared with 2-cyclopentenone (data not shown). The addition of a methyl substituent on the α carbon results in a substantial decrease in activity. This difference in activity is thought to be due to steric hindrance between the substrate methyl group and a number of active site residues identified from the crystal structure, including T26, Y68, W102, Q241 or Y351, preventing access to catalytically superior conformations.

The specific activity of a number of variants with 2-methyl-2-cyclopenten-1-one as substrate (Table 6.7) suggests a wide range of enzymes appear to have improved activities. The specific activity data demonstrate improvement for almost all the enzymes tested. Variants including W102F and W102I have a substantial increase in activity (0.83 ± 0.02 and 1.00 ± 0.04 µmol min^{-1} mg^{-1} respectively), the best representing an almost 6 fold improvement. Additionally a number of variants including T26A, Y68F, Q241W, Q241V and Y351P have specific activities improved but to a far lesser extent.

Biotransformation reactions with 2-methyl-2-cyclopenten-1-one as substrate (Table 6.7) demonstrated that W102F and W102I had marginally improved substrate conversions (%) with slightly poorer product yields (%) than wild-type. Mutant Y351P also has improved substrate conversion and product yield however the magnitude of the specific activity change and the increase in conversion is small.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)	e.e. (%)
WT	0.17	0.02	89	89	69 (*S*)
T26A	0.30	0.02	59	59	56 (*S*)
Y68F	0.43	0.01	59	53	49 (*S*)
W102F	0.83	0.02	93	81	46 (*S*)
W102I	1.00	0.04	93	82	72 (*S*)
Q241W	0.31	0.06	89	89	66 (*S*)
Q241V	0.31	0.01	72	72	62 (*S*)
Y351M	0.16	0.01	84	84	66 (*S*)
Y351P	0.23	0.01	94	94	70 (*S*)

Table 6.7. Specific activity and GC analysis of biotransformation reactions with $PETNR_{His}$ variants identified from screens with 2-methyl-2-cyclopenten-1-one as substrate. Specific activity reactions were composed of 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM 2-methyl-2-cyclopenten-1-one in 50 mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC reactions were composed of 5 mM 2-methyl-2-cyclopenten-1-one, 6 mM NADPH and 2 µM enzyme in 50 mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml.

None of the observed variants have an inversion in stereochemistry as observed with the substrate α-methyl-*trans*-cinnamaldehyde (Section 6.3.2). However, the W102X library screening provides an interesting case study as mutants W102I and W102F produce higher and lower e.e. Considering the rationale described previously, that changes in stereochemistry were attributable to substrate binding to the active site in various conformations, allowing the *trans* addition across the double bond to give alternate results. This suggests that W102F and W102I favour a different distribution of binding modes.

These variants whilst not providing an inversion in stereochemistry potentially represent an intermediate from which to engineer such an enzyme. The W102F variant has also been demonstrated to lead to a complete reversal in stereochemistry with α-methyl-*trans*-cinnamaldehyde as substrate and a decrease (partial reversal) in ee with 2-methyl-2-cyclopenten-1-one which is thought to be

the first step in evolving alternate stereochemistry. The less pronounced inversion with 2-methyl-2-cyclopenten-1-one may be due to the reduced size of this substrate associated with the lack of an aromatic substituent at the β carbon.

The molecular basis of this improvement is most likely that mutagenesis of W102 to smaller side chains removes steric clashes upon substrate binding. W102 has previously been demonstrated to move to perform a range of functions in substrate binding including; i) discrimination between oxidative mechanisms, by removal of a steric clash with TNT [75] and ii) movement to accommodate substrate binding. These effects are discussed in detail in Section 6.3.2.

6.3.6 3-Methyl-2-cyclopenten-1-one

3-Methyl-2-cyclopenten-1-one (0.010 ± 0.001 µmol min^{-1} mg^{-1}) has been demonstrated to be an exceptionally poor substrate, particularly when compared with cyclopentanone (data not shown) and to 2-Methyl-2-cyclopenten-1-one (0.17 ± 0.02 µmol min^{-1} mg^{-1}). This loss of activity is thought to be attributable to the same steric clash involving the methyl substituent on the C_β and residue T26 discussed previously in reference to 3-Methyl-2-cyclohexen-1-one (Section 6.3.4) Notably a comparison between 2-cyclohexen-1-one, 3-methyl-2-cyclohexen-1-one and 2-methyl-2-cyclopenten-1-one suggests that substituents at the β carbon appear to have a more substantial effect on the specific activity than those at the α carbon.

The specific activity of a number of variants with 3-methyl-2-cyclopenten-1-one (Table 6.8) suggest a number of enzymes with slightly improved activity towards 3-methyl-2-cyclopenten-1-one, such as T26L and Y351M (0.015 ± 0.003 µmol min^{-1} mg^{-1} and 0.019 ± 0.006 µmol min^{-1} mg^{-1} respectively). However the remainder of the mutants showed no improvement in specific activity. Biotransformation reactions with 3-methyl-2-cyclopenten-1-one as substrate (Table 6.8) demonstrated that the only mutant with improved activity was Y351M, showing a modest increase in both yield and conversion.

Enzyme	Specific Activity (µmol min^{-1} mg^{-1})	Standard deviation (µmol min^{-1} mg^{-1})	Conversion (%)	Yield (%)
WT	0.010	0.001	9	9
T26L	0.016	0.004	1	0
W102Y	0.012	0.005	1	1
Q241D	0.007	0.001	1	1
Q241V	0.005	0.002	4	4
Y351M	0.020	0.006	11	11

Table 6.8. Specific activity and GC analysis of biotransformation reactions with $PETNR_{His}$ variants identified from screens with 3-methyl-2-cyclopenten-1-one as substrate. Specific activity reactions were composed of 2 µM $PETNR_{His}$ variant, 0.1 mM NADPH, 1 mM 3-methyl-2-cyclopenten-1-one in 50 mM KH_2PO_4/K_2HPO_4 buffer, pH 7 containing 5 % ethanol in a total volume of 1 ml. GC reactions were composed of 5 mM 3-methyl-2-cyclopenten-1-one, 6 mM NADPH and 2 µM enzyme in 50 mM KH_2PO_4/K_2HPO_4 buffer pH 7 containing 5 % ethanol in a total volume of 1 ml.

The molecular basis of the slight improvement of Y351M in both the specific activity and the biotransformation reaction is probably due to the removal of a steric clash between the tyrosine side chain and the β methyl substituent. In the Y351M variant the bulky, polar tyrosine is replaced by the smaller and hydrophobic methionine residue Y351 has been demonstrated to have a role in the binding of PETNR to some compounds for example in the binding of the binding of substrate analogue *p*-hydroxybenzaldehyde in the active site of OYE1 [15]. This effect is unlikely to be the case for 3-methyl-2-cyclopenten-1-one due to lack of the appropriate functionality *para* to the carbonyl functionality.

6.4 Discussion

The screening of mutant libraries against a range of substrates containing either a ketone or an aldehyde-activating group (Table 6.9) yielded a total of 35 potential hits. Kinetic characterisation of these variants confirmed improvements in the specific activity of 14 enzymes (2 enzymes were not fully characterised). This suggests a true positive rate of 40 % for our sample. Of those variants with confirmed specific activity increases only 8 demonstrated improvements in substrate conversion under conditions of biotransformation reactions and represent potentially improved synthetic enzymes (Table 6.9).

Substrate	Enzyme	Improvement in specific activity (fold)
Trans-cinnamaldehyde	T26S	1.1
α-methyl-*trans*-cinnamaldehyde	W102F	9.1
3-Methyl-2-cyclohexen-1-one	T26A	2.5
2-Methyl-2-cyclopenten-1-one	Y68F	2.5
	W102F	4.9
	W102I	5.9
	Y351P	1.4
3-Methyl-2-cyclopenten-1-one	Y351M	1.9

Table 6.9. Summary table detailing confirmed improved variants isolated from the screening of PETNR mutagenic libraries.

The summary data (Table 6.9) suggests a general relationship between the positioning of substituents and the probability of a significant increase in specific activity. In general α-substituted compounds e.g. α methylcinnamaldehyde, 2-methyl-2-cyclopenten-1-one are much better substrates for WT $PETNR_{His}$ than the β substituted compounds 3-methyl-2-cyclohexen-1-one and 3-methyl-2-cyclopenten-1-one (specific activity data described in this chapter). These α-substituted substrates have also been shown to yield far greater increases in specific activity, suggesting acceptance of novel substituents at this position is easier to achieve. This suggests that the protein environment surrounding the β carbon is far more restrictive which is most likely attributable to the T26 residue as discussed

previously. This implies that the improvement of acceptance at the β carbon requires the randomisation of further single site positions, iterative rounds of mutagenesis or the randomisation of multiple sites at the same time.

These results suggest possible mutation hotspots associated with the methyl substituents at the α and β positions (Figure 6.1). Variants isolated from screening with α and β substituted unsaturated carbonyls have highlighted variation in the tyrosine pair (Y68/Y351) which has been suggested to regulate access to the active site (Section 1.4.3.2, Section 1.4.3.6 [40]). These variants involve the substitution of the large, polar tyrosine side chain with a variety of residues including phenylalanine, proline and methionine. The largest improvement in specific activity comes from the substitution of the polar hydroxyl at position 68 with a similar sized non-polar side-chain (Y68F). Smaller improvements are observed by replacement of Y351 with the smaller, non-polar proline and methionine residues. Due to the variation in size of the substituted residues this suggests that size may not be the determining factor. All substitutions change a polar residue for a non-polar residue which suggests that electrostatic interactions between the substrate and the polar tyrosine side-chain could influence substrate binding.

The variant with the largest increase in specific activity for both the 2-methyl-substituted substrates is the substitution of W102 to the smaller non-polar phenylalanine and isoleucine residues. These transformations involve the substitution of a large non polar side-chain with smaller aromatic and aliphatic non-polar side-chains, which suggests that steric constraints between the W102 side-chain and the α substituent are likely to dominate this interaction.

The variant with the largest increase in specific activity for both the 3-methyl-substituted substrates is the substitution of T26 to the smaller non polar residue alanine. This transformation involves the removal of a medium sized polar side-chain and replaced by a small non-polar side chain, whilst the change in polarity is possibly an influencing factor, steric constraints between the β substituent and the additional methyl or the hydroxyl group on the T residue. Interestingly mutations at this residue were also observed to increase activity slightly with *trans*-

cinnamaldehyde, suggesting a similar interaction could be present in non-substituted substrates.

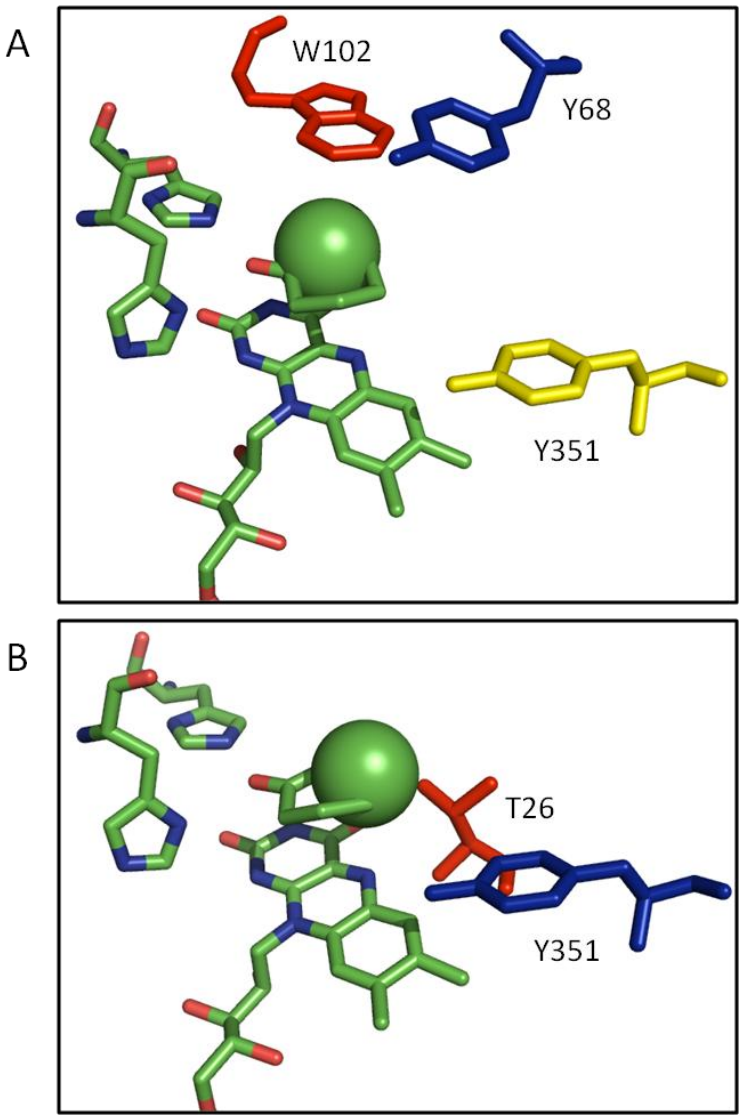


Fig 6.1. Location of residues identified from hits in libraries screened containing with methyl substituted substrates. Panel A α substituted substrates including α-methyl-*trans*-cinnamaldehyde and 2-methyl-2-cyclopenten-1-one, red, W102; blue, Y68; yellow, Y351; panel B β substituted substrates including 3-methyl-2-cyclopenten-1-one and 3-Methyl-2-cyclohexen-1-one, red, T26; blue, Y351. Other active site residues are shown as atom coloured sticks with the substituted carbon shown as a sphere.

Despite the successes demonstrated for screening PETNR libraries against substrates with novel α/β substituents, the experiments to vary the activating group were much less successful. These substrates were screened only against the H181X and H184X libraries due to their observed role in binding the activating

group (Section 1.4.3.4), it was thought that modification of these residues could lead to binding of other activating groups. These residues are largely conserved as His/Asn or His/His pairs in other OYEs and homologues. This outcome is not wholly unexpected due to the conserved nature of the H181 and H184 residues, and the potential roles discussed in other aspects of the PETNR catalytic cycle, such as NAD(P)H binding (Section 1.4.1.3; [69]) and FMN binding (Section 1.4.2; [73]). Possible reasons for this negative outcome include; (i) that there are no improvements present in this library or (ii) the possibility that there are variants which improve the oxidative half reaction however have such compromised reductive half reactions that these effects are masked.

7 Discussion

7.1 Generation of tagged PETNR and initial robotic trials (Chapter 3)

Chapter 3 describes the generation and characterisation of tagged PETNR variants. The best variant was incorporated into a protocol for the growth, extraction, purification and assay of enzyme libraries. Following finalisation of the protocols initial trials on a $PETNR_{WT}$ control block were performed to assess the reproducibility of the experimental protocol.

Variants of PETNR incorporating either histidine or biotin tag nucleotide sequences were generated using a range of molecular biology methods. Initial expression and purification studies showed that the histidine tagged PETNR ($PETNR_{His}$) was expressed to a higher level and easier to purify than the biotin tagged variant ($PETNR_{Bio}$). Due to the problematic expression and purification conditions associated with the $PETNR_{Bio}$ variant, only the $PETNR_{His}$ variant was taken forward to thorough characterisation and multi-well trials. Various aspects of the $PETNR_{His}$ robotic protocols were optimised such as the growth, purification and assay conditions in a 96 well format. Finalised procedures were trialled using a control block containing 96 $PETNR_{His}$ WT clones. This initial trial enabled the quantitation of the sources of error in the protocols. These were estimated to be 14 % and 13 % for errors associated with the growth and purification of the clones, respectively. A proportion of this error has been identified as systematic with a variation of up to 22 % across the block.

A number of potential sources of error associated with the growth, purification and assay in 96 well format were identified. Firstly, the accuracy of robotic colony picking varied, which was dependent on the evenness of the distribution and size of the individual colonies as well as the presence of imperfections in the agar plate (e.g. agar bubbles or particulates in the agar which may be mistaken for colonies). Secondly, a variation in the growth rate of individual clones was dependent on the inoculum size from glycerol stocks combined with the effect the mutation has on *E. coli* growth. Thirdly, mutants can often vary in the degree of protein expression, an effect that was minimised by only performing micro-scale protein purification.

Protein purification conditions are likely to contribute significantly to the overall experimental error including; i) the potential loss of all or part of the cell pellets during removal of growth media ii) variation in the amount of magnetic beads dispensed iii) variation in the volume of individual buffers dispensed and iv) variation in the volume of final eluate. The final potential source of error is the assay conditions including factors such as i) pipetting errors ii) timing of reaction initiation and iii) error in absorbance readings. Whilst a subset of these sources of error is apparent in $PETNR_{WT}$ experiments, there are some sources of error, such as the variation of expression due to mutations, which cannot be characterised. Consequently, the total error associated from the screening reactions cannot be estimated.

7.2 Library generation and initial characterisation (Chapter 4)

Chapter 4 outlines the generation of single site saturation mutagenic libraries at a number of positions in the active site of $PETNR_{His}$. This involves the generation of library DNA by mutagenic PCR, the sequencing of a number of colonies to ensure randomisation, the trial high throughput expression and purification of libraries in a 96 well format, followed by 96 well format activity assays with standard wild type substrates.

When generating genetic diversity there are a number of sources of bias which would increase the proportion of some variants over others, such as the use of degenerate oligos in PCR. Oligos with a greater degree of complementarity to the template are more likely to anneal and hence mutations changing fewer bases will be slightly overrepresented in libraries. An example of this can be seen in the apparent over representation of WT codons (seen in the sequencing results) which result from the completely complementary oligonucleotide binding.

Another source of bias is cellular toxicity, where some mutations generated may be toxic to the cell and hence either do not grow or lead to small colonies that are not picked by the colony picker. This acts as a type of unintended selection pressure as the colony picking software preferentially selects colonies deemed closest to ideal standards of colony size, shape and separation from its neighbours.

It is also important to note that sequencing 10 colonies per 96-well deep well block represents just over 10 % of the library members, and is therefore not a true reflection of the overall randomness of the site-saturation mutagenesis. In the case of the H181X and H184X libraries further sequencing, by co-workers, revealed a highly randomised library with the majority of library clones represented [170].

In addition to DNA level bias attributable to the individual variants, there are a number of factors, involved in the growth, purification and assay which affect the activity demonstrated by each well. Other effects, separate from the property of interest, which are likely to manifest themselves in screening reactions include; i) effect of the mutation on expression; ii) effect of the mutation on stability/folding, leading to increased cellular degradation; iii) effect of the mutation on flavination/FMN binding; iv) effect of the mutation on NADPH activity; v) effect of the mutation on substrate inhibition properties. All of which could possibly serve to mask the true property of interest, that is variation in the steady state reaction rate with the oxidative substrate.

There are a vast number of potential sources of error attributable to the generation of mutagenic libraries. A number of these sources of error come under the control of the experimenter, to some degree, however the majority do not. From the data in this chapter it is clear that reasonable steps have been taken to ensure the quality of the libraries about to be investigated and all but one of these libraries have been demonstrated to be of high quality. Control data were then recorded for these libraries, and they demonstrated a wide range of residual activity with NADPH and CH.

Single site saturation libraries generated in the future could potentially be improved by ensuring all possible variants are represented in the library. The experiments in this chapter indicate that all potential variants at each position have probably not been sampled in all cases. Most obvious of these is the non-random construction of libraries using oligos designed to incorporate specific amino acids at each position.. Alternate methods including the use of MAX libraries [93] and codon mutagenesis

[172] could also yield improved quality libraries. These methods could enable the construction of extremely high quality libraries.

7.3 Summary of screening PETNR libraries against a panel of novel substrates (Chapter 5)

Chapter 5 provides a summary of the screening reactions conducted in an attempt to isolate improved clones from the generated libraries. A number of types of problematic data are detailed in this chapter.

Firstly, there is the occurrence of false negatives. These data represent enzyme variants which have an improved property of interest but appear negative according to the screen employed. Due to their nature these errors are difficult to quantify since they appear negative and are not highlighted by the screen. At a molecular level these results are explained either by; i) positive mutations which are not represented in the library due to bias or probability (as discussed in Section 7.2); ii) positive variants which are not identified due to experimental error (as discussed in Section 7.1); iii) positive variants with improved oxidative half reactions but compromised in other aspects of the catalytic cycle (as discussed in Section 7.2); viii) variants which are not carried further due to the hit identification criteria.

It is usually accepted in screening experiments that a lot of positive variants are missed and present as false negatives. Due to the nature of these data they are difficult to quantify. Experimental conditions are, where possible, tailored to avoid these sources of error, for example, the micro scale purification to avoid variation due to expression levels. Where this is not possible steps are taken to assess their contribution, for example sequencing reactions to assess variability attributable to DNA bias.

Another problematic type of data identified is the occurrence of false positives. These data represent variants which do not have an improved property of interest but appear positive according to the screen and selection criteria employed. As opposed to false negatives discussed previously, false positives can be identified and quantified but this identification generally requires further unproductive work.

At a molecular level these results are explained either by; i) the observation of experimental outliers, for example even in the case of an experiment with no improved variants experimental variation means there is always a top ranked hit; ii) the effect of mixing errors on an assay with a low signal : noise. These type of data are indistinguishable from the true positive data and characterised further using kinetic and analytical chemistry techniques (Chapter 6).

7.4 Confirmation of suspected hits (Chapter 6)

Chapter 6 discussed the further characterisation of potential improved variants. Intermediate scale purifications were used for the spectrophotometric specific activity determination and biotransformation reactions were analysed to determine the percent substrate conversion and the percent product yield.

It was clear from the data that a number of positive variants isolated from the screen were not improved biocatalysts in these subsequent analyses. These false positive results can be explained by consideration of the reasons discussed in Section 6.3, and are particularly prevalent in reaction with poor PETNR substrates, suggesting that low signal : noise can be a problem in some cases.

In order to minimise this problem the Z-factor method of non controls based normalisation was used, but in the circumstance of low signal significant variation due to experimental error can lead to artificially inflated values. However, the alternatives, such as arbitrarily setting a cut-off value, also have their problems. All normalisation and hit definition methods suffer from their own characteristic fallibilities and any single method is unlikely to be completely effective.

A number of variants were identified with improved activity against a range of α and β substituted enones identified as poor substrates for $PETNR_{WT}$, including the 9-fold improvement of the W102F variant with α-methyl-*trans*-cinnamaldehyde. This effectively increases the activity of the enzyme with a poor substrate to a value comparable with the model substrate CH. In addition, mutants W102I and W102F showed a 6 and 5 fold improvement, respectively, with substrate 2-methyl-2-cyclopenten-1-one. A number of other modest improvements; such as T26S with

trans-cinnamaldehyde, Y68F and Y351P with 2-methyl-2-cyclopenten-1-one, and Y351M with 3-methyl-2-cyclopenten-1-one and T26A with 3-methyl-2-cyclohexen-1-one were also found.

These changes in catalytic rate were changes in the property used as a reporter in the screen. However, a number of interesting coincidental changes in the stereochemistry of the reaction were also demonstrated. These include a reversal in stereochemistry associated with the reaction of α-methyl-*trans*-cinnamaldehyde with the W102F enzyme and a diversion in stereochemistry reported for the reactions of 2-methyl-2-cyclopenten-1-one with W102F/I variants.

A number of experiments described above (Chapter 6) demonstrated no improved variants for a particular substrate (e.g. Section 6.2). This lack of positive results is attributable to a number of possible circumstances including; i) the residues targeted could be the wrong residues to affect the acceptance of the substrate; ii) improved variants could be present but show as false negatives for any of the reasons discussed previously (Section 7.1.3); iii) randomisation at a single position may not be insufficient in isolation to affect the acceptance of the substrate, and further iterative or synergistic substitutions may be required.

7.5 Future work

In order to further characterise the effects of these amino acid changes on catalysis, further detailed kinetic and structural studies could be carried out of the variants generated from these experiments to investigate the molecular nature of interactions leading to changes in activity. These studies could provide further insight into the substrate-binding conformations of PETNR, whilst also providing structural information to aid further directed evolution studies.

The obvious extension to the studies reported in this thesis is to continue by generating libraries at further positions around the active site. This allows the extension of the sequence space to be screened. In some cases rationally designing novel variants by combining multiple hits from different libraries could also result in improved variants. Additionally, the variation of multiple positions in the same

library, such as the CASTing and multiple mutagenesis reactions described previously, could allow access to enzyme variants in which some degree of synergy is required for improved function. Randomising the residues identified in this study (such as Y68, W102 and Y351 for α substituted substrates, and T26 and Y351 for β substituted substrates) would represent an excellent starting point for these studies.

These novel libraries, along with those described in this study, could be screened against other potential substrates. The activity of the mutants generated in this report could also be tested against further analogous substrates, for example the W102F mutant could be trialled against a range of α substituted *trans*-cinnamaldehyde substrates. The improved variants generated in these studies could be subjected to further rounds of iterative site directed mutagenesis. This approach might be particularly useful where a number of amino acids have been demonstrated to yield improved variants from the first round of mutagenesis, for example screening using Y68F as a template for randomising W102. All the experiments discussed above could be performed rapidly using the same experimental protocols and equipment as outlined in Chapters 3 and 4 and demonstrated to be effective in Chapters 5 and 6. Works in these directions are essentially inexhaustible and could yield a large number of novel variants with desirable properties.

Further mutagenic libraries could be generated and screened using alternate library generation techniques. These include gene shuffling approaches, multiple mutagenesis reactions or a whole range of the techniques discussed in Section 1.5.2.

There is also potential to vary the screening method to use different properties of the chemical reaction to report on the reaction progress, for example measuring substrate conversion (%) or product yield (%). In particular the incorporation of a chiral analytical chromatography method, similar to that described previously [61], would enable the screening for stereospecific enzymes and potentially lead to an

incredibly powerful platform for the identification of industrially relevant biocatalysts.

Oligonucleotide table

Oligonucleotide identification		Oligonucleotide sequence	T_m (˚C)	2˚ T_m (˚C)
His3T	F	CATCACCATTAATCCCGCTTTGTACATTGATAGCG	56	44
	R	GTGATGATGCAGTGAAGGGTAGTCGGTATAAGGTTC	60	42
His6T	F	CATCATCACCATCACCATTAATCCCGCTTTGTACATTG	59	8
	R	ATGGTGATGGTGATGATGCAGTGAAGGGTAGTCGG	61	6
Bio3	F	CAGAAAATTGAATGGCATGAATAATCCC	56	17
	R	CGCTTCAAAAATATCGTTCAGGCCCTGG	61	30
Bio5	F	AGGGCCTGAACGATATTTTTGAAGCG	58	21
	R	GATTARRCATGCCATTCAATTTTTTG	52	11
Biorep	F	GAAGCGCAGAAAATTGAATGGCATGAATAATCCCGC	78	26
	R	GCGGGATTATTCATGCCATTCAATTTTCTGCGCTTC		36
TagFind	F	CGGGTGCGTATACGGCAGAAAAAGC	61	38.4
	R	GCGCATAATTTCCTCATTAACCAGTCGAATG	60	26
T26X	F	GTTTATGGCCCCACTT**NNK**CGTCTGCGCAGCATCG	75.9	58.8
	R	CGATGCTGCGCAGACG**MNN**AAGTGGGGCCATAAAC		35
Y68X	F	GCTCAGGCAAAAGGC**NNK**GCCGGTGCACCGGG	78.2	52.9
	R	CCCGGTGCACCGGC**MNN**GCCTTTTGCCTGAGC		51.3
W102X	F	CGTATTGCGGTTCAGCTG**NNK**CACACCGGTCGTATC	76.1	45.4
	R	GATACGACCGGTGTG**MNN**CAGCTGAACCGCAATACG		44.9
H181X	F	CGACCTGGTTGAGCTT**NNK**TCTGCGCACGGTTACC	70	41
	R	GGTAACCGTGCGCAGA**MNN**AAGCTCAACCAGGTCG		43
H184X	F	GTTGAGCTTCACTCTGCG**NNK**GGTTACCTGCTGCATCAG	71	34.2
	R	CTGATGCAGCAGGTAACC**MNN**CGCAGAGTGAAGCTCAAC		35.3
Y186X	F	CACTCTGCGCACGGT**NNK**CTGCTGCATCAGTTCCTG	76.1	32.3
	R	CAGGAACTGATGCAGCAG**MNN**ACCGTGCGCAGAGTG		40.7
Q241W	F	CCCCGATCGGTACTTTC**NNK**AACGTCGACAACGGTC	76.1	40.8
	R	GACCGTTGTCGACGTT**MNN**GAAAGTACCGATCGGGG		41.7
Y351X	F	GTCCTGAAAGCTTC**NNK**GGCGGCGGCGCGG	76.3	51.7
	R	CCGCGCCGCCGCC**MNN**GAAGCTTTCAGGAC		49

References

1. Williams, R.E. and N.C. Bruce, *'new uses for an old enzyme' - the old yellow enzyme family of flavoenzymes.* Microbiology, 2002. **148**(6): p. 1607-1614.

2. Vaz, A.D.N., S. Chakraborty, and V. Massey, *Old yellow enzyme - aromatization of cyclic enones and the mechanism of a novel dismutation reaction.* Biochemistry, 1995. **34**(13): p. 4246-4256.

3. Binks, P.R., C.E. French, S. Nicklin, and N.C. Bruce, *Degradation of pentaerythritol tetranitrate by enterobacter cloacae pb2.* Applied and Environmental Microbiology, 1996. **62**(4): p. 1214-1219.

4. Williams, R.E., D.A. Rathbone, N.S. Scrutton, and N.C. Bruce, *Biotransformation of explosives by the old yellow enzyme family of flavoproteins.* Applied and Environmental Microbiology, 2004. **70**(6): p. 3566-3574.

5. French, C.E., S. Nicklin, and N.C. Bruce, *Aerobic degradation of 2,4,6-trinitrotoluene by enterobacter cloacae pb2 and by pentaerythritol tetranitrate reductase.* Applied and Environmental Microbiology, 1998. **64**(8): p. 2864-2868.

6. Muller, A., R. Sturmer, B. Hauer, and B. Rosche, *Stereospecific alkyne reduction: Novel activity of old yellow enzymes.* Angewandte Chemie-International Edition, 2007. **46**(18): p. 3316-3318.

7. Warburg, O. and W. Christian, *Ein zweites sauerstoff bertragendes ferment und sein absorptionsspektrum.* Naturwissenschaften, 1932. **V20**(37): p. 688-688.

8. Akeson, A. and H. Theorell, *Molecular weight and fmn content of crystallin old yellow enzyme.* Arch Biochem Biophys, 1956. **65**(1): p. 439-48.

9. Karplus, P.A., K.M. Fox, and V. Massey, *Flavoprotein structure and mechanism. Structure-function relations for old yellow enzyme.* FASEB Journal, 1995. **9**(15): p. 1518-1526.

10. *Hugo theorell - biography*: Nobelprize.org.

11. Matthews, R.G. and V. Massey, *Isolation of old yellow enzyme in free and complexed forms.* Journal of Biological Chemistry, 1969. **244**(7): p. 1779-1786.

12. Miura, R., T. Yamano, and Y. Miyake, *The heterogeneity of brewer's yeast old yellow enzyme.* Journal of Biochemistry, 1986. **99**(3): p. 901-906.

13. Massey, V. and L.M. Schopfer, *Reactivity of old yellow enzyme with alpha-nadph and other pyridine-nucleotide derivatives.* Journal of Biological Chemistry, 1986. **261**(3): p. 1215-1222.

14. Saito, K., D.J. Thiele, M. Davio, O. Lockridge, and V. Massey, *The cloning and expression of a gene encoding old yellow enzyme from saccharomyces carlsbergensis.* Journal of Biological Chemistry, 1991. **266**(31): p. 20720-20724.

15. Fox, K.M. and P.A. Karplus, *Old yellow enzyme at 2 angstrom resolution: Overall structure, ligand binding, and comparison with related flavoproteins.* Structure, 1994. **2**(11): p. 1089-1105.

16. Stott, K., K. Saito, D.J. Thiele, and V. Massey, *Old yellow enzyme. The discovery of multiple isozymes and a family of related proteins.* Journal of Biological Chemistry, 1993. **268**(9): p. 6097-6106.

17. Niino, Y.S., S. Chakraborty, B.J. Brown, and V. Massey, *A new old yellow enzyme of saccharomyces cerevisiae.* Journal of Biological Chemistry, 1995. **270**(5): p. 1983-1991.

18. French, C.E. and N.C. Bruce, *Purification and characterization of morphinone reductase from pseudomonas putida m10.* Biochemical Journal, 1994. **301**(1): p. 97-103.

19. French, C.E. and N.G. Bruce, *Bacterial morphinone reductase is related to old yellow enzyme.* Biochemical Journal, 1995. **312**(3): p. 671-678.

20. Moody, P.C.E., N. Shikotra, C.E. French, N.C. Bruce, and N.S. Scrutton, *Crystallization and preliminary diffraction studies of morphinone reductase, a flavoprotein involved in the degradation of morphine alkaloids.* Acta Crystallographica Section D: Biological Crystallography, 1997. **53**(5): p. 619-621.

21. French, C.E.C.E., *Sequence and properties of pentaerythritol tetranitrate reductase from enterobacter cloacae pb2.* Journal of bacteriology, 1996. **178**(22): p. 6623-6627.

22. Van Dillewijn, P., R.M. Wittich, A. Caballero, and J.L. Ramos, *Subfunctionality of hydride transferases of the old yellow enzyme family of flavoproteins of pseudomonas putida.* Applied and Environmental Microbiology, 2008. **74**(21): p. 6703-6708.

23. Snape, J.R.J.R., *Purification, properties, and sequence of glycerol trinitrate reductase from agrobacterium radiobacter.* Journal of bacteriology, 1997. **179**(24): p. 7796-7802.

24. Schaller, F.F., *Molecular cloning and characterization of 12-oxophytodienoate reductase, an enzyme of the octadecanoid signaling pathway from arabidopsis thaliana. Structural and functional relationship to yeast old yellow enzyme.* The Journal of biological chemistry, 1997. **272**(44): p. 28066-28072.

25. Straßner, J., A. Fu?rholz, P. Macheroux, N. Amrhein, and A. Schaller, *A homolog of old yellow enzyme in tomato. Spectral properties and substrate specificity of the recombinant protein.* Journal of Biological Chemistry, 1999. **274**(49): p. 35067-35073.

26. Blehert, D.S., B.G. Fox, and G.H. Chambliss, *Cloning and sequence analysis of two pseudomonas flavoprotein xenobiotic reductases.* Journal of Bacteriology, 1999. **181**(20): p. 6254-6263.

27. Fitzpatrick, T.B., N. Amrhein, and P. Macheroux, *Characterization of yqjm, an old yellow enzyme homolog from bacillus subtilis involved in the oxidative stress response.* Journal of Biological Chemistry, 2003. **278**(22): p. 19891-19897.

28. Adalbjornsson, B.V., H.S. Toogood, A. Fryszkowska, C.R. Pudney, T.A. Jowitt, D. Leys, and N.S. Scrutton, *Biocatalysis with thermostable enzymes: Structure and properties of a thermophilic 'ene'-reductase related to old yellow enzyme.* Chembiochem, 2010. **11**(2): p. 197-207.

29. Opperman, D.J., L.A. Piater, and E. van Heerden, *A novel chromate reductase from thermus scotoductus sa-01 related to old yellow enzyme.* Journal of Bacteriology, 2008. **190**(8): p. 3076-3082.

30. Mizugaki, M., M. Nakazawa, H. Yamamoto, and H. Yamanaka, *Purification and characterization of n-ethylmaleimide reducing enzyme from candida lipolytica.* Journal of Biochemistry, 1989. **105**(5): p. 782-784.

31. Cheng, J.Z., C.M. Coyle, D.G. Panaccione, and S.E. O'Connor, *A role for old yellow enzyme in ergot alkaloid biosynthesis.* Journal of the American Chemical Society, 2010. **132**(6): p. 1776-+.

32. Kataoka, M., A. Kotaka, R. Thiwthong, M. Wada, S. Nakamori, and S. Shimizu, *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, 2004. **114**(1-2): p. 1-9.

33. Adachi, O., K. Matsushita, E. Shinagawa, and M. Ameyama, *Occurrence of old yellow enzyme in gluconobacter suboxydans, and the cyclic regeneration of nadp.* Journal of Biochemistry, 1979. **86**(3): p. 699-709.

34. Komduur, J.A., A.N. Lea?o, I. Monastyrska, M. Veenhuis, and J.A.K.W. Kiel, *Old yellow enzymes confers resistance of hansenula polymorpha towards allyl alcohol.* Current Genetics, 2002. **41**(6): p. 401-406.

35. Miranda, M., J. Ramirez, S. Guevara, L. Ongaylarios, A. Pena, and R. Coria, *Nucleotide-sequence and chromosomal localization of the gene encoding the old yellow enzyme from kluyveromyces-lactis.* Yeast, 1995. **11**(5): p. 459-465.

36. Kubata, B.K., Z. Kabututu, T. Nozaki, C.J. Munday, S. Fukuzumi, K. Ohkubo, M. Lazarus, T. Maruyama, S.K. Martin, M. Duszenko, and Y. Urade, *A key role for old yellow enzyme in the metabolism of drugs by trypanosoma cruzi.* Journal of Experimental Medicine, 2002. **196**(9): p. 1241-1251.

37. Brige, A., D. Van Den Hemel, W. Carpentier, L. De Smet, and J.J. Van Beeumen, *Comparative characterization and expression analysis of the four old*

yellow enzyme homologues from shewanella oneidensis indicate differences in physiological function. Biochemical Journal, 2006. **394**(1): p. 335-344.

38. Khan, H., R.J. Harris, T. Barna, D.H. Craig, N.C. Bruce, A.W. Munro, P.C.E. Moody, and N.S. Scrutton, *Kinetic and structural basis of reactivity of pentaerythritol tetranitrate reductase with nadph, 2-cyclohexenone, nitroesters, and nitroaromatic explosives.* Journal of Biological Chemistry, 2002. **277**(24): p. 21906-21912.

39. French, C.E., A.M. Hailes, D.A. Rathbone, and N.C. Bruce, *Morphinone reductase. Characterization, cloning, and application to biocatalytic hydromorphone production.* Annals of the New York Academy of Sciences, 1996. **799**: p. 97-101.

40. Barna, T.M., H. Khan, N.C. Bruce, I. Barsukov, N.S. Scrutton, and P.C.E. Moody, *Crystal structure of pentaerythritol tetranitrate reductase: "Flipped" Binding geometries for steroid substrates in different redox states of the enzyme.* Journal of Molecular Biology, 2001. **310**(2): p. 433-447.

41. Barna, T., H.L. Messiha, C. Petosa, N.C. Bruce, N.S. Scrutton, and P.C.E. Moody, *Crystal structure of bacterial morphinone reductase and properties of the c191a mutant enzyme.* Journal of Biological Chemistry, 2002. **277**(34): p. 30976-30983.

42. Breithaupt, C., R. Kurzbauer, H. Lilie, A. Schaller, J. Strassner, R. Huber, P. Macheroux, and T. Clausen, *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, 2006. **103**(39): p. 14337-14342.

43. Kitzing, K., T.B. Fitzpatrick, C. Wilken, J. Sawa, G.P. Bourenkov, P. Macheroux, and T. Clausen, *The 1.3 angstrom crystal structure of the flavoprotein yqjm reveals a novel class of old yellow enzymes.* Journal of Biological Chemistry, 2005. **280**(30): p. 27904-27913.

44. Nivinskas, H., J. Sarlauskas, Z. Anusevicius, H.S. Toogood, N.S. Scrutton, and N. Cenas, *Reduction of aliphatic nitroesters and n-nitramines by enterobacter*

cloacae pb2 pentaerythritol tetranitrate reductase. FEBS Journal, 2008. **275**(24): p. 6192-6203.

45. Breithaupt, C., J. Strassner, U. Breitinger, R. Huber, P. Macheroux, A. Schaller, and T. Clausen, *X-ray structure of 12-oxophytodienoate reductase 1 provides structural insight into substrate binding and specificity within the family of oye.* Structure, 2001. **9**(5): p. 419-429.

46. Beynon, E.R., Z.C. Symons, R.G. Jackson, A. Lorenz, E.L. Rylott, and N.C. Bruce, *The role of oxophytodienoate reductases in the detoxification of the explosive 2,4,6-trinitrotoluene by arabidopsis.* Plant Physiology, 2009. **151**(1): p. 253-261.

47. Odat, O., S. Matta, H. Khalil, S.C. Kampranis, R. Pfau, P.N. Tsichlis, and A.M. Makris, *Old yellow enzymes, highly homologous fmn oxidoreductases with modulating roles in oxidative stress and programmed cell death in yeast.* Journal of Biological Chemistry, 2007. **282**(49): p. 36010-36023.

48. Reekmans, R., K. De Smet, C.Y. Chen, P. Van Hummelen, and R. Contreras, *Old yellow enzyme interferes with bax-induced nadph loss and lipid peroxidation in yeast.* Fems Yeast Research, 2005. **5**(8): p. 711-725.

49. Uchiyama, N., Z. Kabututu, B.K. Kubata, F. Kiuchi, M. Ito, J. Nakajima-Shimada, T. Aoki, K. Ohkubo, S. Fukuzumi, S.K. Martin, G. Honda, and Y. Urade, *Antichagasic activity of komaroviquinone is due to generation of reactive oxygen species catalyzed by trypanosoma cruzi old yellow enzyme.* Antimicrobial Agents and Chemotherapy, 2005. **49**(12): p. 5123-5126.

50. Stuermer, R., B. Hauer, M. Hall, and K. Faber, *Asymmetric bioreduction of activated c=c bonds using enoate reductases from the old yellow enzyme family.* Current Opinion in Chemical Biology, 2007. **11**(2): p. 203-213.

51. Chaparro-Riggers, J.F., T.A. Rogers, E. Vazquez-Figueroa, K.M. Polizzi, and A.S. Bommarius, *Comparison of three enoate reductases and their potential use for biotransformations.* Advanced synthesis & catalysis, 2007. **349**(8-9): p. 1521-1531.

52. Hall, M., C. Stueckler, H. Ehammer, E. Pointner, G. Oberdorfer, K. Gruber, B. Hauer, R. Stuermer, W. Kroutil, P. Macheroux, and K. Faber, *Asymmetric bioreduction of c=c bonds using enoate reductases opr1, opr3 and yqjm: Enzyme-based stereocontrol.* Advanced synthesis & catalysis, 2008. **350**(3): p. 411-418.

53. Hall, M., C. Stueckler, B. Hauer, R. Stuermer, T. Friedrich, M. Breuer, W. Kroutil, and K. Faber, *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, 2008(9): p. 1511-1516.

54. Stueckler, C., M. Hall, H. Ehammer, E. Pointner, W. Kroutil, P. Macheroux, and K. Faber, *Stereocomplementary bioreduction of alpha,beta-unsaturated dicarboxylic acids and dimethyl esters using enoate reductases: Enzyme- and substrate-based stereocontrol.* Organic Letters, 2007. **9**(26): p. 5409-5411.

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

56. Boonstra, B., D.A. Rathbone, and N.C. Bruce, *Engineering novel biocatalytic routes for production of semisynthetic opiate drugs.* Biomolecular Engineering, 2001. **18**(2): p. 41-47.

57. Wittich, R.M., A. Haidour, P. Van Dillewijn, and J.L. Ramos, *Oye flavoprotein reductases initiate the condensation of tnt-derived intermediates to secondary diarylamines and nitrite.* Environmental Science & Technology, 2008. **42**(3): p. 734-739.

58. Amari, F., A. Fettouche, M. Abou Samra, P. Kefalas, S.C. Kampranis, and A.M. Makris, *Antioxidant small molecules confer variable protection against oxidative damage in yeast mutants.* Journal of Agricultural and Food Chemistry, 2008. **56**(24): p. 11740-11751.

59. Fryszkowska, A., H. Toogood, M. Sakuma, J.M. Gardiner, G.M. Stephens, and N.S. Scrutton, *Asymmetric reduction of activated alkenes by pentaerythritol tetranitrate reductase: Specificity and control of stereochemical outcome by reaction optimisation.* Advanced synthesis & catalysis, 2009. **351**(17): p. 2976-2990.

60. Swiderska, M.A. and J.D. Stewart, *Stereoselective enone reductions by saccharomyces carlsbergensis old yellow enzyme.* Journal of Molecular Catalysis B-Enzymatic, 2006. **42**(1-2): p. 52-54.

61. Bougioukou, D., S. Kille, A. Taglieber, and M. Reetz, *Directed evolution of an enantioselective enoate-reductase: Testing the utility of iterative saturation mutagenesis.* Advanced synthesis & catalysis, 2009. **351**(18): p. 3287-3305.

62. Mueller, N.J., C. Stueckler, B. Hauer, N. Baudendistel, H. Housden, N.C. Bruce, and K. Faber, *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 & Catalysis, 2010. **352**(2-3): p. 387-394.

63. Toogood, H.S., A. Fryszkowska, V. Hare, K. Fisher, A. Roujeinikova, D. Leys, J.M. Gardiner, G.M. Stephens, and N.S. Scrutton, *Structure-based insight into the asymmetric bioreduction of the c=c double bond of α/β - unsaturated nitroalkenes by pentaerythritol tetranitrate reductase.* Advanced Synthesis & Catalysis, 2008. **350**(17): p. 2789-2803.

64. Pak, J.W., K.L. Knoke, D.R. Noguera, B.G. Fox, and G.H. Chambliss, *Transformation of 2,4,6-trinitrotoluene by purified xenobiotic reductase b from pseudomonas fluorescens i-c.* Applied and Environmental Microbiology, 2000. **66**(11): p. 4742-4750.

65. Fox, K.M. and P.A. Karplus, *The flavin environment in old yellow enzyme - an evaluation of insights from spectroscopic and artificial flavin studies.* Journal of Biological Chemistry, 1999. **274**(14): p. 9357-9362.

66. Xu, D., R.M. Kohli, and V. Massey, *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, 1999. **96**(7): p. 3556-3561.

67. Messiha, H.L., A.W. Munro, N.C. Bruce, I. Barsukov, and N.S. Scrutton, *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.* Journal of Biological Chemistry, 2005. **280**(11): p. 10695-10709.

68. Brown, B.J., J.W. Hyun, S. Duvvuri, P.A. Karplus, and V. Massey, *The role of glutamine 114 in old yellow enzyme.* Journal of Biological Chemistry, 2002. **277**(3): p. 2138-2145.

69. Brown, B.J., Z. Deng, P.A. Karplus, and V. Massey, *On the active site of old yellow enzyme - role of histidine 191 and asparagine 194.* Journal of Biological Chemistry, 1998. **273**(49): p. 32753-32762.

70. Marshall, S.J., D. Krause, D.K. Blencowe, and G.F. White, *Characterization of glycerol trinitrate reductase (nera) and the catalytic role of active-site residues.* Journal of Bacteriology, 2004. **186**(6): p. 1802-1810.

71. Scrutton, N.S., A. Berry, and R.N. Perham, *Redesign of the coenzyme specificity of a dehydrogenase by protein engineering.* Nature, 1990. **343**(6253): p. 38-43.

72. Van Den Hemel, D., A. Brige, S.N. Savvides, and J. Van Beeumen, *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, 2006. **281**(38): p. 28152-28161.

73. Khan, H., T. Barna, N.C. Bruce, A.W. Munro, D. Leys, and N.S. Scrutton, *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 and his184.* FEBS Journal, 2005. **272**(18): p. 4660-4671.

74. Pudney, C.R., S. Hay, M.J. Sutcliffe, and N.S. Scrutton, *A-secondary isotope effects as probes of "Tunneling-ready" Configurations in enzymatic h-tunneling: Insight from environmentally coupled tunneling models.* Journal of the American Chemical Society, 2006. **128**(43): p. 14053-14058.

75. Khan, H., T. Barna, R.J. Harris, N.C. Bruce, I. Barsukov, A.W. Munro, P.C.E. Moody, and N.S. Scrutton, *Atomic resolution structures and solution behavior of enzyme-substrate complexes of enterobacter cloacae pb2 pentaerythritol tetranitrate reductase: Multiple conformational states and implications for the mechanism of nitroaromatic explosive degradation.* Journal of Biological Chemistry, 2004. **279**(29): p. 30563-30572.

76. Meah, Y., B.J. Brown, S. Chakraborty, and V. Massey, *Old yellow enzyme: Reduction of nitrate esters, glycerin trinitrate, and propylene 1,2-dinitrate.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(15): p. 8560-8565.

77. Kohli, R.M. and V. Massey, *The oxidative half-reaction of old yellow enzyme: The role of tyrosine 196.* Journal of Biological Chemistry, 1998. **273**(49): p. 32763-32770.

78. Messiha, H.L., N.C. Bruce, B.M. Sattelle, M.J. Sutcliffe, A.W. Munro, and N.S. Scrutton, *Role of active site residues and solvent in proton transfer and the modulation of flavin reduction potential in bacterial morphinone reductase.* Journal of Biological Chemistry, 2005. **280**(29): p. 27103-27110.

79. Fitzpatrick, T.B., S. Auweter, K. Kitzing, T. Clausen, N. Amrhein, and P. MacHeroux, *Structural and functional impairment of an old yellow enzyme homologue upon affinity tag incorporation.* Protein Expression and Purification, 2004. **36**(2): p. 280-291.

80. Sylvestre, J., H. Chautard, F. Cedrone, and M. Delcourt, *Directed evolution of biocatalysts.* Organic Process Research and Development, 2006. **10**(3): p. 562-571.

81. Neylon, C., *Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: Library construction methods for directed evolution.* Nucleic Acids Research, 2004. **32**(4): p. 1448-1459.

82. Stemmer, W.P.C., *Rapid evolution of a protein in vitro by DNA shuffling.* Nature, 1994. **370**(6488): p. 389-391.

83. Zhao, H.M. and W.J. Zha, *In vitro 'sexual' evolution through the pcr-based staggered extension process (step).* Nature Protocols, 2006. **1**(4): p. 1865-1871.

84. Cadwell, R.C. and G.F. Joyce, *Mutagenic pcr.* PCR-Methods and Applications, 1994. **3**(6): p. S136-S140.

85. Cadwell, R.C. and G.F. Joyce, *Randomization of genes by pcr mutagenesis.* PCR Methods Appl, 1992. **2**(1): p. 28-33.

86. Zaccolo, M., D.M. Williams, D.M. Brown, and E. Gherardi, *An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues.* Journal of Molecular Biology, 1996. **255**(4): p. 589-603.

87. Greener, A., M. Callahan, and B. Jerpseth, *An efficient random mutagenesis technique using an e.Coli mutator strain.* Molecular Biotechnology, 1997. **7**(2): p. 189-195.

88. Chen, K.Q. and F.H. Arnold, *Tuning the activity of an enzyme for unusual environments - sequential random mutagenesis of subtilisin e for catalysis in dimethylformamide.* Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(12): p. 5618-5622.

89. Vanhercke, T., C. Ampe, L. Tirry, and P. Denolf, *Reducing mutational bias in random protein libraries.* Analytical Biochemistry, 2005. **339**(1): p. 9-14.

90. Miyazaki, K. and M. Takenouchi, *Creating random mutagenesis libraries using megaprimer pcr of whole plasmid.* Biotechniques, 2002. **33**(5): p. 1033-+.

91. Gupta, R.D. and D.S. Tawfik, *Directed enzyme evolution via small and effective neutral drift libraries.* Nature Methods, 2008. **5**(11): p. 939-942.

92. Miyazaki, K. and F.H. Arnold, *Exploring nonnatural evolutionary pathways by saturation mutagenesis: Rapid improvement of protein function.* Journal of Molecular Evolution, 1999. **49**(6): p. 716-720.

93. Hughes, M.D., D.A. Nagel, A.F. Santos, A.J. Sutherland, and A.V. Hine, *Removing the redundancy from randomised gene libraries.* Journal of Molecular Biology, 2003. **331**(5): p. 973-979.

94. Reetz, M.T. and S. Wu, *Greatly reduced amino acid alphabets in directed evolution: Making the right choice for saturation mutagenesis at homologous enzyme positions.* Chemical Communications, 2008(43): p. 5499-5501.

95. Reetz, M.T., D. Kahakeaw, and R. Lohmer, *Addressing the numbers problem in directed evolution.* Chembiochem, 2008. **9**(11): p. 1797-1804.

96. Reetz, M.T. and J.D. Carballeira, *Iterative saturation mutagenesis (ism) for rapid directed evolution of functional enzymes.* Nature Protocols, 2007. **2**(4): p. 891-903.

97. Reetz, M.T., D. Kahakeaw, and J. Sanchis, *Shedding light on the efficacy of laboratory evolution based on iterative saturation mutagenesis.* Molecular Biosystems, 2009. **5**(2): p. 115-122.

98. Reetz, M.T., M. Bocola, J.D. Carballeira, D. Zha, and A. Vogel, *Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test.* Angewandte Chemie - International Edition, 2005. **44**(27): p. 4192-4196.

99. Reetz, M.T., L.W. Wang, and M. Bocola, *Directed evolution of enantioselective enzymes: Iterative cycles of casting for probing protein-sequence space.* Angewandte Chemie - International Edition, 2006. **45**(8): p. 1236-1241.

100. Hames, C., S. Halbedel, O. Schilling, and J. Stulke, *Multiple-mutation reaction: A method for simultaneous introduction of multiple mutations into the glpk gene of mycoplasma pneumoniae.* Applied and Environmental Microbiology, 2005. **71**(7): p. 4097-4100.

101. Hidalgo, A., A. Schliessmann, R. Molina, J. Hermoso, and U.T. Bornscheuer, *A one-pot, simple methodology for cassette randomisation and recombination for focused directed evolution.* Protein Engineering Design & Selection, 2008. **21**(9): p. 567-576.

102. Bosley, A.D. and M. Ostermeier, *Mathematical expressions useful in the construction, description and evaluation of protein libraries.* Biomolecular Engineering, 2005. **22**(1-3): p. 57-61.

103. Patrick, W.M., A.E. Firth, and J.M. Blackburn, *User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries.* Protein Engineering, 2003. **16**(6): p. 451-457.

104. Morley, K.L. and R.J. Kazlauskas, *Improving enzyme properties: When are closer mutations better?* Trends in Biotechnology, 2005. **23**(5): p. 231-237.

105. Boersma, Y.L., M.J. Droge, and W.J. Quax, *Selection strategies for improved biocatalysts.* Febs Journal, 2007. **274**(9): p. 2181-2195.

106. Griffiths, J.S., M. Cheriyan, J.B. Corbell, L. Pocivavsek, C.A. Fierke, and E.J. Toone, *A bacterial selection for the directed evolution of pyruvate aldolases.* Bioorganic & Medicinal Chemistry, 2004. **12**(15): p. 4067-4074.

107. Reetz, M.T., H. Hobenreich, P. Soni, and L. Fernandez, *A genetic selection system for evolving enantioselectivity of enzymes.* Chemical Communications, 2008(43): p. 5502-5504.

108. Baker, K., C. Bleczinski, H.N. Lin, G. Salazar-Jimenez, D. Sengupta, S. Krane, and V.W. Cornish, *Chemical complementation: A reaction-independent genetic assay for enzyme catalysis.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(26): p. 16537-16542.

109. Lin, H.N., H.Y. Tao, and V.W. Cornish, *Directed evolution of a glycosynthase via chemical complementation.* Journal of the American Chemical Society, 2004. **126**(46): p. 15051-15059.

110. Fernandez-Gacio, A., M. Uguen, and J. Fastrez, *Phage display as a tool for the directed evolution of enzymes.* Trends in Biotechnology, 2003. **21**(9): p. 408-414.

111. Daugherty, P.S., *Protein engineering with bacterial display.* Current Opinion in Structural Biology, 2007. **17**(4): p. 474-480.

112. Yan, X.H. and Z.R. Xu, *Ribosome-display technology: Applications for directed evolution of functional proteins.* Drug Discovery Today, 2006. **11**(19-20): p. 911-916.

113. Belien, T., K. Hertveldt, K. Van den Brande, J. Robben, S. Van Campenhout, and G. Volckaert, *Functional display of family 11 endoxylanases on the surface of phage m13.* Journal of Biotechnology, 2005. **115**(3): p. 249-260.

114. Strobel, H., D. Ladant, and J.L. Jestin, *In vitro selection for enzymatic activity: A model study using adenylate cyclase.* Journal of Molecular Biology, 2003. **332**(1): p. 1-7.

115. Leemhuis, H., R.M. Kelly, and L. Dijkhuizen, *Directed evolution of enzymes: Library screening strategies.* Iubmb Life, 2009. **61**(3): p. 222-228.

116. Zhang, J.H., G. Dawes, and W.P.C. Stemmer, *Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(9): p. 4504-4509.

117. Van Loo, B., J.H.L. Spelberg, J. Kingma, T. Sonke, M.G. Wubbolts, and D.B. Janssen, *Directed evolution of epoxide hydrolase from a.Radiobacter toward higher enantioselectivity by error-prone pcr and DNA shuffling.* Chemistry & Biology, 2004. **11**(7): p. 981-990.

118. Ben-David, A., G. Shoham, and Y. Shoham, *A universal screening assay for glycosynthases: Directed evolution of glycosynthase xynb2(e335g) suggests a general path to enhance activity.* Chemistry & Biology, 2008. **15**(6): p. 546-551.

119. Shi, C., X.Z. Lu, C.P. Ma, Y.M. Ma, X.Y. Fu, and W.G. Yu, *Enhancing the thermostability of a novel beta-agarase agab through directed evolution.* Applied Biochemistry and Biotechnology, 2008. **151**(1): p. 51-59.

120. Xia, T. and Q. Wang, *Directed evolution of streptomyces lividans xylanase b toward enhanced thermal and alkaline ph stability.* World Journal of Microbiology & Biotechnology, 2009. **25**(1): p. 93-100.

121. Li, H.M., L.H. Mei, V.B. Urlacher, and R.D. Schmid, *Cytochrome p450bm-3 evolved by random and saturation mutagenesis as an effective indole-hydroxylating catalyst.* Applied Biochemistry and Biotechnology, 2008. **144**(1): p. 27-36.

122. Mayer, C., D.L. Jakeman, M. Mah, G. Karjala, L. Gal, R.A.J. Warren, and S.G. Withers, *Directed evolution of new glycosynthases from agrobacterium beta-glucosidase: A general screen to detect enzymes for oligosaccharide synthesis.* Chemistry & Biology, 2001. **8**(5): p. 437-443.

123. Janes, L.E., A.C. Lowendahl, and R.J. Kazlauskas, *Quantitative screening of hydrolase libraries using ph indicators: Identifying active and enantioselective hydrolases.* Chemistry-a European Journal, 1998. **4**(11): p. 2324-2331.

124. Connolly, A.R. and J.D. Sutherland, *Catalyst screening using an array of thermistors.* Angewandte Chemie-International Edition, 2000. **39**(23): p. 4268-+.

125. Klein, G. and J.L. Reymond, *An enzyme assay using pm.* Angewandte Chemie-International Edition, 2001. **40**(9): p. 1771-1773.

126. Miller, O.J., E.G. Hibbert, C.U. Ingram, G.J. Lye, and P.A. Dalby, *Optimisation and evaluation of a generic microplate-based hplc screen for transketolase activity.* Biotechnology Letters, 2007. **29**: p. 1759-1770.

127. Klein, G. and J.L. Reymond, *Enantioselective fluorogenic assay of acetate hydrolysis for detecting lipase catalytic antibodies.* Helvetica Chimica Acta, 1999. **82**(3): p. 400-407.

128. Reetz, M.T., M.H. Becker, H.W. Klein, and D. Stockigt, *A method for high-throughput screening of enantioselective catalysts.* Angewandte Chemie-International Edition, 1999. **38**(12): p. 1758-1761.

129. Carlon, R.P., N. Jourdain, and J.L. Reymond, *Fluorogenic polypropionate fragments for detecting stereoselective aldolases.* Chemistry-a European Journal, 2000. **6**(22): p. 4154-4162.

130. Lafferty, M. and M.J. Dycaico, *Gigamatrix: A novel ultrahigh throughput protein optimization and discovery platform*, in *Protein engineering*. 2004, Elsevier Academic Press Inc: San Diego. p. 119-134.

131. Polizzi, K.M., M. Parikh, C.U. Spencer, I. Matsumura, J.H. Lee, M.J. Realff, and A.S. Bommarius, *Pooling for improved screening of combinatorial libraries for directed evolution.* Biotechnology Progress, 2006. **22**(4): p. 961-967.

132. Yang, G.Y. and S.G. Withers, *Ultrahigh-throughput facs-based screening for directed enzyme evolution.* Chembiochem, 2009. **10**(17): p. 2704-2715.

133. Hao, J. and A. Berry, *A thermostable variant of fructose bisphosphate aldolase constructed by directed evolution also shows increased stability in organic solvents.* Protein Engineering, Design and Selection, 2004. **17**(9): p. 689-697.

134. Reetz, M.T., J. D Carballeira, and A. Vogel, *Iterative saturation mutagenesis on the basis of b factors as a strategy for increasing protein thermostability.* Angewandte Chemie-International Edition, 2006. **45**(46): p. 7745-7751.

135. Bommarius, A.S., J.M. Broering, J.F. Chaparro-Riggers, and K.M. Polizzi, *High-throughput screening for enhanced protein stability.* Current Opinion in Biotechnology, 2006. **17**(6): p. 606-610.

136. Marx, C.K., T.C. Hertel, and M. Pietzsch, *Random mutagenesis of a recombinant microbial transglutaminase for the generation of thermostable and heat-sensitive variants.* Journal of Biotechnology, 2008. **136**(3-4): p. 156-162.

137. Reetz, M.T., P. Soni, and L. Fernandez, *Knowledge-guided laboratory evolution of protein thermolability.* Biotechnology and Bioengineering, 2009. **102**(6): p. 1712-1717.

138. You, L. and F.H. Arnold, *Directed evolution of subtilisin e in bacillus subtilis to enhance total activity in aqueous dimethylformamide.* Protein Eng., 1996. **9**(1): p. 77-83.

139. Ogino, H., T. Uchiho, N. Doukyu, M. Yasuda, K. Ishimi, and H. Ishikawa, *Effect of exchange of amino acid residues of the surface region of the pst-01 protease on its organic solvent-stability.* Biochemical and Biophysical Research Communications, 2007. **358**(4): p. 1028-1033.

140. Qin, Y.Q., X.M. Wei, X. Song, and Y.B. Qu, *Engineering endoglucanase ii from trichoderma reesei to improve the catalytic efficiency at a higher ph optimum.* Journal of Biotechnology, 2008. **135**(2): p. 190-195.

141. Ehren, J., S. Govindarajan, B. Moron, J. Minshull, and C. Khosla, *Protein engineering of improved prolyl endopeptidases for celiac sprue therapy.* Protein Engineering Design & Selection, 2008. **21**(12): p. 699-707.

142. Wan, L.L., M.B. Twitchett, L.D. Eltis, A.G. Mauk, and M. Smith, *In vitro evolution of horse heart myoglobin to increase peroxidase activity.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(22): p. 12825-12831.

143. Lingen, B., J. Grotzinger, D. Kolter, M.R. Kula, and M. Pohl, *Improving the carboligase activity of benzoylformate decarboxylase from pseudomonas putida by a combination of directed evolution and site-directed mutagenesis.* Protein Engineering, 2002. **15**(7): p. 585-593.

144. Spadiut, O., I. Pisanelli, T. Maischberger, C. Peterbauer, L. Gorton, P. Chaiyen, and D. Haltrich, *Engineering of pyranose 2-oxidase: Improvement for biofuel cell and food applications through semi-rational protein design.* Journal of Biotechnology, 2009. **139**(3): p. 250-257.

145. Sohn, H., Y.S. Kim, U.H. Jin, S.J. Suh, S.C. Lee, D.S. Lee, J.H. Ko, and C.H. Kim, *Alteration of the substrate specificity of thermus caldophilus adp-glucose*

pyrophosphorylase by random mutagenesis through error-prone polymerase chain reaction. Glycoconjugate Journal, 2006. **23**(9): p. 619-625.

146. Hibbert, E.G., S.A. Tarik, E. Mark, S.J. Costelloe, J.M. Ward, H.C. Hailes, and P.A. Dalby, *Directed evolution of transketolase substrate specificity towards an aliphatic aldehyde.* Journal of Biotechnology, 2008. **134**(3-4): p. 240-245.

147. Wu, Z.L., L.M. Podust, and F.P. Guengerich, *Expansion of substrate specificity of cytochrome p450 2a6 by random and site-directed mutagenesis.* Journal of Biological Chemistry, 2005. **280**(49): p. 41090-41100.

148. Padhi, S.K., D.J. Bougioukou, and J.D. Stewart, *Site-saturation mutagenesis of tryptophan 116 of saccharomyces pastorianus old yellow enzyme uncovers stereocomplementary variants.* Journal of the American Chemical Society, 2009. **131**(9): p. 3271-3280.

149. Williams, G.J., T. Woodhall, L.M. Farnsworth, A. Nelson, and A. Berry, *Creation of a pair of stereochemically complementary biocatalysts.* Journal of the American Chemical Society, 2006. **128**(50): p. 16238-16247.

150. Williams, G.J., S. Domann, A. Nelson, and A. Berry, *Modifying the stereochemistry of an enzyme-catalyzed reaction by directed evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(6): p. 3143-3148.

151. Reetz, M.T., A. Zonta, K. Schimossek, K. Liebeton, and K.E. Jaeger, *Creation of enantioselective biocatalysts for organic chemistry by in vitro evolution.* Angewandte Chemie-International Edition, 1997. **36**(24): p. 2830-2832.

152. Reetz, M.T., M. Bocola, L.W. Wang, J. Sanchis, A. Cronin, M. Arand, J.Y. Zou, A. Archelas, A.L. Bottalla, A. Naworyta, and S.L. Mowbray, *Directed evolution of an enantioselective epoxide hydrolase: Uncovering the source of enantioselectivity at each evolutionary stage.* Journal of the American Chemical Society, 2009. **131**(21): p. 7334-7343.

153. Mugford, P.F., U.G. Wagner, Y. Jiang, K. Faber, and R.J. Kazlauskas, *Enantiocomplementary enzymes: Classification, molecular basis for their enantiopreference, and prospects for mirror-image biotransformations.* Angewandte Chemie-International Edition, 2008. **47**(46): p. 8782-8793.

154. Strickland, S., G. Palmer, and V. Massey, *Determination of dissociation-constants and specific rate of constants of enzyme-substrate (or protein-ligand) interactions from rapid reaction kinetic data.* Journal of Biological Chemistry, 1975. **250**(11): p. 4048-4052.

155. Leslie, A.G.W., *Recent changes to the mosflm package for processing film and image plate data.* Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26, 1992.

156. Dodson, E.J., M. Winn, and A. Ralph, *Collaborative computational project, number 4: Providing programs for protein crystallography.* Macromolecular Crystallography, Pt B, 1997. **277**: p. 620-633.

157. Perrakis, A., R. Morris, and V.S. Lamzin, *Automated protein model building combined with iterative structure refinement.* Nature Structural Biology, 1999. **6**(5): p. 458-463.

158. Emsley, P. and K. Cowtan, *Coot: Model-building tools for molecular graphics.* Acta Crystallographica Section D-Biological Crystallography, 2004. **60**: p. 2126-2132.

159. Brideau, C., B. Gunter, B. Pikounis, and A. Liaw, *Improved statistical methods for hit selection in high-throughput screening.* Journal of Biomolecular Screening, 2003. **8**(6): p. 634-647.

160. Hengen, P.N., *Purification of his-tag fusion proteins from escherichia coli.* Trends in Biochemical Sciences, 1995. **20**(7): p. 285-286.

161. Hofmann, K. and Y. Kiso, *An approach to the targeted attachment of peptides and proteins to solid supports.* Proceedings of the National Academy of Sciences of the United States of America, 1976. **73**(10): p. 3516-3518.

162. Basran, J., R.J. Harris, M.J. Sutcliffe, and N.S. Scrutton, *H-tunneling in the multiple h$^-$ transfers of the catalytic cycle of morphinone reductase and in the reductive half-reaction of the homologous pentaerythritol tetranitrate reductase.* Journal of Biological Chemistry, 2003. **278**(45): p. 43973-43982.

163. Hottenrott, S., T. Schumann, A. Pluckthun, G. Fischer, and J.U. Rahfeld, *The escherichia coli slyd is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase.* Journal of Biological Chemistry, 1997. **272**(25): p. 15697-15701.

164. *Clustal*. [cited; Available from: www.ebi.ac.uk/Tools/clustalw2

165. Qiagen, *The qiaexpressionist*. 2003.

166. Schatz, P.J., *Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in escherichia coli.* Nature Biotechnology, 1993. **11**(10): p. 1138-1143.

167. *Avidity*. [cited; Available from: http://www.avidity.com/.

168. Studier, F.W., *Protein production by auto-induction in high-density shaking cultures.* Protein Expression and Purification, 2005. **41**(1): p. 207-234.

169. Toogood, H.S., J.M. Gardiner, and N.S. Scrutton, *Biocatalytic reductions and chemical versatility of the old yellow enzyme family of flavoprotein oxidoreductases.* ChemCatChem, 2010. **9999**(9999): p. NA.

170. Toogood, H., A. Fryszkowska, M. Hulley, M. Sakuma, D. Mansell, G. Stephens, J. Gardiner, and N. Scrutton, *A site-saturated mutagenesis study reveals that residues 181 and 184 influence ligand binding, stereochemistry and oxime formation in pentaerythritol tetranitrate reductase. .* Submitted to Chembiochem (Sept 2010), 2010.

171. Pudney, C.R., S. Hay, J.Y. Pang, C. Costello, D. Leys, M.J. Sutcliffe, and N.S. Scrutton, *Mutagenesis of morphinone reductase induces multiple reactive configurations and identifies potential ambiguity in kinetic analysis of enzyme tunneling mechanisms.* Journal of the American Chemical Society, 2007. **129**: p. 13949-13956.

172. Kegler-Ebo, D.M., C.M. Docktor, and D. DiMaio, *Codon cassette mutagenesis: A general method to insert or replace individual codons by using universal mutagenic cassettes.* Nucleic Acids Research, 1994. **22**(9): p. 1593-1599.