THE UNIVERSITY
of EDINBURGH

Biocatalytic application of rare PLP-dependent aminotransferases for the synthesis of high value amino acids and amines

Annabel Serpico

A Thesis Submitted for the Degree of

Doctor of Philosophy

The University of Edinburgh

2018

Lay Summary

Traditional chemical approaches for the production of compounds often rely on harsh conditions such as high temperature and pressure and create toxic waste and pollutants. Moreover, these reactions are often non efficient, creating unwanted side products that need to be disposed of. A very important branch of chemistry, called green chemistry, supports the design of products and processes that use the least amount of dangerous substances as they can. An important aspect of green chemistry is biocatalysis, which consists of the use of biological systems, either whole cells or enzymes for the production of compounds. Indeed, enzymes represent a very valid alternative for the production of chemicals in a more 'green', eco-friendly way. Enzymes have the main advantage of requiring mild conditions (room temperature, atmospheric pressure) and being extremely efficient, minimizing the production of unwanted side products. Moreover, enzymes are biodegradable which means that they are degraded over time.

An important class of compounds that find a broad range of applications from pharmaceuticals to personal care products are the chiral amines. The traditional chemical approach for their production requires dangerous metals and is very inefficient. A very valid alternative for their production is represented by aminotransferase enzymes that are able to yield the desired product in high purity.

In this work two aminotransferases, the D-Phenylgylcine aminotransferase (D-PhgAT) and the amino-pentol aminotransferase FumI are explored as tools for the production of high value compounds. These compounds are the D-phenylglycines, which are key molecules for the production of antibiotics and the fatty amines that are used as detergents and are widely present in cosmetic formulations.

Abstract

Optically active amines and amino acids are ubiquitously distributed in nature where they play many crucial roles. Moreover, it was recently estimated that around 40% of blockbuster drugs and 20% of agrochemicals contain chiral amines in their structure, thus there has been considerable effort in developing efficient, low cost and widely-applicable methods for their production. Compared to classical chemical synthesis, aminotransferase (AT) enzymes have been widely explored as a more efficient and sustainable method for the preparation of optically pure amines from the corresponding ketones. In several cases these biocatalysts have replaced the existing chemical catalysts, which are limited by their high cost and poor regio- and stereoselectivity.

Two very interesting, distantly-related class III bacterial ATs have been investigated in this study: the D-phenylglycine (D-Phg) aminotransferase (D-PhgAT) and the amino-pentol AT (FumI). The D-PhgAT from *Pseudomonas stutzeri* ST-201 catalyses the reversible transamination from L-glutamic acid to benzoylformate, yielding α-ketoglutarate and D-Phg with high (>99%) enantiomeric excess (% *ee*). The D-PhgAT possesses the unique feature that the amino acid donor and amino acid product display an inverted absolute configuration. Thus, D-PhgAT is a very promising biocatalyst as it yields high value D-amino acids from inexpensive L-amino acid donors.

By carrying out a detailed kinetic analysis of recombinant D-PhgAT the best substrates were identified and an optimized method, delivering a range of enantiopure aromatic D-amino acids at 1 g scale, has been developed. Moreover, the x-ray crystal structure of D-PhgAT at 2.25 Å resolution has been determined with its pyridoxal 5'-phosphate (PLP) cofactor bound as an internal aldimine. The active site architecture highlights various residues potentially involved in catalysis and illuminates the basis of the exquisite enantioselectivity of this unique member of the AT superfamily. These studies promote D-PhgAT as a useful tool for the sustainable production of high value, aromatic D-amino acids.

Fatty amines (C_8-C_{20}) are high value products with a wide variety of applications in fabric softeners, detergents and cosmetic formulations. The unusual *Sphingopyxis* FumI detoxifies the C_{20} mycotoxin natural product fumonisin B1 (FB_1). FumI catalyses

transamination at the C2 position of a hydrolyzed fumonisin B1 (HFB_1) derivative with pyruvate as the amino acceptor. A screen of prochiral C2 ketones of varying chain length (C3-C17 2K) with L-alanine as amino donor has revealed that the recombinant FumI is capable of producing alkyl-amines of various chain lengths (C_3-C_{17}) by working in the reverse direction. Chiral HPLC analysis has revealed that the enzyme generates *(S)*-amines with high *ee* (>99%). The regioselectivity of FumI was also determined using prochiral keto-substrates with substitutions along the chain (such as 3-decanone (C10-3K) and 5-decanone (C10-5K)). Interestingly, C10-3K, but not C10-5K, was converted to the corresponding amine. A range of long chain aldehydes (C_{10}-C_{20}) were also converted by FumI to the corresponding amines with L-Ala as the amino donor. The x-ray crystal structure of the PLP-bound, internal aldimine form of FumI was solved to a resolution of 1.6 Å by molecular replacement using a *Bacillus subtilis* AT as a model. The structure suggested residues potentially involved in catalysis and revealed a potential hydrophobic binding site for the long alkyl chain substrates. The combination of aldehyde and ketone alkyl chain substrate promiscuity, the high enantioselectivity and the 3D structures of this rare AT suggests that FumI is a very useful addition to the biocatalytic toolbox for the synthesis of high value fatty amines.

Acknowledgements

First of all, I would like to express my biggest gratitude to my supervisor Prof. Dominic Campopiano who gave me the opportunity to join his research group four years ago. Thanks for the trust, the support, the challenges, the responsibilities and the rewards, the enthusiasm and the inspiration. It has been a great pleasure and a great honour.

Thanks to Dr. Jon Marles-Wright for his amazing support and his help with the crystallography. He has been a great collaborator and a great person to work with.

Thanks to all the past and present members of the Campopiano group. I have been lucky to share this experience with some amazing people and scientists, some of which I am sure will be part of my life for much longer. It has been a great privilege and great fun.

Thanks to Piera, who has shared this journey with me since day one until the very last day, from the first coffee on the royal mile to the burger at 8 p.m. in the office the day before submitting this thesis. Thanks for being my partner in crime in this journey, without her it wouldn't have been the same.

Thanks to Guio, for being the light when I saw just dark and for always taking care of me, no matter in which country she is. I cannot wait to work with her again.

Thanks to Pete, for the chats at 7 in the morning and the 'posh' coffees at any times. Thanks for fixing the unfixable and always make me feel special.

Thanks to Alexis, for the laugh, the dance, the craziness and the joy that he brought into my life and in the lab.

Thanks to Jo, for the hugs and the laughs, for bringing roller skating back in my life and sharing amazing adventures with me.

Thanks to Gary, for being always there with hugs, massages and sweets; to Ben(ito), for the mess and the laughs; to Silvia, for the amazing cakes; to Stacie, for her smile; to Catherine, for keeping this project alive with my same enthusiasm and Mark for the endless distraction.

Thanks go also to Amanda, Bohdy, Alex, Chris, Van, Lily, your help was really appreciated.

Thanks to all my project students, most of them drove me mad but I have learnt something from each one of them. Thanks to all the other students who have been part of this group, some of you are really special to me and I hope we'll keep in touch.

A special thanks to Vero for being my family, my rock, my shoulder and my arm. I have to thank her for surviving fire, water and every possible challenge. Thanks for being amazing without even noticing, I cannot wait for her to realize it.

Thanks to my girls, Alisia and Sally, for our adventures around Europe and the unsuccessful time spent at the gym. I cannot wait for more to come.

Thanks to all the Skye crew (in particular Lindsey and Stefan) for sharing with me an incredible weekend that I will never forget.

Thanks to my favourite Italians in Edinburgh for making me feel at home when I was homesick with the food, the laughs and the loudness.

Thanks to everyone that always cared about me, no matter where I was. A special thanks to who found time to visit me, in particular Laura and Simona, I am really lucky to have found you in Naples 6 years ago and I am sure I will have you in my life for much longer.

A special thanks to my family. It's hard to put into words exactly how thankful I am for them and for everything they have ever done for me. Life without them wouldn't be the same and neither would I.

Thanks to my mum for her endless love and her unconditional support even when she didn't agree with me (most of the time). Thanks for always taking care of me, even from far away.

Thanks to my dad for being my biggest fan and my first love. Thanks for never stop counting the days until the next time he would have seen me again and make me feel always the luckiest daughter ever.

Thanks to my sister, my best friend, my soulmate and my other half .Thanks for always being able to find the words, no matter which language I am speaking.

Thanks to my little brother, my biggest weakness, because he makes my heart melt every time he smiles, no matter what happened.

And finally, thanks to YOU, my lover, my partner, my best friend. Thank you for teaching me that distance means nothing when someone means everything. Thanks for believing in me even when I didn't. Thanks for never stop fighting for me and for us. Thanks for supporting me in every step, every choice, every mistake and every success. Without you this wouldn't have been possible. I cannot wait to share the rest of my life with you.

Declaration

I, Annabel Serpico, declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree or professional qualification. Except where otherwise stated by reference or acknowledgment, the work presented is entirely my own.

Annabel Serpico

The University of Edinburgh

2018

List of abbreviations

2,3-DHIA	5-nitro-2, 3-dihydro-1H-inden-2-amine
4-AEB	4-(2-Aminoethyl) benzonitrile
4-NPEA	2-(4-nitrophenyl) ethan-1-amine
Aba	2 Aminobutyric Acid
ACN	Acetonitrile
ADH	Alanine Dehydrogenase
AdK	Adenylate Kinase
A-domain	Adenylation Domain
ADP	Adenosine 5-diphosphate
AKG	α-ketoglutarate
AlkJ	Alcohol Dehydrogenase
AmDH	Amine Dehydrogenases
AMP	Adenosine 5'-Monophosphate
AO	Amine Oxidases
API	Active Pharmaceutical Ingredients
AspAT	Aspartate Aminotransferases
AT	Aminotransferase
ATP	Adenosine 5'-Triphosphate
AUC	Area Under the Curve
Bm-AT	*Bacillus megaterium* Aminotransferase
BSA	Serum Albumin Bovine
BZALD	Benzaldehyde
BZF	Benzoylformate
C. violaceum	*Chromobacterium violaceum*
C10-1A	Decylamine
C10-2A	2-Decanamine
C10-2K	2-Decanone
C10-3K	3-Decanone
C10-5K	5-Decanone

C10-Al	Decanal
C10-FA	Decanoic acid
C12-1A	Dodecylamine
C12-Al	Dodecanal
C12-FA	Lauric acid
C13-2A	2-Tridecanamine
C13-2K	2-Tridecanone
C14-FA	Myristic Acid
C15-2A	2-Pentadecanamine
C15-2K	2- Pentadecanone
C16-1A	hexadecylamine
C16-Al	Hexadecanal
C16-FA	Palmitic acid
C17-2A	2-Heptadecanamine
C17-2K	2-Heptadecanone
C18-FA	Stearic Acid
C20-FA	Arachidic acid
C3-2A	Isopropylamine
C3-2K	Acetone
C4-2A	2-Butylamine
C4-2K	Butanone
C6-2A	2-Aminohexane
C6-2K	Hexanone
C8-2A	2-Aminooctanone
C8-2K	Octanone
CAPS	N-Cyclohexyl-3-Aminopropanesulfonic Acid
CAR	Carboxylic Acid Reductase
cHPLC	Chiral HPLC
CoASH	Coenzyme A
CPG	Controlled Pore Glass
CSA	Camphor-8-Sulfonic Acid

CV	Column Volume
Cv-AT	*Chromobacterium violaceum* Aminotransferase
DH	Dehydrogenase
D-Hpg	D-4-Hydroxyphenylglycine
DKR	Dynamic Kinetic Resolution
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
DOPA	L-3,4-Dihydroxyphenylalanine
Dpg	3,5-Dihydroxyphenylglycine
D-PhgAT	D-Phenylglycine Aminotransferase
E. coli	*Escherichia Coli*
EcALDH	Aldehyde Dehydrogenase
EDTA	Ethylene Diamine Triacetic Acid
ee	Enantiomeric Excess
ESI	Electrospray Ionisation
EWG	Electron Withdrawing Group
FB_1	Fumonisin B_1
FDA	Food and Drug Administration
FDAA	1-fluoro-2, 4-dinitrophenyl-5-L-alanine amide (Marfey's reagent)
FDH	Formate Dehydrogenase
FMN	Flavin Mono-Nucleotide
FT	Fourier Transform
Fum	Fumonisin
GABA	ϒ aminobutyric Acid Aminotransferase
GDH	Glucose Dehydrogenase
GluDH	Glutamate Dehydrogenase
GRC	Gordon Research Conference
GSK	GlaxoSmithKline
h	Hour
HBF	4-Hydroxybenzoylformate

HEPES	4-(2-Hydroxyethyl) Piperazine-1-Ethanesulfonate
HFB_1	Hydrolysed Fumonisin B_1
His6	Hexahistidine Tag
HMA	L-4-Hydroxymandelate
HmaS	L-4-Hydroxymandelate Synthase
HmO	L-4-Hydroxymandelate Oxidase
Hpg	4-Hydroxyphenylglycine
HPLC	High Performance Liquid Chromatography
HPPA	4-Hydroxyphenylpyruvic Hcid
HSAN1	Human Autonomic Neuropathy Type 1
HTS	High Throughput Screening
IMAC	Immobilized Metal Affinity Chromatography
IPA	Isoproprylamine
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
IPyA	Indole Pyruvic Acid
IRED	Imine Reductase
k_{cat}	Catalytic Constant
K_d	Dissociation Constant
KDS	3-Ketodihydrosphingosine
K_m	Michaelis-Menten Constant
KPhos	Potassium Phosphate
LB	Luria Bertani
LC-MS	Liquid Chromatography – Mass Spectrometry
LDH	Lactate Dehydrogenase
LMW	Low Molecular Weight
MBA	α-Methylbenzylamine
McCAR	*Mycobacterium chlorophenolicum* CAR
MeOH	Methanol
min	Minutes
MmCAR	*Mycobacterium marinum* CAR
mp	Base-pairs

MS	Mass Spectrometry
MsCAR	*Mycobacterium smegmatis* CAR
NAD^+	Nicotinamide Adenine Dinucleotide
NADH	Nicotinamide Adenine Dinucleotide Reduced Form
$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Reduced Form
NBA	N-Butyraldehyde
NiCAR	*Nocardia iowensis* CAR
NMR	Nuclear Magnetic Resonance
NRPS	Non-Ribosomal Peptide Synthetase
OD	Optical Density
ON	Overnight
OXD	*Ortho*-Xylylene Diamine
P	Phosphate
P. putida	*Pseudomonas putida*
P. stutzeri	*Pseudomonas stutzeri*
PAGE	Polyacrylamide Gel Electrophoresis
PAP	AMP phosphotransferase
PCP	Pepidyl Carrier Protein
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
Pdh	Prephenate Dehydrogenase
PDH	Pyruvate Dehydrogenase
PEG	Polyethylene Glycol
Phg	Phenlyglycine
PL	Pyridoxal
PLP	Pyridoxal 5'-Phosphate
PM	Pyridoxamine
PMP	Pyridoxamine-5'-Phosphate
PN	Pyridoxine
PNP	Pyridoxine 5'-Phosphate

PPA	Phenylpyruvic Acid
r.m.s.	Root Mean Square
RedAm	Reductive Aminase
RNA	Ribonucleic Acid
rpm	Revolution Per Minute
SDM	Site Directed Mutagenesis
SDS	Sodium-Dodecyl Sulphate
sec	Seconds
SEC	Size Exclusion Chromatography
SOC	Super Optimal Culture
TA	Transaminase
TCA	Tricarballylic Acid
TEAA	Triethylammonium Acetate
TFA	Trifluoroacetic Acid
TR-domain	Thioester Reductase Domain
Tris	Tris(hydroxymethyl)aminomethane
UV-Vis	Ultraviolet Visible
Vf-AT	*Vibrio fluvialis*-Aminotransferase
v_{max}	Maximum Velocity
WT	Wild type

Table of contents

1 Introduction

1.1 Biocatalysis

1.1.1 A Background to biocatalysis

Biocatalysis, or the practise of using enzymes to catalyse reactions between organic components, has a long history. Indeed, the first use of enzymes can be dated back to 5000 B.C. when the ancient Egyptians used them for the preservation and production of food and alcoholic drinks.[1] At that time there was no concept of enzymes and the science behind those transformations was a mystery to our ancestors. The relationship between humans and enzymes has evolved since then and now enzymes are used to carry out a wide variety of chemical biotransformations, defining the field of biocatalysis.[2–4]

The biocatalysis field has flourished in recent years due to an array of factors. Unlike homo- and heterogeneous catalysts that rely on harsh and energy-intensive processing, enzymes are renewable, biodegradable and biocompatible. Enzymes require mild reaction conditions; a typical enzyme-catalysed procedure is usually carried out at room temperature, neutral pH and atmospheric pressure, without requirement of particular instruments (such as high pressure hydrogenation equipment).[5–7] Biocatalysts also possess chemo-, regio- and enantio-selectivity, ensuring that the desired product is produced in one step and without the need for protection and deprotection steps, common in enantio- and regioselective organic synthesis. Biocatalysis therefore has features that are very attractive in the context of green chemistry and sustainable development.[8]

Biocatalysis has reached its present relevance through waves of innovation and technology development. In 2012, Bornscheuer and colleagues described the development of modern biocatalysis over three main periods of times, named the 'three waves of biocatalysis'.[9] The first wave can be traced back to over a century ago (1910-1980) when plants, animal tissue extracts or microbial strains having the enzymatic activity of interest were applied to useful chemical transformations such as the hydroxylation of steroids using different microorganisms.[10] The second wave of biocatalysis, which began in the 1980s, saw the expanding of the number of gene sequences and the determination and submission of

high resolution protein structures to the Protein Data Bank (PDB). Initial protein engineering, typically structure based, allowed the expansion of the substrate scope of enzymes and the synthesis of unusual intermediates. The third wave of biocatalysis started in 1990s and runs until today. It is defined by the development of pioneering molecular biology techniques by Stemmer and Arnold.[11] These techniques, known as directed evolution, mimic the Darwinian process of natural selection to evolve proteins to a user defined goal. Thanks to the continued development of these techniques, the timescale for the development of enhanced biocatalysts is now measured in months rather than in years.[12] At the Biocatalysis Gordon Research Conference (GRC), held in Maine on July 2018, speculations to what will constitute the 'fourth wave of biocatalysis' have been made, suspecting that it will be focused on the development of combinations of multiple enzymes together which are linked *via* cascade processes or metabolic engineering and by combining chemo- and biocatalytic approaches.[13]

In a broad spectrum of modern industrial applications there is a continual increase in the use of biocatalytic processes, not only in pharmaceuticals but also in food, feed and fine chemical sectors.[12–14] The global enzyme market in 2018 has been estimated at $4.4bn and is shared between a wide varieties of sectors, from organic synthesis to personal care (Figure 1.1, Table 1.1).[17]

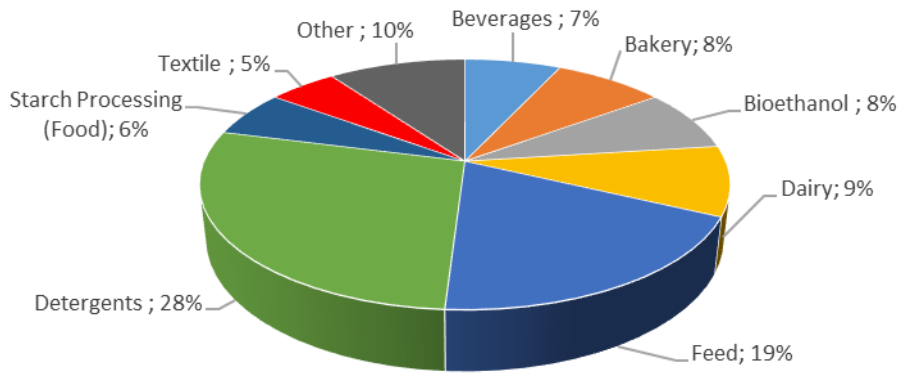
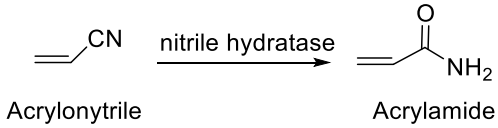


Figure 1.1 Industrial application of enzymes according to their sector.

Currently, well over 500 distinct industrial products are made using enzymes and approximately 150 industrial processes use enzymes or whole cells catalysis.[18] One such example is the manufacture of acrylamide, a key monomer for non-ionic and ionic

polyacrylamides. The chemical synthesis of acrylamide requires a copper catalyst, but this has been completely supplanted by a biocatalytic approach based on the use of a nitrile hydratase enzyme, conducted on a scale of tens of thousands of metric tons per year (Figure 1.2).[19,20] Other examples include the production of semisynthetic penicillin by employing penicillin amidases[21] or the production of high corn syrup by the action of xylose isomerase[22] (Table 1.1).



Acrylonytrile Acrylamide

Figure 1.2 Example of a bulk chemical produced by enzymatic catalysis. This process, used for the commercial production of pure acrylamide has been developed by Nitto Chemical Industry Company.

T/a	Product	Enzyme
> 1,000,000	High-fructose corn syrup	Glucose isomerase
> 100,000	Lactose-free milk	Lactase
> 10,000	Acrylamide	Nitrilase
	Cocoa butter	Lipase
> 1,000	Nicotinamide	Itrilase
	D-Pantothenic acid	Aldonolactonase
	(*S*)-Chloropropionic acid	Lipase
	6-Aminopenillanic acid	Penicillin amidase
	Aspartame	Thermolysin
	L-Aspartate	Aspartase
	D-Phenylglycine	Hydantoinase
	4-hydroxy-D-Phenylglycine	Hydantoinase
>100	Ampicillin	Penicillin amidase
	L-Methionine, L-Valine	Aminoacylase
	L-Carnitine	Dehydratase/hydroxylase
	L-DOPA	B-Tyrosinase
	L-Malic acid	Fumarase
	(*S*)-Methoxyisopropyl-amine	Lipase
	(R)-Mandelic acid	Nitrilase
	L-Alanine	L-Aspartate-β-decarboxylase

Table 1.1 Biotransformation on industrial scales. Adapted from Sánchez-Carrón.[17]

1.1.2 The biocatalysis cycle

The development of a biocatalytic process can be illustrated as the 'biocatalytic cycle' (Figure 1.3).[17,23] The starting point usually consists of the identification of a product of interest, which ideally can be produced biosynthetically using different enzymes. Such enzymes can either be commercially available/described in the literature or identified by screening for organisms (cultured or uncultured) catalysing the desired reaction. The genes encoding the desired enzymes can be codon optimised for expression in the selected host organism and purchased (e.g. GenScript). Once one or more biocatalysts are obtained, a kinetic characterisation is carried out and structural information are acquired by solving of the 3D structure or by homology modelling.

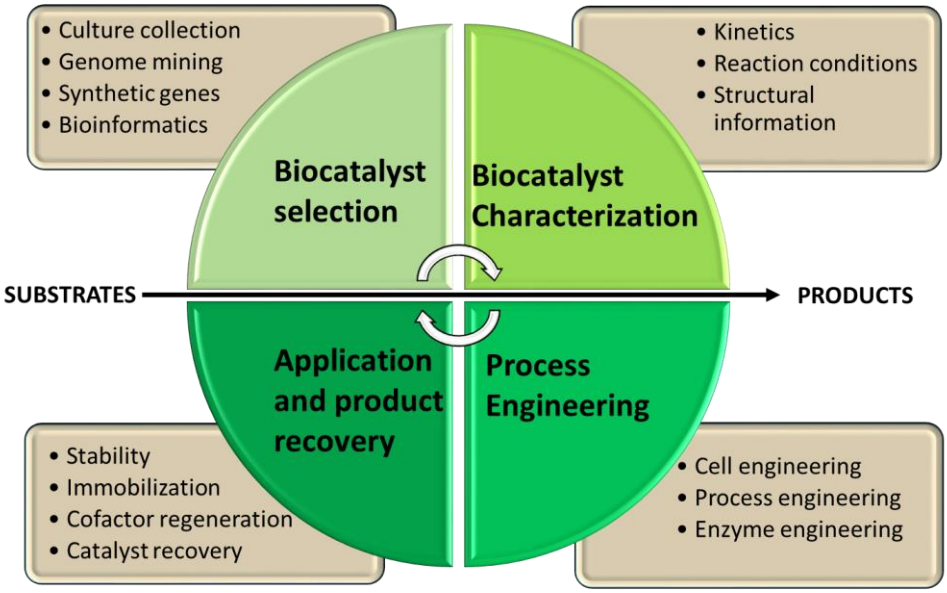


Figure 1.3 The biocatalysis cycle. Adapted from Sánchez-Carrón and Schmid *et al.*[17,23]

The biocatalyst and the biocatalytic process are then engineered for optimal performances at the level of the enzyme and/or the process. Commonly, iterative cycles of protein engineering are required to improve a specific enzyme to the desired standards (activity, substrate range, etc.). Once the biocatalyst has the desired characteristic other issues such as stability, cofactor recycling and/or biocatalyst recovery can be addressed. The reaction medium, consisting of an aqueous, organic or two-liquid phase system is usually optimised to dissolve the components of the reaction while retaining enzyme activity. Many biocatalysts are used in immobilized form as it eliminates the drawbacks associated with

substrate and product inhibition and allows a very easy product separation.[24,25] If immobilised, the enzyme can be re-used multiple times, drastically reducing the reaction costs. This biocatalytic cycle can be iterated multiple times to reach an efficient industrial process that meets the requirements in terms of catalyst stability, conversion and cost-efficiency.[17,23]

1.1.3 Case study - sitagliptin manufacture: a tale of biocatalysis

For a given biotransformation, the performances of the wild type (WT) enzymes with the target compounds under the desired reaction conditions are often not satisfactory and enzyme evolution is frequently necessary. Indeed, in several cases, the WT enzyme does not show any activity towards the desired substrate whatsoever and several rounds of evolution are necessary to achieve conversions up to industrial standards.[26]

A spectacular example of a biocatalyst redesigned by directed evolution for the production of the active pharmaceutical ingredient (API) is the ATA-117 for Sitagliptin manufacture. Sitagliptin is the active ingredient of Januvia™, a top selling drug for type II diabetes developed by Merck and synthesised by high pressure rhodium mediated asymmetric hydrogenation.[27] The Rh mediated process was highly stereospecific (> 97% enantiomeric excess (*ee*)) but the residual catalyst required removal upon carbon treatment and the overall process was not 'green'.

Merck, together with the leading biocatalysis company Codexis, investigated the use of an aminotransferase (AT) for the conversion of the key pro-sitagliptin ketone into the sitagliptin amine. An (*R*)-selective AT (ATA-117) from *Arthrobacter sp.* was selected for enzyme evolution using a combined *in silico* design and directed evolution approach. Several rounds of mutagenesis were required to achieve a satisfying conversion of the bulky pro-sitagliptin ketone under the desired reaction conditions (40,000-fold activity increase in 50% DMSO, 40 °C). The final variant, which held 27 mutations, was able to convert 200 g L^{-1} prositagliptin with 99.95% *ee* at 92% yield (Figure 1.4).[28] These 27 mutations are found in several positions, not only in the active site but also at the enzyme dimer interface which most likely stabilise the dimeric form of the enzyme under conditions (such as 50% DMSO) that would normally cause enzyme denaturation.[27,28]

The biocatalytic approach has several benefits compared to the chemocatalytic process, such as a productivity increase of 50%, waste reduction of 19%, heavy metal elimination and overall cost reduction. Furthermore, no high-pressure hydrogenation equipment was needed. This 2010 study still represents one of the flagship applications of an industrial biocatalyst today, highlighting the advantages of adopting the biocatalytic strategy over the chemical counterpart and was recognized with the 2010 Presidential Greener Reaction Conditions Award from the United States Environmental Protection Agency.

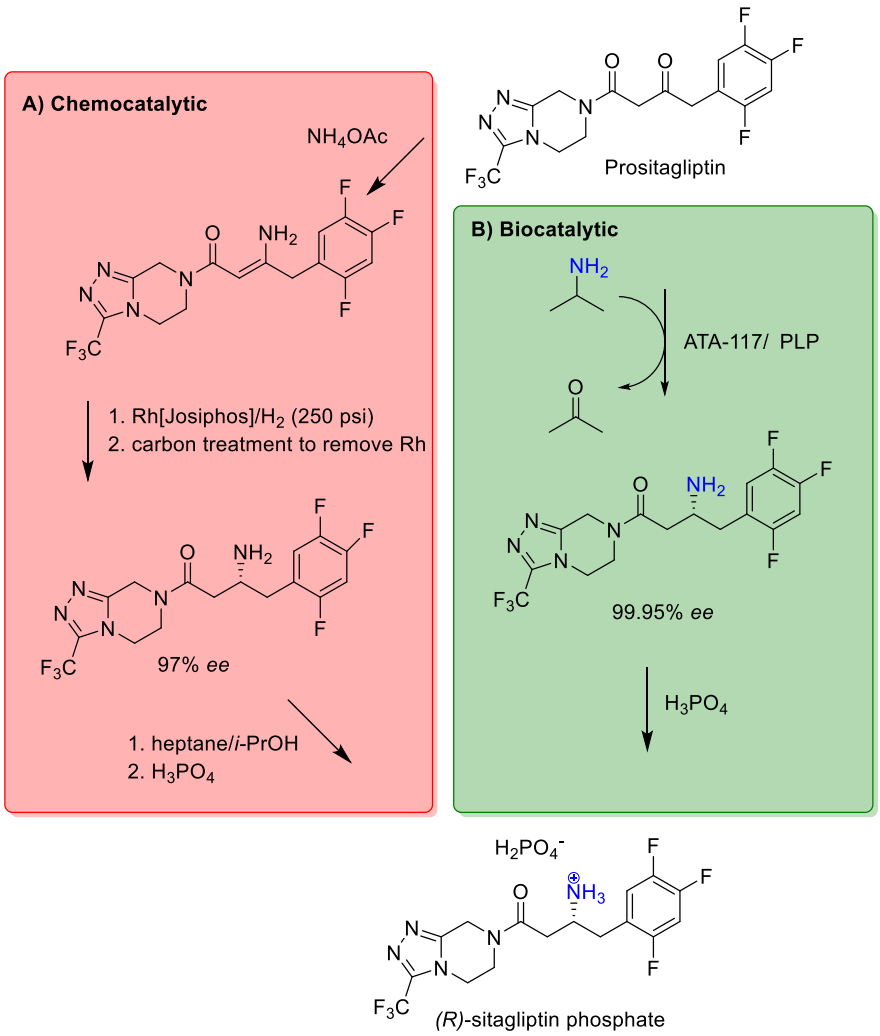


Figure 1.4 Comparison between the **A)** chemocatalytic and **B)** biocatalytic route for sitagliptin manufacturing. The prositagliptin ketone is converted into the Sitagliptin amine in high yield and *ee* Adapted from Savile *et al.*[27]

1.1.4 Limitations of biocatalysis

Despite the advantages of biocatalysis and its increased application in many sectors, there are still a number of drawbacks associated with biocatalysis as a generally applicable technique. Problems such as limited stability of the biocatalysts under operative conditions, an insufficient number of well-characterised, ready-to-use enzymes and high development costs have to be addressed.[29,30] Due to the existence of well-developed chemical tools, the improvement and optimisation of traditional chemical transformations are generally still faster when compared to the biocatalytic approach. There is still a lack of optimal, off-the-shelf biocatalysts that can be used for multiple biotransformations and protein engineering is usually necessary to ensure an enzyme works for a given substrate and in the desired reaction conditions (e.g. high solvent, high temperature). Thus, for a biocatalytic process to be established it should bring significant improvements and typically represents a second-generation method, supplanting a pre-existing chemical process.

For some reactions, such as the enzymatic reduction of ketones, biocatalytic processes are competitive with their chemical counterparts. However, for many others, the chemical approach is still preferred. Despite this, the use of enzymes is broadening and is already quite prominent in the production of fine-chemicals and pharmaceuticals (Table 1.1). Opportunities are available for combining chemo- and bio-catalysis and overcoming some of the limits associated with both approaches.

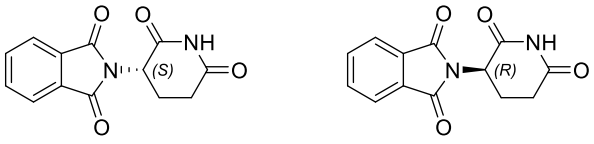
Currently, a growing number of companies see biocatalysis as an appealing option and are investing in developing a core suite of biocatalytic tools to optimise production schemes. In particular, enzyme panels from different suppliers (e.g. Prozomix, Codexis) are increasingly becoming available to allow for screening of interesting enzyme activities, which could be useful for the synthesis of APIs. Biocatalysis is poised to play an ever-increasing role in the manufacture of chemicals, and in the next decade more processes are expected to adopt it. This paradigm shift will be facilitated by high throughput screening (HTS), which will be crucial for testing large libraries of mutant enzymes in short time frames. Additionally, bioinformatic software that will accurately predict hot spots for mutations will reduce the time required to obtain new biocatalysts and improved variants.[31]

1.2 The importance of chirality

1.2.1 Chiral drugs

Biocatalysis is gaining increased attention for its ability to perform stereoselective transformations in an environmentally friendly way.[32] The creation of a stereogenic centre *via* organic synthesis is not trivial, and multiple approaches have been described based on organometallic and organocatalytic methods.[33,34] In 2006, 90% of the Food and Drug Administration (FDA) approved drugs were chiral compounds and thus the development of novel, green approaches to the synthesis of chiral centres is of paramount importance.[35] As such, the intrinsic chirality of enzymes makes them a powerful tool for the development of stereoselective transformations.

Drug isomerism is a crucial consideration in drug development. This is because although isomers have the same chemical structure, most drug isomers exhibit differences with respect to the pharmacology, toxicology, pharmacokinetics, and metabolism.[36] This is exemplified by the tragic history of thalidomide (thalidomid, Contergan), which shows why chiral separation and enantioselectivity is crucial in drug development chemistry.[37] Thalidomide was a racemic sedative launched in 1957 and prescribed to pregnant women to alleviate morning sickness. Unfortunately fetuses exposed to thalidomide were born with catastrophic limb malformation (phocomelia). This was compounded by a mortality rate of 50%. These horrific side effects were found to be due to one of the two enantiomers of thalidomide, which has a stereogenic carbon atom on the glutarimide ring adjacent to succinamide ring. The (*S*)-enantiomer of the drug has the teratogenicity effect while the (*R*)-enantiomer has the pharmacologic effect (Figure 1.5).[37]



(S)-thalidomide *(R)*-thalidomide

Figure 1.5 The two enantiomers of thalidomide, (*S*)-thalidomide was found to cause malformations in embryos while (*R*)-thalidomide relieved the symptoms of morning sickness.

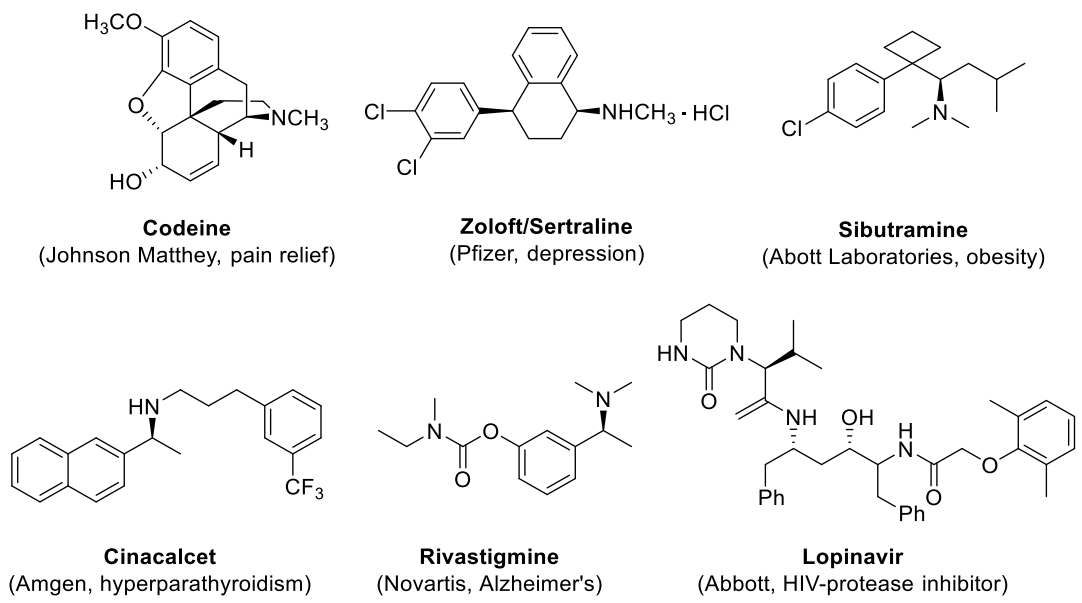
This painful episode illustrates the importance of chiral separation and enantioselective production of drugs. As a consequence, in 1992, the FDA required that if approval for a

racemic drug is being sought, then both enantiomers must be tested for their pharmacological effects. Following this, in 1997 a fast track program was introduced by the FDA, which allowed faster approval of a single enantiomer. This translated into the general decision for synthesising and marketing single enantiomer drugs. Additionally, a lower dosage is required to achieve the desired effects, and single enantiomer drugs present better pharmacological profiles in comparison to their racemates.[36] The most pertinent example of this is the pain relief drug ibuprofen, which was switched to a single-enantiomer because of the increased effectiveness of the (*S*) enantiomer, which is over 100-fold more potent as an inhibitor of cyclo-oxygenase 1 than the (*R*)-ibuprofen enantiomer. Ibuprofen was followed by ketoprofen and fifteen other drugs in the period between 1994 and 2011 in the so called 'chiral switching'. Thus the demand for enantiopure drugs is continuously increasing, and the development of methods for their production is vital.[38]

1.2.2 Chiral amines

Optically active amines and amino acids are ubiquitously distributed in nature where they play a central role in living systems. L-amino acids, for instance, are the building blocks of proteins, function as precursors of DNA and RNA bases, and are involved in the biosynthesis of all the fundamental biological macromolecules.[39] They are also involved in secondary metabolism and signaling pathways. Chiral amines are highly biologically active; this is thought to be due to the density of structural information they carry and their ability to hydrogen bond with targets.[39] Thus, it is not surprising that chiral amines are key scaffolds for the synthesis of natural products, pharmaceuticals and agrochemicals. It was recently estimated that currently around 40% of pharmaceuticals and 20% of agrochemicals contain a chiral amine moiety in their structure.[40] Pharmaceutical compounds containing a chiral amine moiety have a broad range of applications, treating conditions such as pain, depression, HIV, obesity and Alzheimer's (Figure 1.6).[41]

However, their application is not limited to the pharmaceutical and agrochemical sectors, they also find application as chiral resolving agents and chiral auxiliaries. Thus, due to their crucial roles, the demand for chiral amines is steadily increasing and high chemical and enantiomeric purity (>99%) are usually required, but their synthesis with chemo-, regio- , diastero- and enantiocontrol is still challenging.



Codeine
(Johnson Matthey, pain relief)

Zoloft/Sertraline
(Pfizer, depression)

Sibutramine
(Abott Laboratories, obesity)

Cinacalcet
(Amgen, hyperparathyroidism)

Rivastigmine
(Novartis, Alzheimer's)

Lopinavir
(Abbott, HIV-protease inhibitor)

Figure 1.6 Example chiral amine pharmaceutical drugs. These find application to treat several kinds of condition such as Alzheimer's, depression, obesity and HIV.

1.2.2.1 Organic Synthesis of chiral amines

The importance of chiral amines and the strong need for enantiomerically pure substances has led to the development of a pool of methodologies for their synthesis. Two main chemical approaches can be listed for the preparation of optically pure chiral amines: (**1**) resolution of racemates and (**2**) asymmetric synthesis. The resolution technique is the 'classical' route to an enantiopure product. It differs from the asymmetric synthesis as it separates two enantiomers rather than aiming to produce a single enantiomer from the beginning. Although the major drawback is a maximum yield of 50%, classical resolution *via* the formation of diastereomeric salt pairs and precipitation of the least soluble salt is still used industrially to produce optically active pure amines as it is relatively inexpensive and can be easily scaled-up.[42] In this approach a racemate interacts with an enantiopure compound to form a diastereomeric salt which, unlike the enantiomer crystals, can be easily separated while the other enantiomer is generally discarded. This leads to poor atom economy and so asymmetric synthesis is usually preferred as it can yield 100% of the desired product. This can be generally achieved by imine and enamide reduction (Figure 1.7).[39]

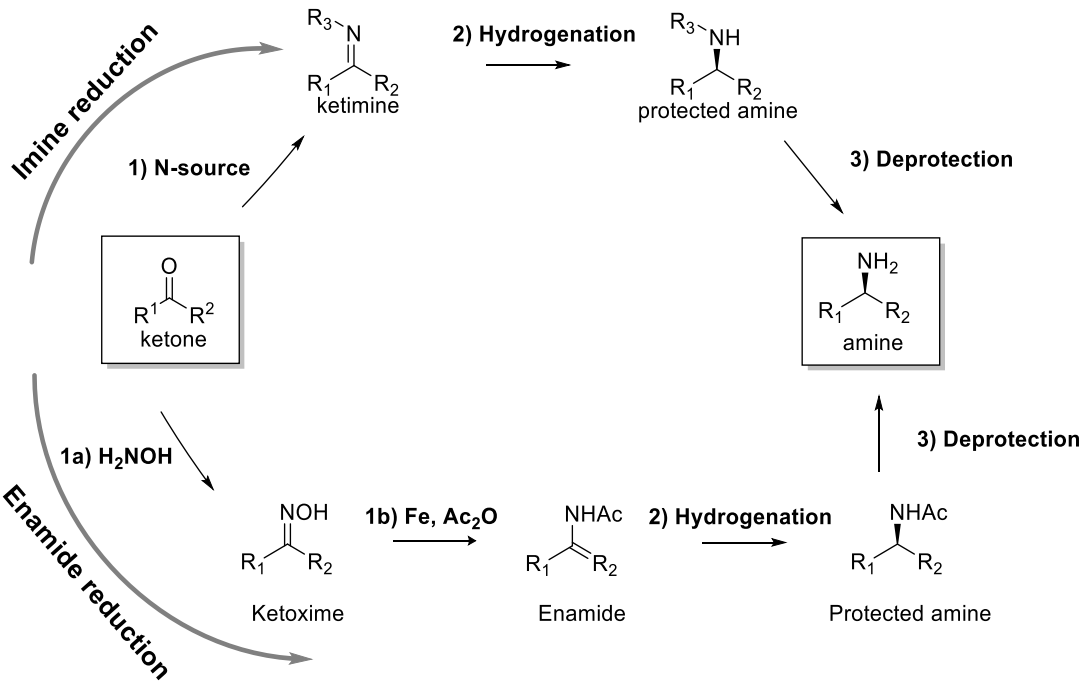


Figure 1.7 Most used chemical approaches for the synthesis of chiral amines: imine and enamide reduction.

The imine reduction process requires multiple steps: (1) condensation with an N-source to generate the imine, (2) reduction to generate an amine and (3) cleavage of the auxiliary group to release the free amine. A similar alternative approach is enamide reduction. However, both these methods require the use of expensive metal catalysts and the removal of auxiliary groups to give the free amine can be problematic.

1.2.2.2 Biocatalytic route to non-racemic chiral amines

Several biocatalytic approaches are available for the production of enantiopure amines. As for the chemical approach, two main strategies are commonly used: (**1**) enzymatic kinetic resolution of racemates and (**2**) enzymatic asymmetric synthesis from prochiral substrates (Figure 1.8).[43] In the enzymatic kinetic resolution approach, two enantiomers react with the biocatalysts with different reaction rates, yielding an enantio-enriched sample of the less reactive enantiomer (**1**) (Figure 1.8A). The main limitation of this process is a maximum theoretical yield of 50%. In order to overcome this technical limitation and achieve a yield of 100%, the addition of an extra enzyme, a racemase, is necessary to convert the unwanted enantiomer and facilitate a dynamic kinetic resolution (DKR) (Figure 1.8A). Using

this approach, several enzymes such as lipases, hydantoinases and amidases have been successfully employed for the production of a wide range of chiral amines. Alternatively, enzymatic asymmetric synthesis can be employed (**2**), allowing the direct conversion of a prochiral substrate into the desired single enantiomer in one step (Figure 1.8B). This provides the greatest atom economy and higher efficiency than the alternative DKR approach as no requirement of additional enzyme is necessary to achieve a potential yield of 100%.

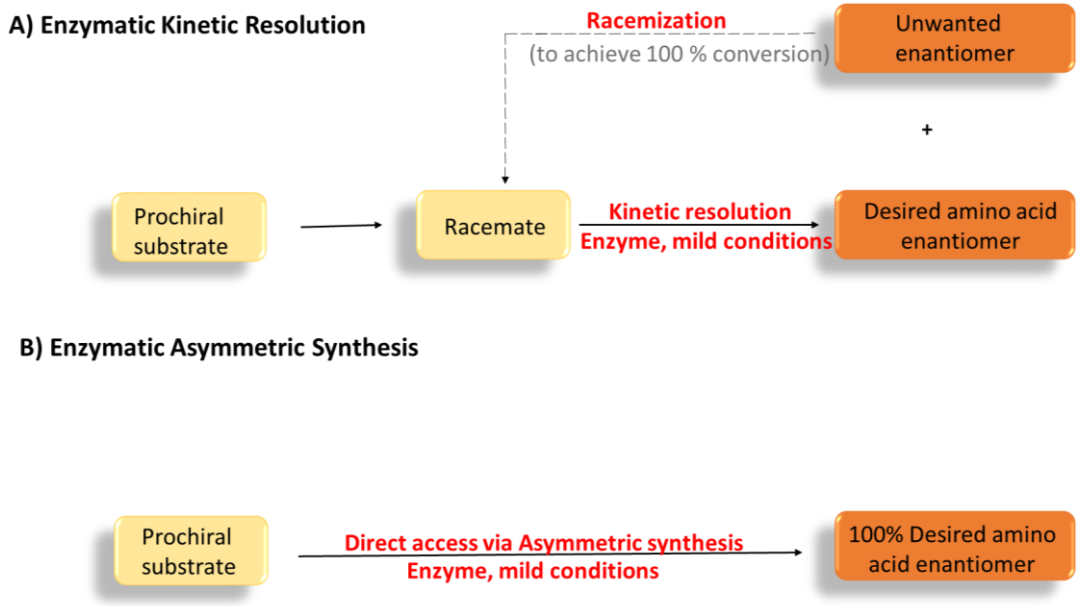


Figure 1.8 A) Enzymatic kinetic resolution and **B)** enzymatic asymmetric synthesis approaches for the production of chiral amines. Adapted from Xue *et al.* [43]

Nature has evolved a wide variety of enzymes which can be used to produce enantiopure chiral amines from various starting materials using both approaches outlined above. These include lipases (**1**), amine oxidases (AOs) (**2**), amine dehydrogenases (AmDHs) (**3**), imine reductases (IReds) (**4**), reductive aminases (RedAms) (**5**) and aminotransferases (ATs or TAs) (**6**) (Figure 1.9).[41] Amongst the principal methods for the enzymatic synthesis of chiral amines, depicted in Figure 1.9, the most widely applied strategy is the resolution of racemic starting material (low cost, produced from inexpensive bulk chemicals) employing hydrolytic enzymes (*e.g.* lipases, 1).[41] However, in the past decade ATs have started to become the elected biocatalyst for the production of primary amines.

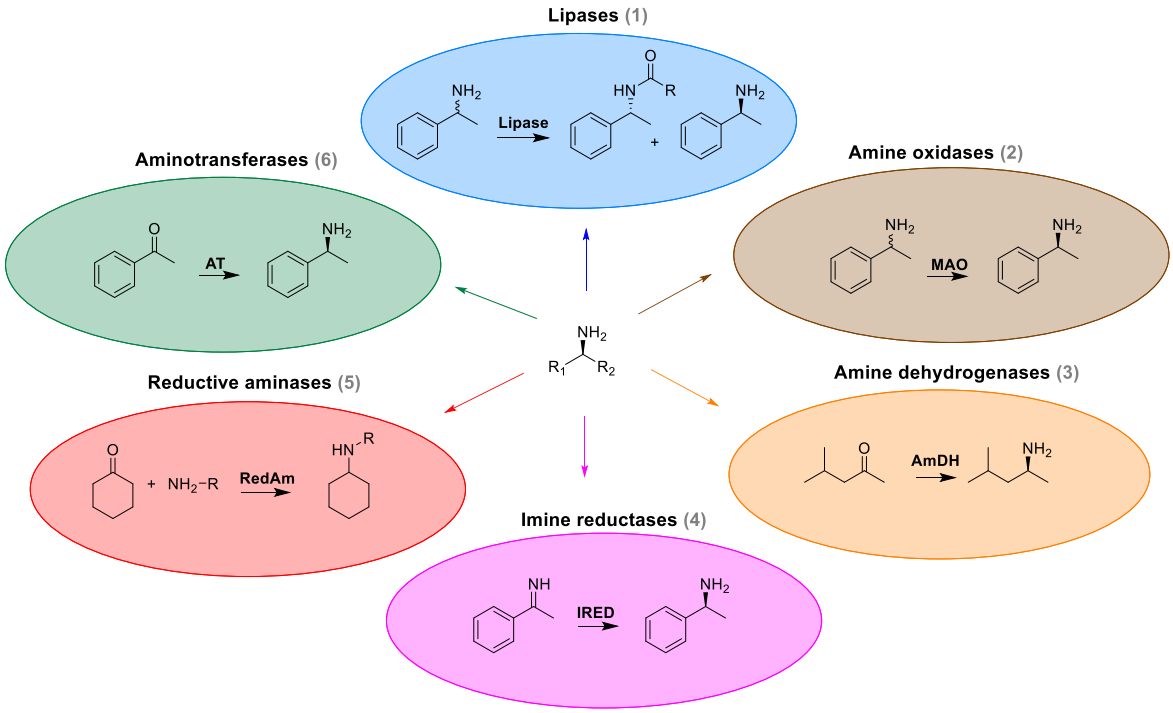


Figure 1.9 Different biocatalyst that catalyse the production of amines: lipases (1), amine oxidases (AOs) (2), amine dehydrogenases (AmDHs) (3), imine reductases (IReds) (4), reductive aminases (RedAms) (5) and aminotransferases (ATs or TAs) (6).

Lipases (**1**)

Lipases belong to the hydrolase family that catalyse the formation or cleavage of esters and amides. Lipases have found their application in industry in the production of optically pure aliphatic amines, benzyl amines and amino alcohols. They are widely employed for their flexibility as they are not restricted to the 20 proteogenic amino acids, but they also accept non-natural analogues. They are the first enzymes to have been applied on an industrial scale for the enantioselective acylation of racemic amines.[44,45]

Amine oxidases (AOs) (**2**)

The AOs catalyse the oxidation of amines to imines with simultaneous reduction of oxygen to hydrogen peroxide. They are divided in two classes according to the cofactor used: type I AOs are copper dependent whilst type II AOs are flavin dependent. The AOs represent a valuable alternative to lipase mediated kinetic resolution and de-racemization of racemic amines.[46,47]

Amine dehydrogenases (AmDHs) (**3**)

The ADHs catalyse the NAD(P)H dependent reversible asymmetric reductive amination of ketones with ammonia to yield amines.[48] Even though the AmDH reaction is reversible, the equilibrium reaction is in favour of amination ($10^{14}<K_{eq}< 10^{18}$). The main drawback in utilising AmDHs is the requirement of expensive cofactor such as NADH and NADPH. For industrial applications, *in situ* cofactor regeneration is essential for the production of chiral amines, allowing their use in catalytic amounts. For the cofactor recycling system, formate dehydrogenase (FDH) is often used, which catalyses the oxidation of formate to carbon dioxide with the concurrent reduction of NAD^+ to NADH (Figure 1.10A). Alternatively, glucose dehydrogenase (GDH) is also commonly used for its ability to accept both NADH and NADPH substrates [49] (Figure 1.10B).

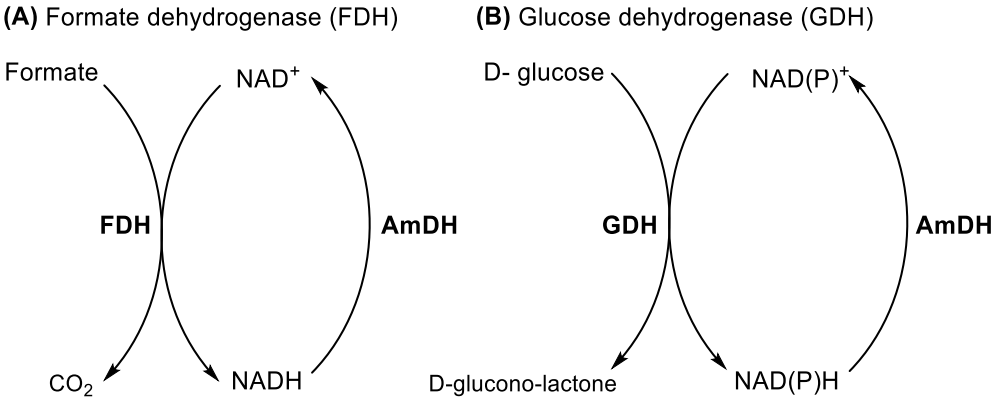


Figure 1.10 Reaction scheme for NADH and NADPH recycling systems employing **(A)** formate dehydrogenase (FDH) **(B)** glucose dehydrogenase (GDH).

Imine reductases (IREDs) (**4**) and *reductive aminases* (RedAms) (**5**)

The IREDs are a recently discovered class of NAD(P)H dependent enzymes that catalyse the asymmetric reduction of imines to amines *via* oxidoreduction. A breakthrough in their application was the identification of an IRED homologue from *Aspergillus oryzae* by Turner and colleagues displaying remarkable activity for reductive amination (RedAm).[50,51] This RedAm is able to accept a wide range of carbonyl acceptors and amine nucleophiles and hints at the potential application of these IRED homologues for the reductive amination of a wide range of substrates such as aliphatic, cyclic and aromatic ketones. Since they require NADH and/or NADPH, recycling systems would be necessary for their industrial application.

Aminotransferases (6) (ATs or TAs)

The ATs are pyridoxal 5'-phosphate (PLP) dependent enzymes that catalyse the reversible transfer of an amine group to a carbonyl containing compound (amino acceptor) such as ketones, aldehydes and ketoacids.[52] Unlike other biocatalysts used for the preparation of chiral amines, TAs have the main advantage of not relying on a costly cofactor that requires *in situ* regeneration. TAs can be used for the asymmetric synthesis of amines from the corresponding ketone or kinetic resolution of racemic amines. The first process is usually preferred as it offers the potential of achieving 100% yield. However, problems such as unfavourable reaction equilibrium, and substrate and/or (co)product inhibition have hampered the widespread use of ATs in industry.[53]

1.3 Aminotransferases, a class of PLP-dependent enzymes

1.3.1 Vitamin B6 vitamers

Vitamin B_6 is one of nature's most versatile organic enzyme cofactors. Six biologically interconvertible compounds constitute the vitamin B_6 group: pyridoxine (PN), pyridoxamine (PM), pyridoxal (PL) and the phosphorylated forms pyridoxine 5'-phosphate (PNP), pyridoxamine 5'-phosphate (PMP) and pyridoxal 5'-phosphate (PLP).[54,55,56] These differ from each other in the functional group present at the 4' position of the pyridine ring (Figure 1.11).[57,58] Amongst them, PLP is the biologically active form of vitamin B_6, which catalyses over 100 reactions inside the human body, including the production and degradation of several neurotransmitters.[59–61]

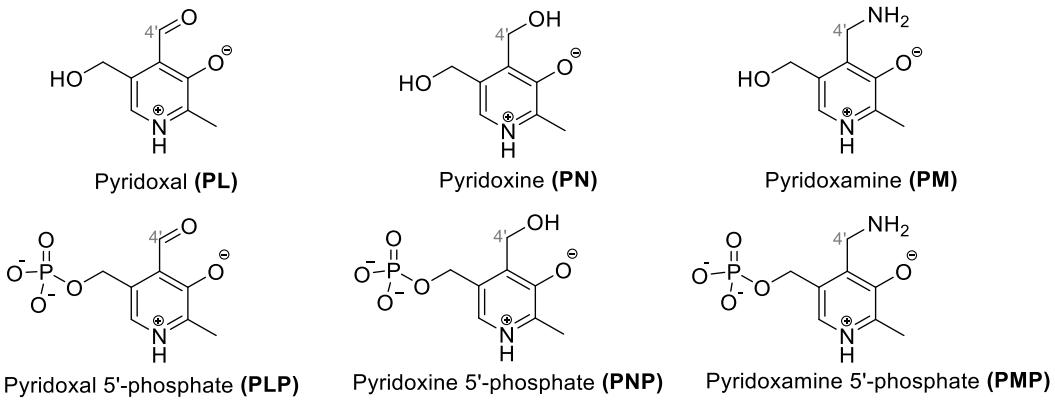


Figure 1.11 Structures of the six vitamin B6 derivatives and their phosphorylated forms.

PLP dependent enzymes are principally involved in the biosynthesis of amino acids and amino acid derivatives, but also in the biosynthesis of amino sugars and other amine-containing compounds.[57,59] As they are widespread in cellular processes, many of them have been identified as drug targets. Indeed, ϒ aminobutyric acid AT is the target of inhibitors to treat epilepsy, serine hydroxyl methyltransferase is a cancer target and L-3,4-dihydroxyphenylalanine carboxylase is a target for the treatment of Parkinson's disease.[61,62]

Genetic defects affecting PLP dependent enzymes are implicated in a variety of diseases such as primary hyperoxaluria type I and autonomic neuropathy type 1 (HSAN1) and inadequate levels of PLP in the brain have been shown to causes neurological dysfunctions like epilepsy.[63,64] The versatility of the PLP as cofactor, derived from its excellent electron sink properties, has also inspired the development of biomimetic catalysts such as N-quaternized catalyst for the production of amines.[65]

1.3.2 The chemistry of PLP

PLP-dependent enzymes are found in five of the six classes defined by the Enzyme Commission (EC), and catalyse a multitude of different reaction representing 4% of all enzyme activities.[66] Despite the variety of PLP-catalysed reactions, the chemical mechanism is broadly consistent. PLP acts upon amino containing substrates (mostly amino acids) and its aldehyde group is invariably bound to the ε- amine group of an active site lysine residue through a Schiff base linkage, forming the internal aldimine species.[67,68] This represents the *holo*-form of the enzyme while the lack of the cofactor is referred to as the *apo*-form.[69] The internal aldimine can be distributed between two tautomeric forms, the ketoenamine form when the imine nitrogen is protonated (400-420 nm) and the enolimine form when the alcohol oxygen is protonated (320-340 nm) (Figure 1.12).[70]

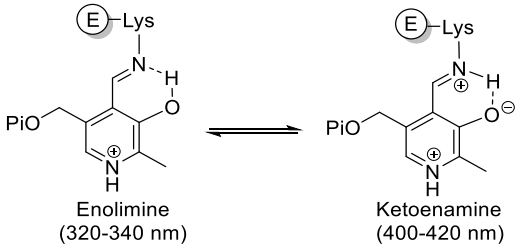


Figure 1.12 Enolimine and ketoenamine equilibrium.

The ketoenamine and enolimine distribution is characteristic of each PLP dependent enzyme and depends on the active site environment. Upon addition of an amino acid, the lysine residue is exchanged for the amino group of the substrate *via* a transaldimination reaction, forming a non-covalently bound external aldimine. This causes a change in the ketoenamine-enolimine equilibrium. These changes in the UV-Vis spectrum can be used to monitor external aldimine formation and offer a window into the biochemistry of PLP dependent enzymes.[71]

PLP's reaction versatility arises from this common external aldimine intermediate as is followed by cleavage of one of the bonds adjacent to the C_α atom (α, β or γ). PLP catalysed reactions can therefore be subgrouped according the position at which the net reaction occurs (the α, β or ϒ carbon of the amino acid substrate) (**Figure 1.13**). Reactions at the α-position include transamination, decarboxylation, racemization, elimination and replacement of an electrophilic R group. Reactions at the β or ϒ position include elimination or replacement.[72]

The precision architecture of the enzyme active site and organisation of the PLP cofactor allow only one reaction to occur and prevent all other possible reactions. This is described by the Dunathan hypothesis (Figure 1.14).[73] In 1966, Dunathan proposed that the specific reaction of a PLP-dependent enzyme is controlled by the orientation of the substrate. The bond (α, β or ϒ) to be broken will be the one perpendicular to the plane of the conjugated π system of the PLP ring (Figure 1.14A and B).[73] The only exception is represented by the PLP-dependent glycogen phosphorylases, which use a different type of PLP chemistry, utilising the phosphate group of the PLP cofactor for catalysis.[74,75] The bond breaking steps are facilitated by the formation of a carbanion that is stabilised by the electron sink property of the pyridine ring of the PLP cofactor (Figure 1.14C).[76,77] The stabilised carbanion is described as a quinonoid as its structure resembles a quinone and can be monitored spectrophotometrically as it has a strong absorption at 500 nm.[78,79]

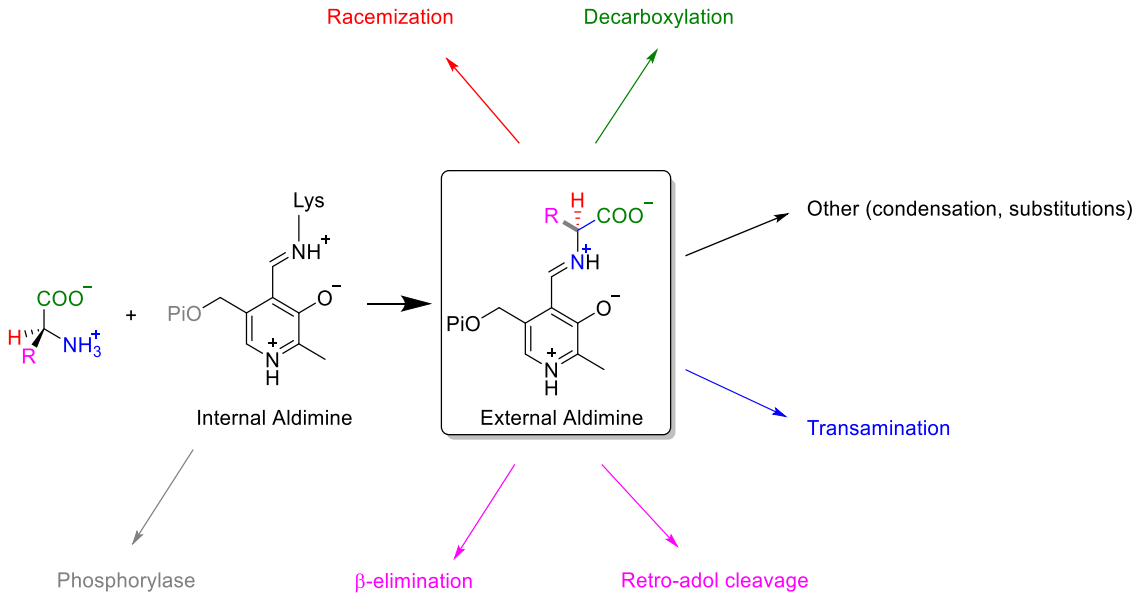


Figure 1.13 PLP dependent enzymes catalyse a variety of chemical reactions after the substrate forms a covalent aldimine intermediate with PLP. Different bonds can be broken (shown in different colours) leading to distinct enzymatic activities.

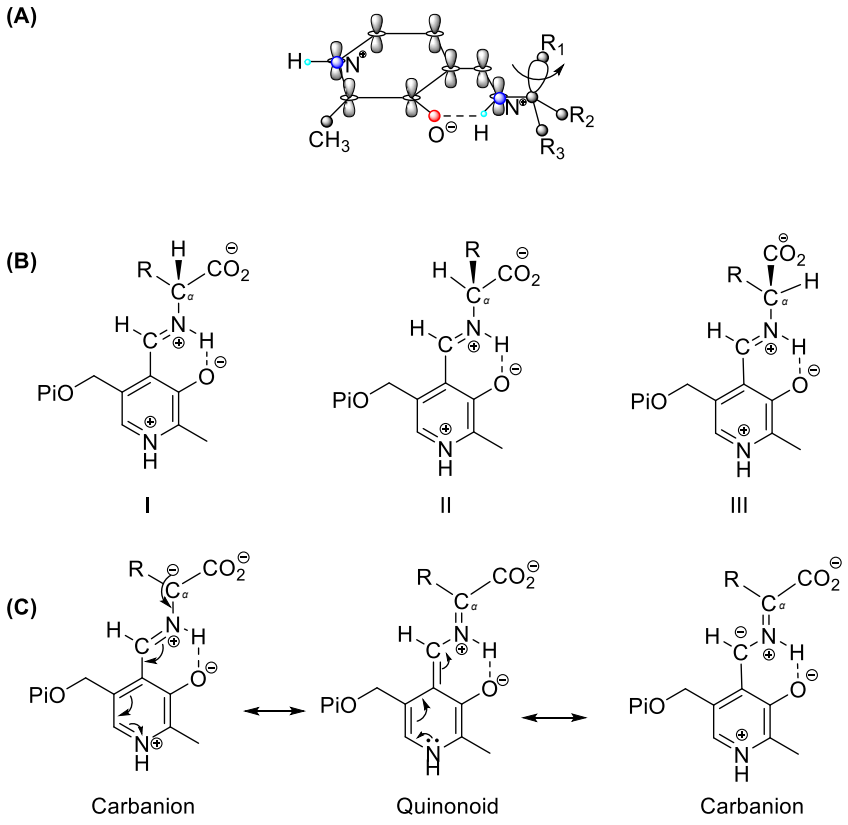


Figure 1.14 Illustration of Dunathan hypothesis. **(A)** Stereoelectronic control of reaction specificity in PLP-catalysed reactions. **(B)** Possible conformation of the external aldimine after cleavage of the three possible bonds adjacent to the Cα. **(C)** Carbanion and quinonoid resonance forms obtained after the external aldimine deprotonation

1.3.3 PLP enzyme classification

Despite the variety of the PLP-catalysed reactions, all structurally characterised PLP-dependent enzymes have been classified into seven structural groups, presumably corresponding to independent evolutionary lineages.[59,60] The capability of proteins with similar structures to catalyse such a diverse range of chemical reactions represents a remarkable example of divergent evolution.[59]

Several classification schemes have been proposed for PLP enzymes.[80] However, the commonly accepted classification scheme is the one proposed by Goldsmith[81] and recently redefined by Peracchi *et al.*[60], in which PLP-enzymes are divided into five different fold types based on sequence similarity and 3D structures (Figure 1.15).[55] Individual folds are named according to the enzyme first identified with such a fold and consist of: the aspartate aminotransferases (AspAT, fold type I), the tryptophan β-synthase (fold type II), the alanine racemase (fold type III), the D-amino acid aminotransferase (fold type IV) and the glycogen phosphorylase (fold type V) families. Two additional clusters have subsequently been added and include the lysine 5, 6 aminomutase (fold type VI) and the lysine 2, 3 aminomutase (VII) familes, whose structures have been recently determined and found to differ from other fold types (Figure 1.15). Amongst the PLP superfamily, the vast majority (approximately two-thirds) of enzymes belong to fold type I, followed by group II and group IV. Group V PLP enzymes are rare whilst groups VI and VII are almost absent.

The aspartate aminotransferase family (fold type I)

The group I PLP fold, also known as the AspAT family, is structurally the most well characterised fold in the PLP superfamily. Enzymes in this class are catalytically active as homodimers, where each subunit consists of two domains: a large domain, in which the central feature is a seven stranded β sheet and a small domain consisting of a three-stranded β sheet packed against helices on one side. The PLP cofactor is covalently attached to the large domain *via* a lysine residue at the N-terminus, with the active site located in a cleft between the two domains. Residues from both domains/subunits are involved in cofactor binding and stabilisation. Structural analysis of the members of this fold has revealed differences between the N-terminal domains, dividing the fold type I into

six subclasses of which three are aminotransferases subclasses as defined by Grishin *et al.*[81] and Metha *et al.* and will be discussed in detail later.[82]

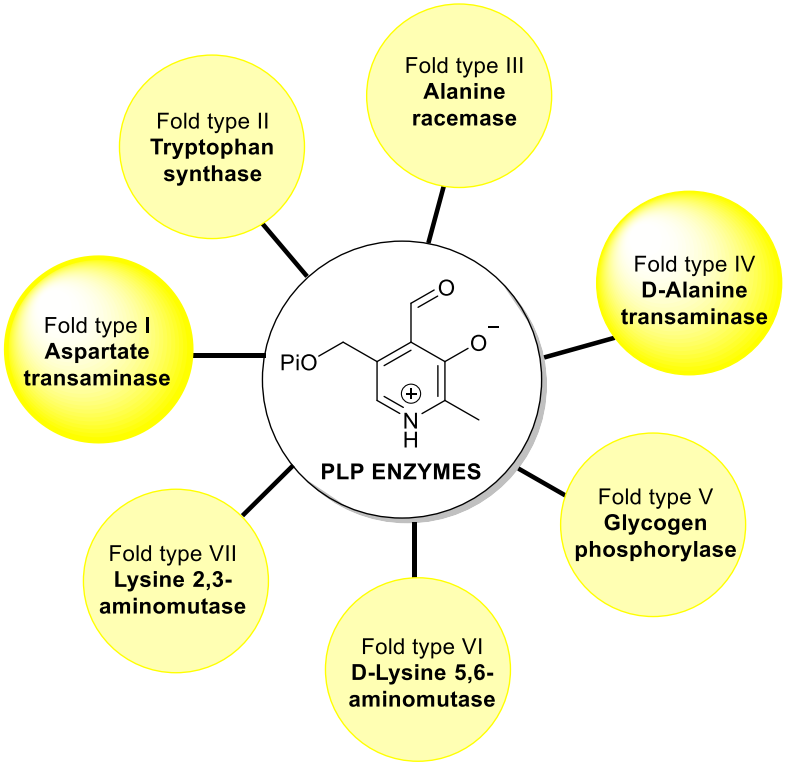


Figure 1.15 Classification of PLP dependent enzymes based on fold types. The fold type I and IV include transaminases and are highlighted in bright yellow. Adapted from Slabu *et al.*[83]

The tryptophan synthase β family (fold type II)

The tryptophan synthase family includes the tryptophan synthase β subunit, threonine deaminase and *O*-acetylserine sulfhydrilase. The catalytic unit of these type II PLP enzymes is composed entirely of residues from one monomer. Each monomer consists of a PLP-binding site in the N-terminal domain which forms the catalytic part, and a regulatory domain situated at the C-terminus. However the functional form is either homodimeric or a higher-order oligomer.

The alanine racemase family (fold type III)

Fold type III, known as the alanine racemase superfamily, includes amino acid racemases and decarboxylases. Enzymes of this group are homodimers with each monomer containing two domains, one with an α/β barrel and the other mainly comprising β strands.

Interestingly the mode of binding of PLP is similar to other fold types, with the phosphate group anchored at the N-terminus of an α-helix and residues from both monomers contributing to the active site formation.

The D-amino acid aminotransferase family (fold type IV)

Not only does the fold type IV family include D-amino acid ATs, but also lyases. Fold type IV enzymes are superficially similar to fold type I, as they function as homodimers and the catalytic unit of each monomer is composed of a small and large domain. In addition, the active sites of D-amino acid transferases are mirror images of the active sites of the fold type I aminotransferases such that the *re*-face rather than the *si*-face of the PLP cofactor is solvent exposed (where the *re*- and *si*-faces are defined in relation to the C-4' of the cofactor as the centre of the plane).[84]

Glycogen phosphorylase family (fold type V)

Fold type V PLP enzymes are multidomain proteins which contain three distinctive domains; the C- and N-terminal domains and a glycogen-binding domain. The C-terminal domain binds PLP and also has a dinucleotide-binding fold. Type V enzymes are strikingly different from the other PLP families as they utilise the phosphate group of PLP for catalysis.

Lysine 5, 6 aminomutase (fold type VI)

The overall structure of type VI PLP enzymes is an $\alpha_2\beta_2$ tetramer. In this structure, a Rossmann domain covalently binds the PLP cofactor while simultaneously positioning the other cofactor, adenosylcobalamin (coenzyme B_{12}) far from the active site in order to prevent radical generation in absence of substrate.[85]

Lysine 2, 3 aminomutase (fold type VII)

Lysine 2, 3 aminomutase is a homotetramer consisting of a dimer of tightly associated dimers. The subunits of the dimers are linked through a zinc ion which coordinates residues from both subunits, with the PLP located at the N-terminus of the channel. Each subunit consists of three domains, the C-terminal domain, the central globular domain and the N-terminal domain.[86]

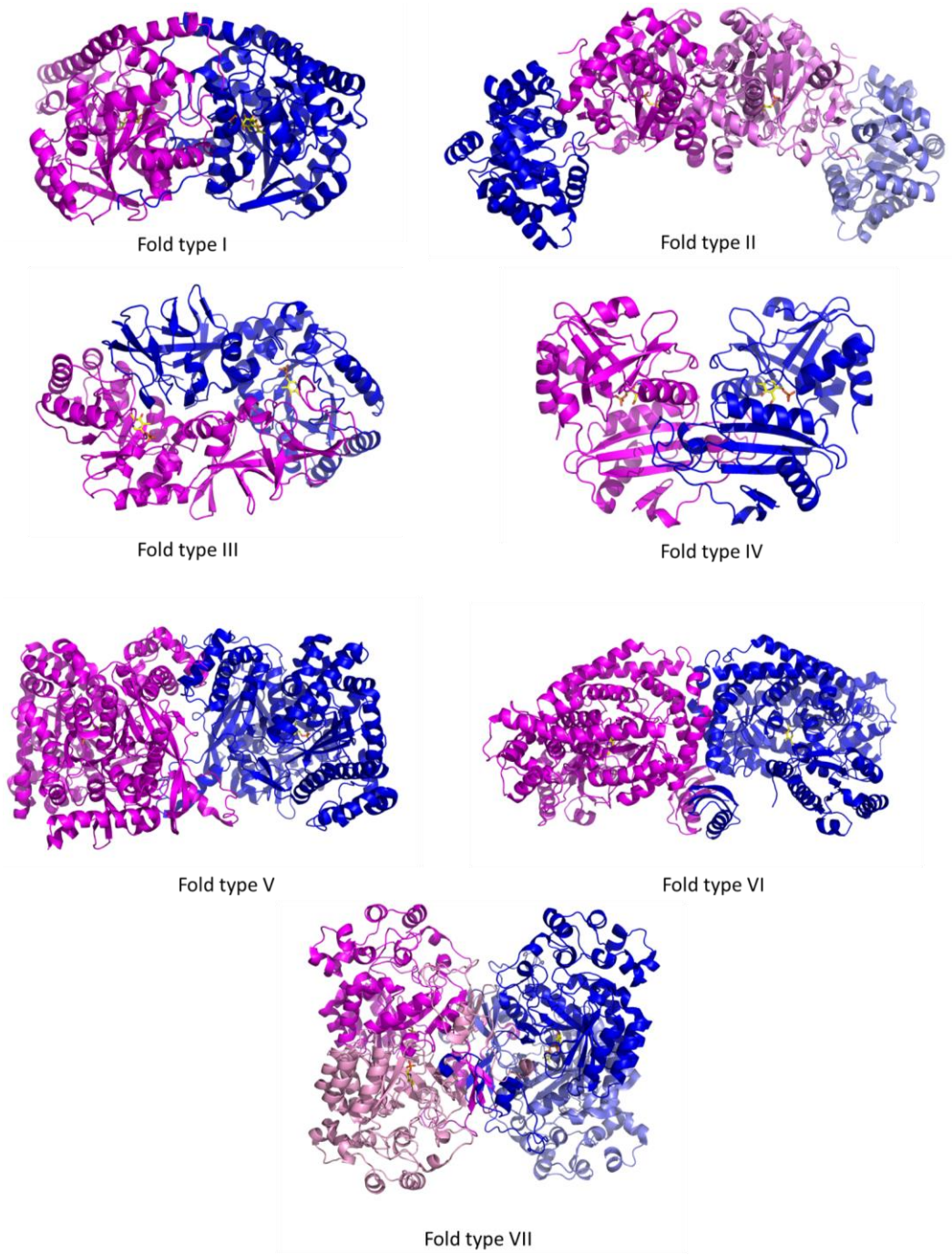


Figure 1.16 PLP fold type examples. Type I Aspartate aminotransferase from *Gallus gallus* (PDB code: 1IVR),[87] Type II tryptophan β-synthase from *Salmonella typhimurium* (PDB code: 2RHG), Type III Alanine racemase from *Geobacillus stearothermophilus* (PDB code: 1SFT),[88] Type IV D-amino acid aminotransferase from *Bacillus* sp. (strain YM-1) (PDB code: 1DAA),[89] Type V Glycogen phosphorylase from *Oryctolagus cuniculus* (PDB code: 1GPA),[90] Type VI lysine 5, 6 aminomutase (PDB code: 1XRS),[85] Type VII the lysine 2, 3 aminomutase (PDB code: 2A5H).[86] Monomer A is coloured in magenta while monomer B in blue. PLP is shown as sticks and highlighted in yellow.

1.4 Aminotransferases

ATs are enzymes which catalyse the interconversion of amino acids and oxoacids by transfer of the amino group. The catalytic mechanism can be split in two half reactions that see PLP recycling between its aldehyde and amino forms (Figure 1.17).[69]

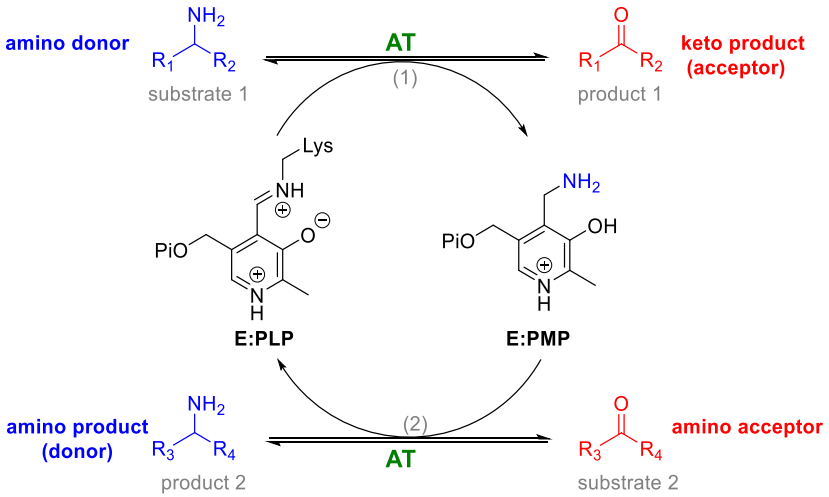


Figure 1.17 Schematic view of the AT reaction. In the first half-reaction (1) an amino donor donates the amino group to the cofactor, generating PMP and a keto co-product. In the second half of the reaction (2) the PMP acts as amino donor, transaminating a keto substrate, generating the final amine product (which can also be a donor in the reverse direction) and regenerating the PLP cofactor.

Based on the position of the transferred amino group, ATs can be grouped into α-ATs and ω- AT subfamilies; α-ATs only accept substrates with the amino group in the α-position to the carbonyl carbon while ω-AT have a much broader substrate scope as they accept substrates with at least one methylene group between the α-carbon and the carboxylate moiety (Figure 1.18).

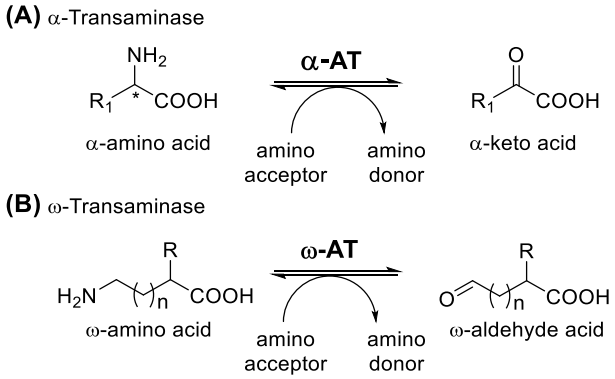


Figure 1.18 ATs are divided into α-ATs **(A)** and ω-ATs **(B)** according to whether the transfer of the amino group occurs at the α-carbon (α-ATs) or further away from the carboxylic moiety (ω-ATs).

For this reason ω-ATs are more appealing for industrial biocatalysis as they have a broader substrate range and are not limited to α-amino acids. A further distinction based on the product enantioselectivity can also be made, defined as (*S*)-AT and (*R*)-AT. Several (*S*)-ATs have been characterised, while a more limited number of (*R*)-AT isoforms have been identified.[91–93]

The application of ATs for the stereoselective amination of prochiral ketones represents an environmentally benign and economically appealing alternative to transition metal catalysed asymmetric synthesis.[51–54] ATs display high enantio- and regio-selectivity, broad substrate specificity and high reaction rate and stability, all attractive qualities for industrial applications. Because of these features, ATs have been widely used in recent years for the production of a wide range of products such as natural and unnatural amino acids, amino alcohols, amino sugars and chiral amines.[59,63–65]

The application of an engineered AT for the synthesis of 'blockbuster' drugs such as Januvia™(Sitagliptin) (Figure 1.4) has shown the enormous power and efficiency of this important class of biocatalysts. However, despite the enormous potential of ATs, challenges associated with substrate and co-product inhibition, challenging reaction equilibria and substrate scope have limited the widespread use of ATs. Many researchers are addressing these problems from different angles, trying to expand substrate scope and develop amine donors which can drive the equilibrium towards product formation.[52,101–105] Some of these developments will be addressed in section 3.3.4.

Although significant progress has been made towards expanding the substrate/product scope and overcoming the unfavourable equilibrium of ATs, there is still a list of challenging substrates which would be exciting to aminate (from a both commercial and scientific perspective) using this class of biocatalysts. The discovery of novel ATs through metagenomic approaches, combined with enzyme evolution will most likely expand the application of ATs. Moreover, the availability of structural information will no doubt help to further open the door to the expansion of the AT substrate scope.

1.4.1 Reaction mechanism

ATs are assumed to follow a ping-pong bi-bi mechanism. This means that the first substrate (amino donor) enters in the active site and only after the first product leaves the active site,

the second substrate (amino acceptor) can bind generating the second product and regenerating the enzyme. A detailed AT reaction mechanism is described in **Figure 1.19**. In the resting state, the ε group of catalytic lysine is covalently bound to the PLP cofactor to form an internal aldimine. In the first half-reaction an amino donor, generally an amino acid, displaces the internal aldimine *via* a transaldimination reaction to form the external aldimine. The released lysine residue can now act as a catalytic base in the next step abstracting the α-proton from the external aldimine to form a planar quinonoid. After hydrolysis, a keto co-product and pyridoxamine 5'-phosphate (PMP) are formed. In the second half-reaction the inverse of the first half reaction occurs, where after deprotonation by the lysine residue, the amino acceptor reacts with PMP to form a second quinonoid and water. In the last step, the chiral amine forms an internal aldimine and the final chiral product is released and the cofactor regenerated. Each of the steps in this reaction are reversible resulting in an equilibrium between the products and reactants.

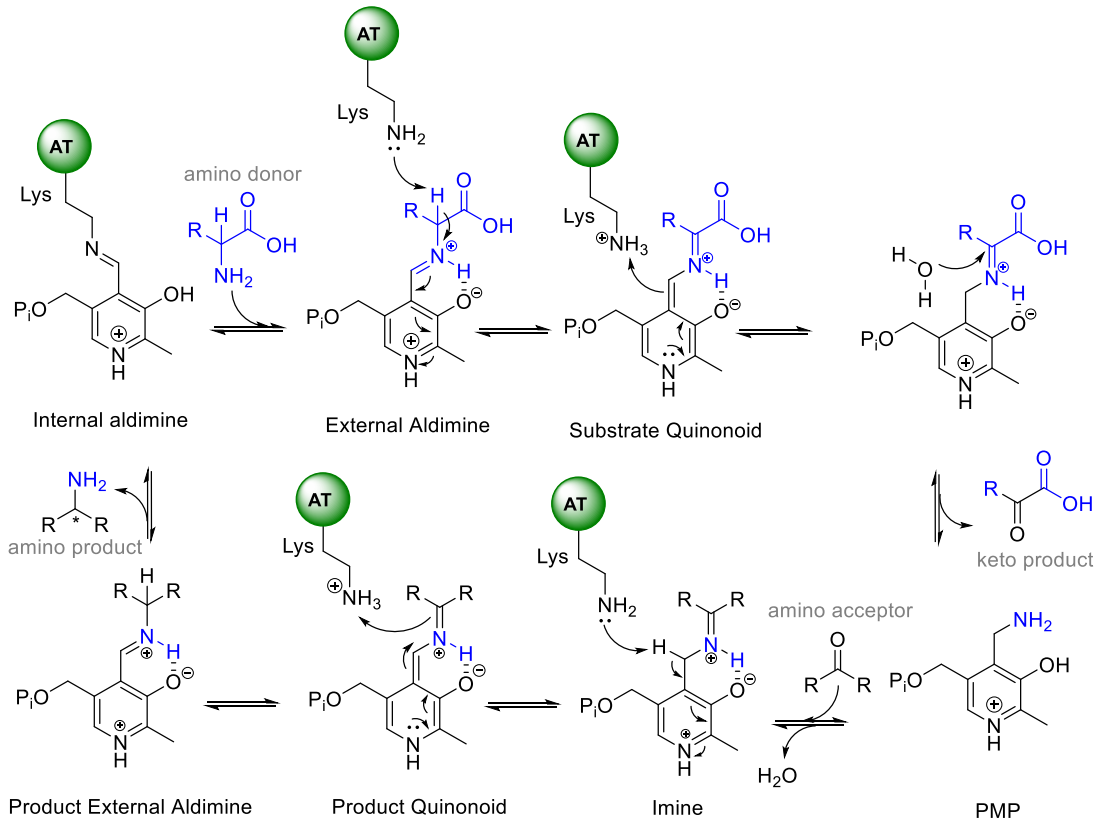


Figure 1.19 AT reaction mechanism. Steps include, transaldimination, deprotonation of the substrate, protonation of PLP, hydrolysis of the amino donor to yield PMP, condensation of the amino acceptor, deprotonation of PLP, protonation of the product and transaldimination to release the product.

The stereospecificity of ATs depends on the position of the catalytic lysine residue in comparison with PLP. According to the reaction mechanism of ATs, the reaction proceeds *via* a planar quinonoid intermediate, and the stereocenter is formed by protonation from the catalytic lysine on either the *si*- or the *re*-face of PLP (Figure 1.20). The catalytic lysine in (*S*)-ATs (fold type I) acts from below the quinonoid intermediate on the *si*-face, while the catalytic lysine in (*R*)-ATs and in branched-chain ATs (fold type IV) acts on the *re*- face above the quinonoid intermediate.[106]

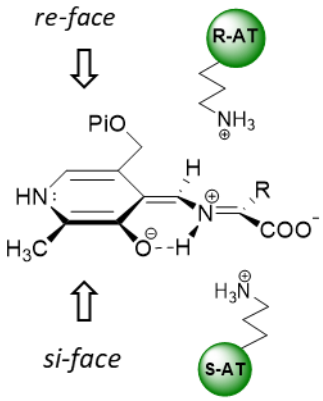


Figure 1.20 The position of the lysine compared to PLP determines the stereospecificity of the AT catalysed reaction. In (*S*)-AT the catalytic Lys acts from below the quinonoid intermediate on the *si*-face, thus yielding the (*S*)-enantiomer. In contrast, in (*R*)-AT the Lys acts from above the quinonoid intermediate on the *re*-face, producing the (*R*)-enantiomer.

1.4.2 Structural analysis of transaminases

TAs are found in the fold types I and IV families of PLP enzymes, where fold-type I is the most common fold for ATs and includes aspartate, ornithine, and ω-amino acid ATs. D-amino acid ATs and branched-chain ATs belong to the fold type IV. In addition to the traditional classification of PLP enzyme fold types, Mehta *et al.* further classified ATs into four subgroups based on sequence similarity and secondary structure predictions. Subgroup I ATs comprise aspartate, alanine, tyrosine, histidinol-phosphate-, and phenylalanine- ATs; subgroup II consists of acetylornithine, ornithine, 4-aminobutyrate and diaminopelargonate ATs; subgroup III contains D-alanine and branched-chain amino acid ATs, and finally subgroup IV comprises of serine and phosphoserine ATs.[82]

More recently, Steffen-Munsberg *et al.* have proposed a new classification of ATs based on sequence and structural similarity and grouped ATs in six subgroups or classes;

class I: L-aspartate ATs; class II: L-alanine ATs; class III: ω-ATs; class IV: D-amino acid ATs and branched chain ATs (BCAT); class V: L-serine ATs; class VI: sugar ATs.[107,108] This classification will be adopted in this thesis, focusing on the class III, ornithine-like, AT family as it presents features attractive for industrial biocatalysis such as broader substrate scope.

ATs are interesting enzymes as they are able to accommodate different substrates in their active site while discriminating against all others. This phenomena is referred as 'dual substrate recognition'.[109,110] This dual specificity could either be achieved by movements of the PLP cofactor between two different binding sites or by the movement of flexible side chains to accommodate different substrates. The latter has been observed for the first time in the *E. coli* aspartate AT (AspAT) and is achieved by the so called 'arginine switch'. [113,114] The solving of the structure of a hexamutant of *E.coli* AspAT which displayed higher activity towards aromatic substrates, highlighted the movement of an active site arginine residue (Arg292) out of the active site. The Arg292 residue either shifted its position to directly interact with charged substrates or moved aside to allow access of aromatic ligands (Figure 1.21).[69,111,112]

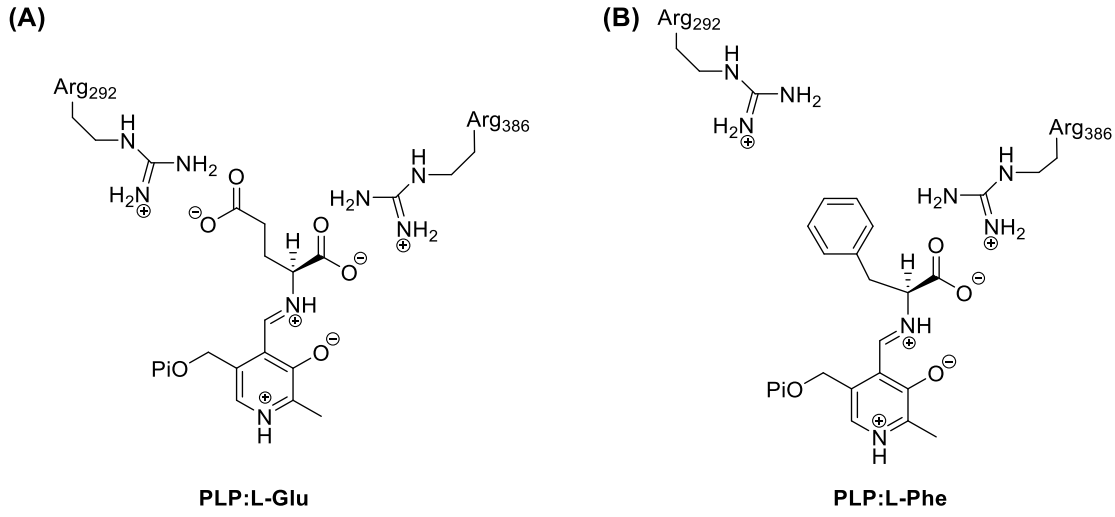


Figure 1.21 Schematic view of the arginine switch adopted by *E. coli* AspAT. **(A)** The γ carboxylate of L-Glu interacts with Arg292. **(B)** The Arg292 points out the active site when aromatic substrates such as L-Phe bind. The Arg386 interacts with the carboxylate of both substrates. This is how ATs display their dual substrate recognition and are capable of accepting both types of substrates. Adapted from Eliot *et al.*[69]

In a more recent report, Steffen-Munsberg *et al.* identified an analogous 'flipping' arginine residue in *Paracoccus denitrificans* AT which plays the same role, directly interacting with

the carboxylate of the substrates or providing a hydrophobic environment by moving out of the active site.[104,113,114]

A typical class III AT is *Chromobacterium violaceum* AT (Cv-AT), which has been widely used for biocatalytic application as it displays a quite broad substrate scope. Its structure was successfully solved highlighting key features of this AT.[115,116,117] Cv-AT is a homodimer in which the PLP cofactor lies at the dimer interface (Figure 1.22A). Each monomer is composed of a small and a large domain (Figure 1.22B). A three-layered α/β/α sandwich comprising seven-stranded β-sheets constitutes the large domain. All the β-sheets except one are parallel, with the catalytic Lys located in the loop preceding the only antiparallel strand. The small domain consists of residues from both the C- and N-termini which can be divided into two lobes (Figure 1.22 B).

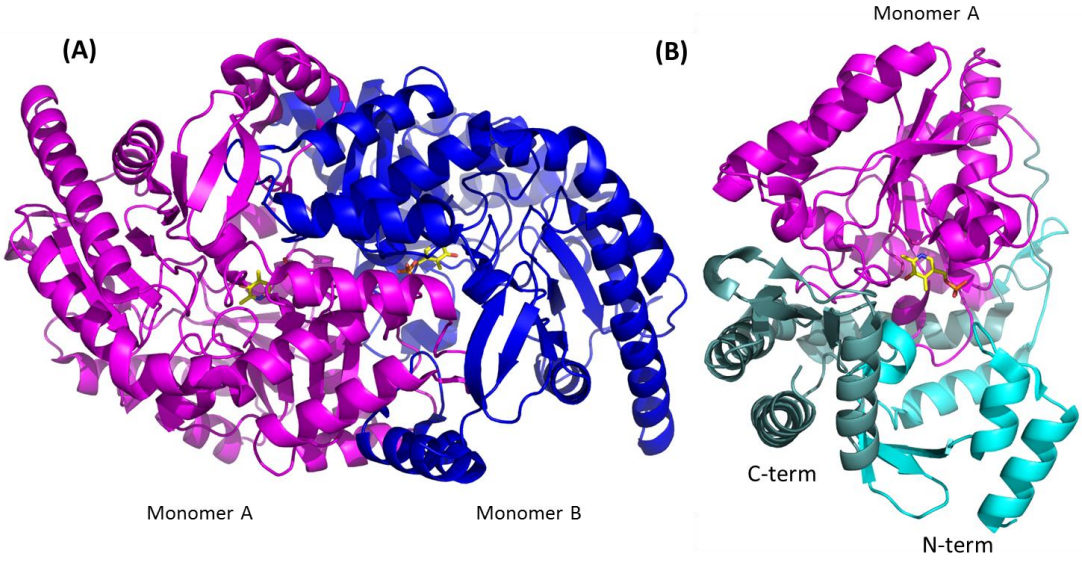


Figure 1.22 Crystal structure of the Cv-AT obtained at 1.8 Å (PDB code: 4A6T). **(A)** Overall structure of Cv-AT, monomer A is highlighted in magenta while monomer B is in blue. **(B)** Monomer A of Cv-AT where the large subunit is in magenta, the C-terminal lobe of the small subunit is in light teal while the N-terminal lobe is in cyan. PLP is shown as sticks and highlighted in yellow.

The interplay of residues from both monomers is crucial for cofactor stabilisation. Typical residues, highly conserved amongst the family, include the Tyr residue, perpendicular to the PLP cofactor and the Val residues that sandwich the PLP. The electron sink nature of the cofactor is enhanced by a hydrogen bond interaction of the pyridinium nitrogen with an aspartic acid residue (D259), which maintains the cofactor in the protonated form. The phosphate group of the cofactor is coordinated by a series of hydrogen bonds with Ser118,

Gly119 and Ser120 from one monomer and Thr320 and Phe319 from the other one, providing a very stable anchoring point for PLP also known as the 'phosphate cup'.[118] These general features are highly conserved amongst the class III family and will be further discussed in section 3.6 and 4.9.

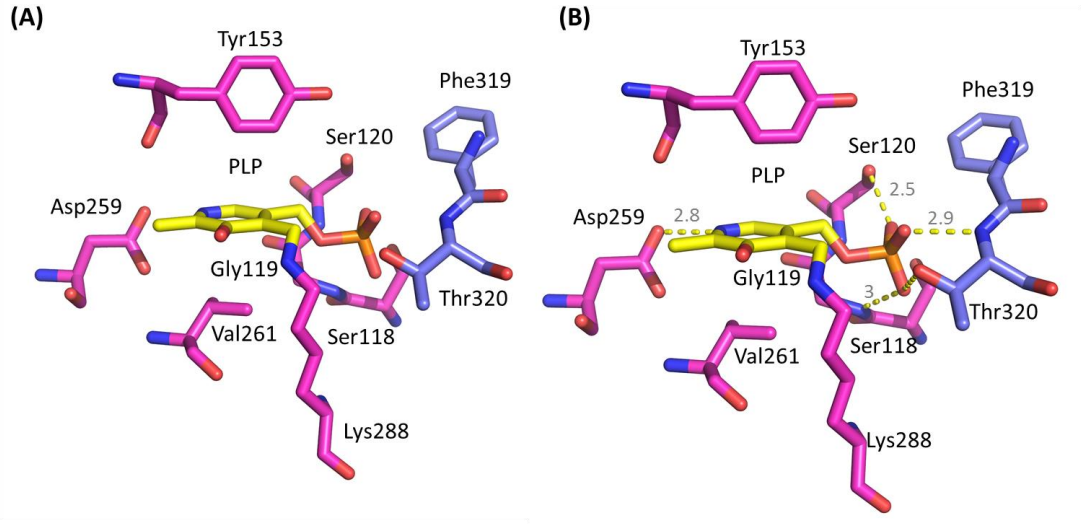


Figure 1.23 (A) Cv-AT PLP internal aldimine. The PLP internal aldimine and residues involved in cofactor stabilization are shown as stick representations, with carbon atoms shown in magenta and blue for the two monomers in the active dimer, oxygens are shown in red, nitrogen in blue and phosphorous in orange. Coordinated water molecules are shown as red spheres. The aromatic ring of PLP is sandwiched between Val261 and Tyr153. The nitrogen of the pyridine ring is fixed by hydrogen bond coordination to Asp259. The phosphate group is stabilized by a network of interactions with Ser118, Gly119, and Ser120 from Monomer A; and Phe319 and Thr320 from monomer B. **(B)** Hydrogen bonds, direct or through water molecules, are shown as dotted lines and distances (Å) in grey.

1.5 Target compounds and biocatalysts

This study is focused on the biocatalytic production of two main classes of molecules: the phenylglycine compounds and the fatty amines. These classes of molecules and the catalysts chosen for their biocatalytic production will be introduced in this section with details about the current synthetic methods and understanding of these enzymes.

1.5.1 The phenylglycine family

A noteworthy family of non-proteinogenic amino acids that occur in natural products, but also find broad application in synthetic compounds, is the phenylglycine family. This family includes D- phenylglycine (D-Phg) and D-4-hydroxyphenylglycine (D-Hpg).[119]

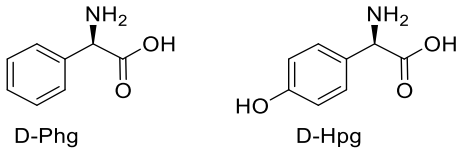


Figure 1.24 D-Phg and D-Hpg structure.

In contrast to proteinogenic amino acids, in phenylglycines the bulky aromatic sidechain is directly attached to the α-carbon. This strongly decreases the degrees of freedom of rotation for the aromatic side chain, structurally influencing phenylglycine containing peptides.

Phenylglycine-type amino acids occur in a wide variety of peptide natural products, including glycopeptide antibiotics and biologically active linear and cyclic peptides. In vancomycin, for instance, D-Phg plays a crucial role in the structure and the function of the final glycopeptide. Similarly, synthetic drugs such as amoxicillin, one of the top-selling antibiotics developed by GSK, contain D-Hpg in its structure (Figure 1.25).

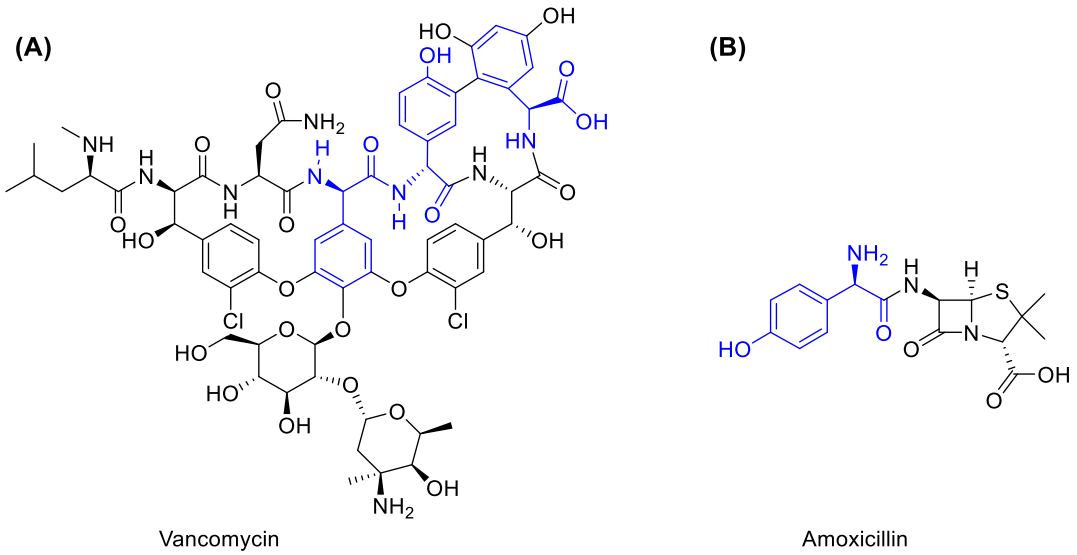


Figure 1.25 Natural and synthetic compounds containing the D-Phg and D-Hpg moiety. **(A)** Vancomycin is a natural glycopeptide antibiotic produced by *Amycolatopsis orientalis* whilst **(B)** amoxicillin is an antibiotic produced by GSK.

It is interesting to characterise how nature has evolved a biosynthetic pathway to produce these compounds. Despite the fact that Phg is structurally the simplest member of the phenylglycine family, its biosynthetic pathway has been only recently identified following the sequencing of the pristinamycin biosynthetic pathway in *Streptomyces pristinaespiralis.*[120] An operon of five genes (*plgA-E*) was identified and their functions have

been assigned. Bioinformatic analysis of the pathway suggests that synthesis starts from phenylpyruvic acid (PPA), which is converted into phenylacetyl-CoA (PA-CoA) by PlgB and PlgC, assumed to work as a pyruvate dehydrogenase complex. The hydroxylacyl-dehydrogenase PlgA converts PA-CoA into benzoylformyl-CoA (BZF-CoA) before the thioesterase PlgD catalyses the hydrolysis of BZF-CoA, releasing BZF and CoA. The final step is an L-specific AT PlgE which yields L-Phg from the keto-acid substrate BZF (Figure 1.26).

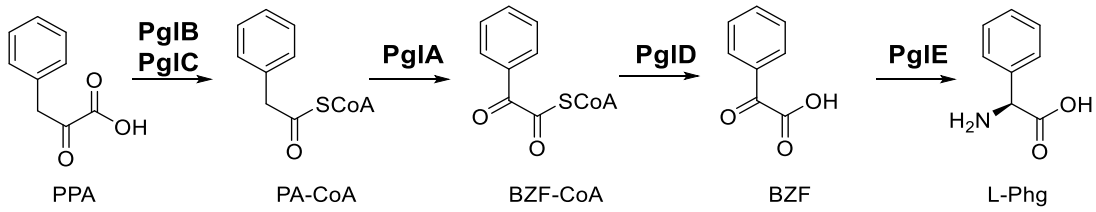


Figure 1.26 Biosynthesis pathway of L-Phg (Phg) hypotesized from the pristinamycin biosynthesis gene cluster (*pglA*, *pglB*, *pglC*, *pglD* and *pglE*) of *Streptomyces pristinaespiralis*. Adapted from Al Toma *et al.*[121]

The Hpg synthetic pathway has been elucidated following sequencing of the chloroeremomycin biosynthetic cluster *in Amycolaptosis orientalis*.[122,123] The first step is conversion of prephenate (PPA) into 4-hydroxyphenylpyruvic acid (HPPA) by the action of the prephenate dehydrogenase (Pdh). L-4-hydroxymandelate synthase (HmaS) then catalyses the conversion of HPPA into L-4-hydroxymandelate (L-HMA) by hydroxylating the benzylic position of HPPA. Subsequently L-HMA oxidase (HmO), a flavin mono-nucleotide (FMN) dependent enzyme, catalyses the oxidation of L-HMA to 4-hydroxybenzoylformate (HBF). Finally, a transamination reaction catalysed by L-Hpg AT (HpgT), which uses L-Tyr as amino donor (Figure 1.27) yields the final product L-Hpg.

It is important to underline that both biosynthetic pathways produce the L-enantiomer of the desired product D-Phg/D-Hpg. An epimerization (E) domain is found in the non-ribosomal peptide synthetase (NRPS) module where the Phg residue is incorporated in the peptide as the D-isomer. E-domains are structurally related to condensation domains that catalyse peptide bond formation and have been shown to epimerise the residue once it is incorporated within the peptide, explaining why D-enantiomers are found in the final natural products.

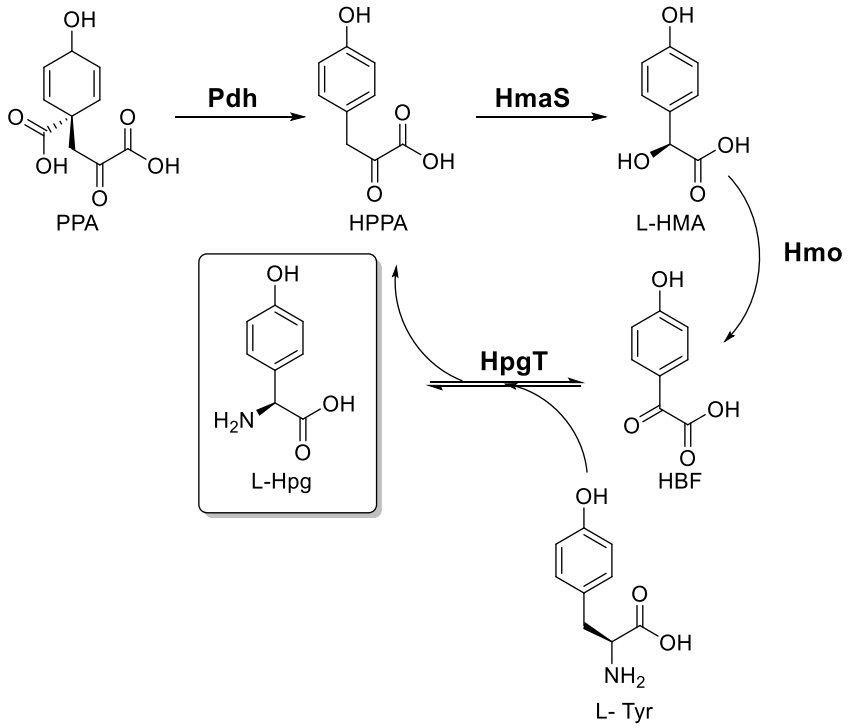


Figure 1.27 Hpg biosynthetic pathway in *Amycolaptosis orientalis*. Pdh: prephenate dehydrogenase, HmaS: L-4- hydroxymandelate synthase, HmO: L-4-hydroxymandelate oxidase, HpgT: 4-hydroxybenzoylformate/L-4-tyrosine AT. Adapted from Al Toma *et al.*[121]

Synthetically, the conventional preparation of D-Phg in industry, developed in the early 1970's, is a diastereomeric salt recrystallization of the phenylglycine racemic mixture using (+)-camphor-8-sulfonic acid (CSA) as the resolving agent (Figure 1.28, Route 1).[124] This method, as described previously, presents the main disadvantage of a maximum theoretical yield of 50%.

As D-Phg and D-Hpg are high value compounds, several biocatalytic approaches have been attempted for an efficient and enantioselective production of D-Phg and D-Hpg. Amongst these, the most widely used is a hydantoinase/carbomylase system.[125] Here, a chemically synthesised racemic hydantoin is first hydrolysed by a hydantoinase to yield the corresponding carbamoyl derivative, which is subsequently cleaved by a carbamoylase to afford the desired optically pure amino acid. Incorporation of a hydantoin racemase to perform dynamic kinetic resolution (DKR) affords theoretically a 100% yield when starting from the racemate. The hydantoin process has the further advantage that nature has provided us with D- or L-selective hydantoinases/carbamoylases, with broad substrate specificity and hence access to either D-or L-amino acids, including non-proteinogenic

derivatives, is possible. However, the substrate 5-phenylhydantoin has to be synthesized by a Bucherer–Bergs reaction with benzaldehyde, urea, and highly toxic cyanide, and the low solubility of 5-phenylhydantoin in aqueous solution results in a low overall reaction rate.

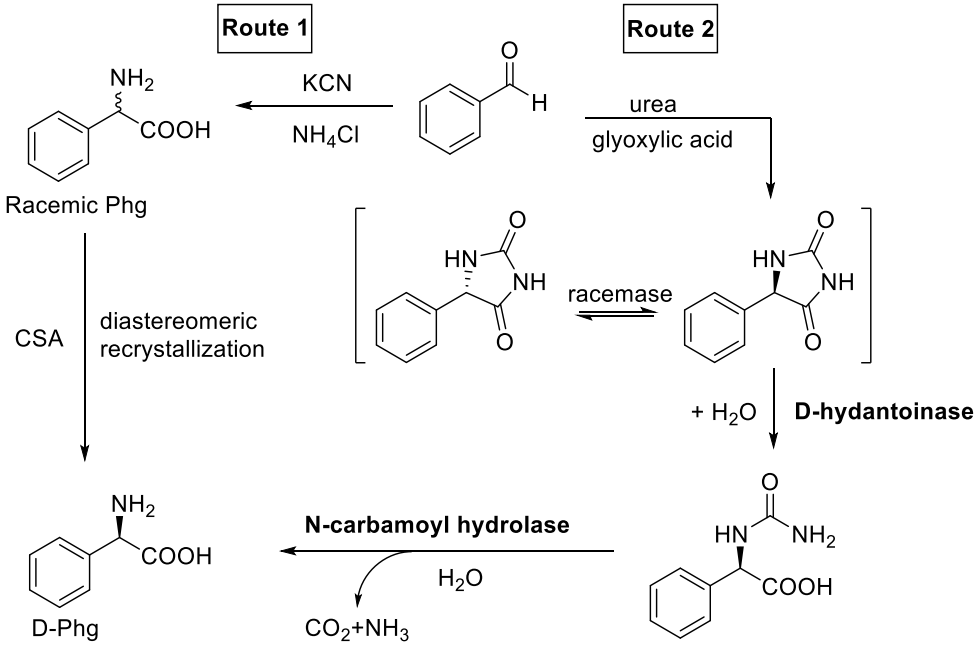


Figure 1.28 Conventional routes for D-Phg production. Route **(1)** Phenlyglycine diastereomeric salt crystallization using CSA as the resolving agent. Route **(2)** D-hydantoinase/carbomylase approach for the synthesis of D-Phg.

Several other DKR approaches have been reported applying lipases, nitrilases or amidases. However, all those methods fail to produce an *ee* greater than 88% and therefore are not suitable for direct applications.[125,126] Similar approaches are used to produce D-Hpg. As all the current methods present several drawbacks, the development of an efficient alternative for the synthesis of these compounds would be very advantageous.

1.5.1.1 The D-phenylglycine aminotransferase from *Pseudomonas stutzeri*

The D-phenylglycine AT (D-PhgAT), isolated from the soil bacterium *Pseudomonas stutzeri* ST-201, catalyses the reversible transamination of L-glutamic acid (L-glu) with benzoylformate (BZF) and 4-hydoxybenzolyformate (HBF), as amino acceptors, yielding D-Phg and D-Hpg respectively, and α-ketoglutarate (AKG) (Figure 1.29).[127] The unique feature of this AT is that it is a so-called ‘‘stereo-inverting’’ AT, a property whereby the

amino donors (L-Glu and D-Phg) in this reversible reaction exhibit inverse absolute configuration.

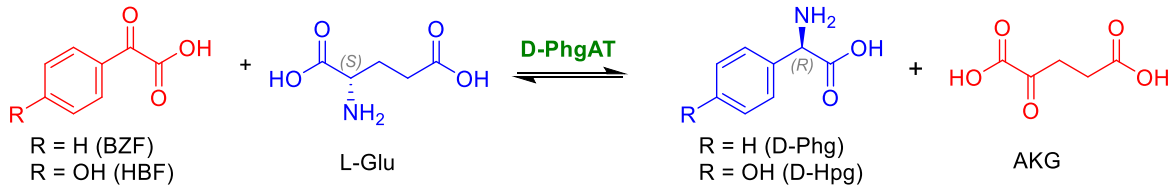


Figure 1.29 Reaction catalysed by the D-PhgAT. L-Glu donates the amino group to the pyridoxal 5'-phosphate (PLP) cofactor generating pyridoxal amine (PMP) and α-ketoglutarate (AKG). In step 2, benzoyl formate (BZF) or its 4-hydroxy derivative (HBF) accepts the amino group from PMP yielding D-Phg or D-Hpg respectively.

From a biocatalysis perspective, D-PhgAT has attractive features such as high turnover and stability and has been used in qualitative and quantitative analysis of L-Glu in food when coupled with L-glutamate dehydrogenase, and for quality control of amoxicillin in pharmaceuticals when coupled with a penicillin acylase.[128,129] The homologous D-PhgAT from *P. putida* (82% sequence identity) has been successfully combined with HmaS and Hmo to engineer an *E. coli* strain that is able to produce D-Phg.[130]

Wiyakrutta *et al.* have reported preliminary characterisation of D-PhgAT with the determination of K_M for D-Phg and AKG (1.1 and 2.4 mM respectively). The optimum pH and temperature have been determined to be 9.5 and 35 °C, and it was found that 10 mM BZF was converted with a yield of 80% after 3 h. The substrate scope, however, has been only marginally explored, suggesting a very narrow substrate scope. In addition, an *apo*-structure, lacking the PLP cofactor and several residues at the N-terminus has been deposited in the PDB in 2006 (PDB code: 2CY8, Figure 1.30).[131] The enzyme is a homodimer, whose structure belongs to PLP fold type I and class III ATs and which displays the typical features of these families. However the lack of the cofactor and active site residues do not allow for the identification of the molecular reasoning of the D-PhgATs unique enantioselectivity.

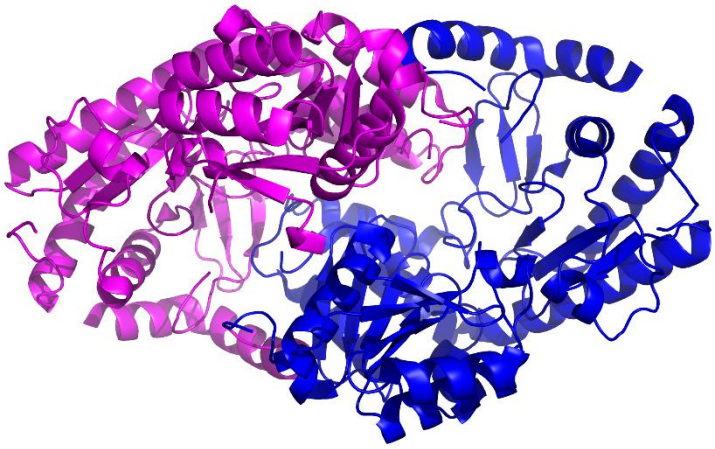


Figure 1.30 Incomplete structure **(**PDB code: 2CY8) of the apo-form of D-PhgAT deposited in the PDB database in 2006. Monomer A is shown in magenta while monomer B is shown in blue.

The ability of D-PhgAT to use an L-amino acid substrates and yield a final D-amino acid product represents a great advantage for industrial biocatalysis and therefore D-PhgAT was chosen as target for the synthesis of high value D-amino acids (chapter 3). A structural understanding of the active site architecture is required to understand the enantioselectivity of this enzyme and enhance its directed evolution.

1.5.2 Fatty amines

Aliphatic long chain or 'fatty' amines are a class of primary, secondary and tertiary amines from C_8 to C_{22}.[132] Fatty amines have a broad range of application but are predominately found in the manufacturing of emulsifying, cleaning and cosmetic formulations.[133] The global fatty amine market size was estimated at 670 kilo tons in 2016 and the global demand for these compounds is expected to grow (https://www.grandviewresearch.com/industry-analysis/fatty-amines-market).

Commercially available fatty amines consist of either a mixture of different carbon chain lengths or solely a specific chain length. Primary amines are the most common, but secondary and tertiary substitutions are also available.

Two main production routes are used to prepare fatty amines. One route is the hydrogenation of nitriles, used since the 1940s.[133] This involves the hydrolysis of naturally occurring materials such as fats and fatty oils at high temperatures (200-280 °C) to obtain saturated and unsaturated fatty acids and glycerol. A fatty nitrile intermediate produced

using ammonia is then hydrogenated using a variety of catalysts, first to an imine and then to a primary amine. An alternative method uses fatty alcohols and ammonia. Natural fatty acids can be catalytically reduced to alcohols by high pressure (>20 MPa) and temperature (250–300 °C) hydrogenation using a copper chromite catalyst. Fatty amines can then be produced from the fatty alcohol using excess ammonia, at elevated temperature and high pressure (50-340 °C, 3.5 MPa). Neither of these approaches are eco-friendly and both require very harsh conditions. Thus, the development of a suitable biocatalytic alternative for their production would be of great interest.

1.5.2.1 Fumonisin and the aminotransferase FumI

Mycotoxins are secondary metabolites of fungi which are harmful to humans and animals. Ingestion of these mycotoxins causes mycotoxicoses, which can result in death.[134] A particular class of mycotoxins, produced by the maize pathogen *Fusarium verticilloides* and other fungi, are the fumonisins.[135] As their chemical structures resemble the sphingolipids, fumonisins are able to bind the mammalian ceramide synthase, disrupting sphingolipid metabolism.[136,137] Amongst the fumonisins, fumonisin B_1 (FB_1) is the most abundant and has been shown to be carcinogenic, teratogenic and is suspected to cause oesophageal cancer (Figure 1.31).[138,139] In addition, the FB_1 degradation intermediate, hydrolysed FB_1 (HFB_1), is shown to cause defects in neonates.[139]

In 2005, by studying the degradation of this harmful mycotoxin in the bacterium *Sphingomonas sp.* MTA14, two genes responsible for FB_1 degradation were identified.[140] These genes, *fumD* and *fumI*, code for a carboxylesterase and an AT, respectively. Recombinant FumD was shown to catalyse the first step of FB_1 degradation, the hydrolytic release of the two tricarballylic acid (TCA) side chain of FB_1 to produce HFB_1 while FumI was shown to catalyse the transfer of the 2-amino group from HFB_1 to pyruvate, producing 2-keto- HFB_1 (Figure 1.31).[141]

The research organization BIOMIN (Austria) has investigated the application of FumD and FumI as a tool for food detoxification and this application is covered by a patent (WO2004085624). Despite the interesting feature of FumI being able to accept long chain amines, the substrate scope of this interesting enzyme has not been examined.[142,143] Studies have been limited to the determination of its kinetic properties towards HFB_1 and

pyruvate, and the biosynthetic potential of this enzyme has not been explored. Thus, the application of FumI for the synthesis of fatty amines from aldehydes and ketones represents an attractive alternative to the existing synthetic methods.

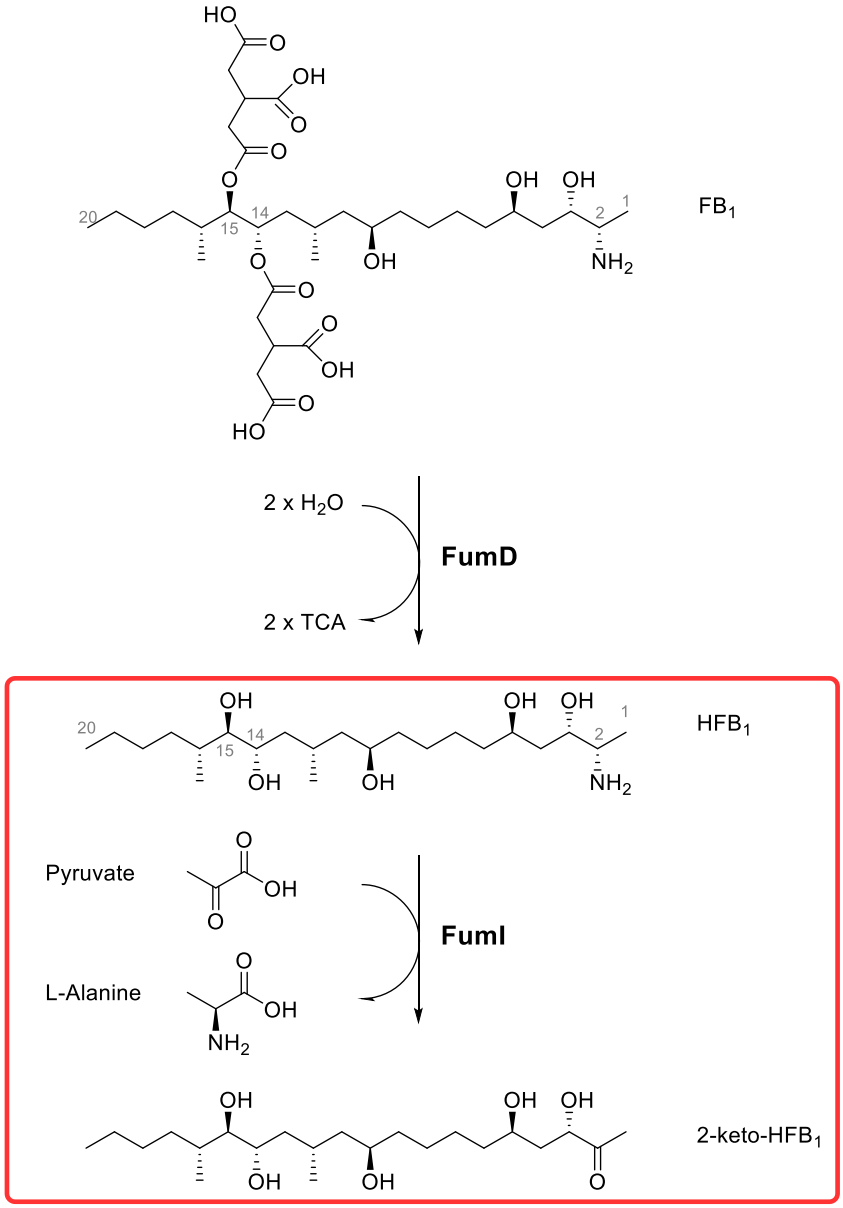


Figure 1.31 Initial steps of the FB_1 degradation pathway of *Sphingopyxis* sp. MTA144FB1. The carboxylesterase FumD catalyses the hydrolysis of FB_1 to HFB_1 and the aminotransferase FumI catalyses the transfer of the 2-amino group from HFB_1 to pyruvate, producing 2-keto- HFB_1. Adapted from Hartinger *et al.*[142]

2 Aims

The overall aim of this work is to investigate the substrate scope and the biosynthetic potential of the D-PhgAT and FumI by detailed structural and kinetic characterisation. As such, the main aims for each enzyme are as follows:

1) D-PhgAT:

- Express and purify recombinant *P.stutzeri* D-PhgAT
- Characterise D-PhgAT for the natural reaction
- Develop a suitable assay and optimise the biotransformation conditions
- Explore the amino donor and amino acceptor scope
- Use structural studies to understand the molecular basis of the unique 'stereo-inverting' property of the enzyme
- Immobilize D-PhgAT using a commercial resin to explore the possibility of recycling multiple times the catalyst

2) FumI:

- Express and purify recombinant *Sphingopyxis* FumI
- Investigate the 'reverse' FumI reaction
- Explore the ability of FumI to accept long chain alkyl ketones for the biosynthesis of high value fatty amines
- Explore the substrate (C_3-C_{20}) chain length tolerance
- Use structural studies to elucidate the residues involved in catalysis and substrate specificity

3 The D-Phenylglycine aminotransferase

The D-Phenylglycine aminotransferase (D-PhgAT) was firstly isolated in 1997 from *Pseudomonas stutzeri* ST-201, a soil bacterium capable of utilising D-phenylglycine as the sole carbon and nitrogen source.[127] D-PhgAT has the unique capability of catalysing a transamination reaction where the two amino donors L-Glu and D-phenylglycine (D-Phg) in the 'forward' and 'reverse' direction) exhibit opposite absolute configuration (Figure 3.1) and therefore is addressed in the literature as a 'stereo-inverting' AT. D-PhgAT channels the amino nitrogen directly between L-Glu and D-Phg without the requirement of additional enzymes (e.g. amino acid racemase for D-Phg or L-Glu). Thus, the enzyme represents an excellent target enzyme for the synthesis of enantiomerically pure D-Phg and derivatives such as D-4-hydroxyphenylglycine (D-Hpg) in a single step using a commercially available cheap amino donor such as L-Glu, providing a more concise and sustainable route to their biosynthesis than has previously been available.

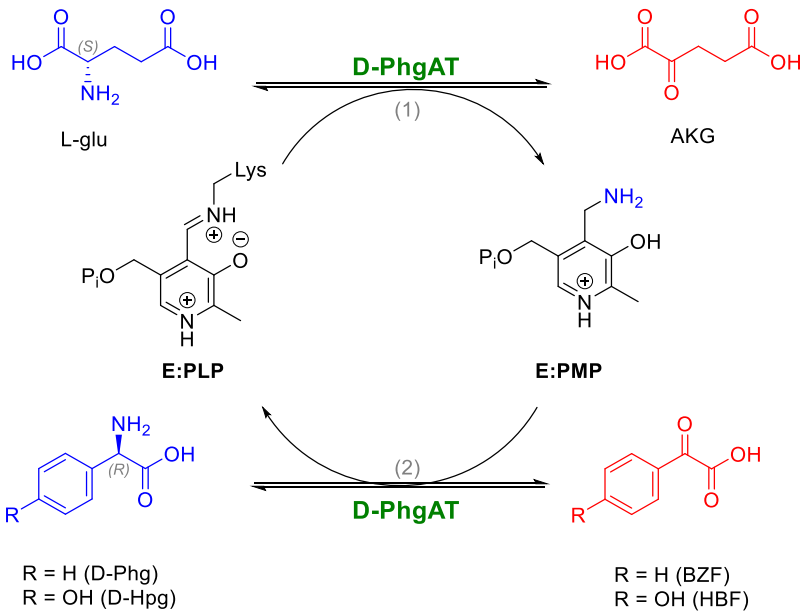


Figure 3.1 Schematic view of the reversible D-PhgAT reaction. In step 1 L-Glu (blue) donates the amino group to the 5'-pyridoxal phosphate (PLP) cofactor generating pyridoxal amine (PMP) and α-ketoglutarate (AKG) (red). In step 2 benzoyl formate (BZF) (red) or its 4-hydroxy derivative (HBF) accepts the amino group from PMP yielding D-Phg or D-Hpg respectively (blue).

3.1.1 Aims

The aims of this project are to express, purify and characterise D-PhgAT and assess its suitability for the synthesis of high value, enantiopure products. Very limited information is available in the literature about transamination reactions between two amino acids with inverted absolute stereochemistry, therefore a crystallographic investigation is required in order to fully illuminate the mechanistic origin of the enantioselectivity of this unique member of the AT superfamily. Determination of the 3D D-PhgAT structure will be central to expand and explore the biosynthetic potential of D-PhgAT, applying the enzyme to the synthesis of non-natural products.

3.2 D-PhgAT aminotransferase expression and purification

Full-length, codon optimized *dpgA* (*Pseudomonas stutzeri* ST-201, UNIPROT code: Q6VY99, Appendix 8.1) was purchased from GenScript and cloned into a pET-15b plasmid with NcoI/ BamHI restriction sites, to give recombinant D-PhgAT with a non-cleavable N-terminal His_6-tag (pET15b-dpgA) (Appendix 8.2).

Wiyakrutta *et al*. had reported that, in their hands, D-PhgAT exhibited low *in vitro* solubility and a tendency to aggregate and precipitate during purification, which limited its large-scale application.[144] The *dpgA* gene over expression in *E. coli* was reported to yield D-PhgAT mostly in the insoluble fraction and subsequent expression in *Pichia pastoris* with chaperones improved the solubility by just 20%.[145] In an effort to overcome this problem, different isopropyl β-D-1-thiogalactopyranoside (IPTG) concentrations and temperatures were tested in order to identify optimal conditions for protein expression. However, in contrast to Wiyakrutta *et al*., good conditions for protein expression were found to be fairly standard 0.1 mM IPTG with 16 h induction at 20 °C. Gene expression was induced under these conditions in 1 L of BL21 (DE3) *E. coli* cells.

The protein was then purified from the cell lysate using nickel immobilized metal affinity chromatography (IMAC), where it eluted at a concentration of 100 mM imidazole. A second purification step, size exclusion chromatography (SEC), yielded a symmetrical peak with an elution volume of 78.0 mL (Figure 3.2B). Based on the calibration curve associated with the gel filtration column used (Superdex 16/60 S200, Appendix 8.5), D-PhgAT eluted at

a retention volume corresponding to a homodimer (~ 100 kDa) in solution with the SDS-PAGE analysis showing a monomeric mass of ~ 51 kDa (Figure 3.2 B). Typical yields were on the order of ~ 70 mg protein per litre of culture.

To ensure the correct protein was expressed and purified, the mass of purified D-PhgAT was obtained using denaturing liquid chromatography electrospray – mass spectrometry (LC ESI-MS). The resulting spectrum displayed several peaks, representing the many charge states of the protein (Figure 3.3). The protein mass was deconvoluted from this spectrum using the maximum entropy (MaxEnt) algorithm from smoothed and centroided data. The observed mass of 50109.22 ± 0.53 Da was in keeping with the expected mass based on the recombinant protein sequence minus the initial methionine (50109.60 Da, Figure 3.3), calculated from the expressed protein sequence using ExPaSy ProtParam tool (https://web.expasy.org/protparam/). The loss of the methionine is very common as a post translational modification.[146] An additional peak was also visible which is believed to correspond to the gluconoylation of the histidine tag (+ 178 Da), a very common phenomenon observed in recombinant proteins expressed in *E. coli*.[147] A smaller additional peak can also be seen at each of the many charge states which corresponds to the PLP cofactor (230 Da).

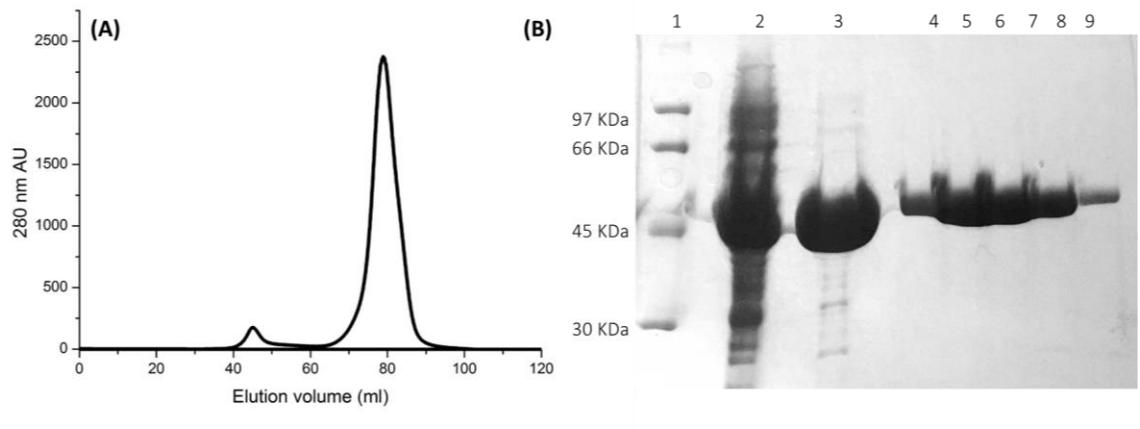


Figure 3.2 (A) Chromatogram from Superdex 16/60 S200 gel filtration chromatography of D-PhgAT. D-PhgAT elution volume corresponds to a dimer (79 mL). An aggregate is also visible at ~ 40 mL. **(B)** 12% SDS-PAGE gel of D-PhgAT purification steps. Lane 1: low molecular weight (LMW) marker (GE), Lane 2: Soluble fraction of cell lysate Lane 3: D-PhgAT after IMAC, Lanes 4-9: elution fractions from SEC (3 mL fractions from 69-87 mL)**.**

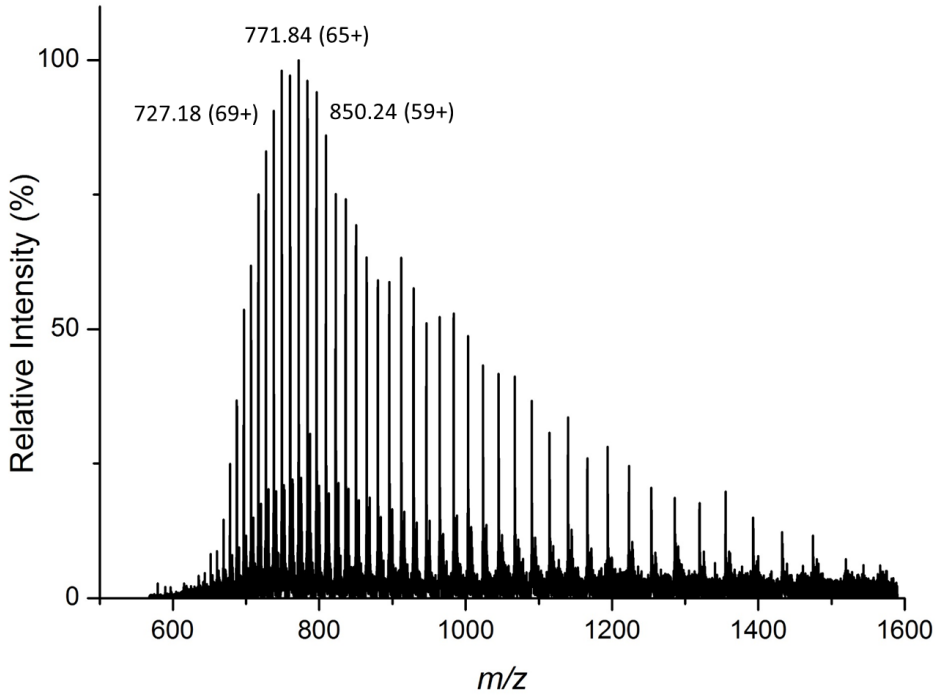


Figure 3.3 Denaturing LC ESI-MS analysis of recombinant D-PhgAT (20 μM). The deconvoluted mass of 50109.22 ±0.53 Da agrees with the expected value based on the sequence of 50109.60 Da. The values are *m/z* with the charge states given in brackets.

3.2.1 UV-Vis studies

Binding of substrates to PLP-dependent enzymes can be monitored through ultraviolet–visible (UV-Vis) spectroscopy of the PLP molecule where two absorbance maxima are usually observed. This is because the PLP molecule exists in two isoforms when bound as a Schiff base: the enolimine and ketoenamine forms (Figure 1.12). The enolimine form absorbs around 350 nm and the ketoenamine form around 420 nm. To investigate PLP binding, the enzyme was dialysed against buffer containing PLP (0.1 M CAPS, 150 mM NaCl, 50 μM PLP, pH 9.5) for 2 h in order to generate the *holo*-form of the enzyme. Excess PLP was then removed using a PD-10 desalting column and the resulting sample was analysed by UV-Vis spectroscopy.

The enzyme shows a distinct UV-Vis spectroscopic profile with a λ_{max} at 411 nm, indicative of PLP binding (Figure 3.4A). The formation of different intermediates results in distinct changes to the UV-Vis spectrum of PLP, allowing these intermediates to be followed spectroscopically. Upon addition of the amino donor (L-Glu), a shift towards 330 nm is

observed in the first instance, corresponding to the formation of the external aldimine (Figure 3.4B). Over time, an additional shift toward 314 nm is observed, indicating the formation of the PMP intermediate (Figure 3.4C). Interestingly, if the amino acceptor (BZF) is added to the reaction mixture, a shift back towards 330 nm is observed suggesting that the PMP has donated the amino group to BZF and reverted to the PLP external aldimine (Figure 3.4D). An overview of all the spectral changes is shown in Figure 3.4. From these spectral changes, the reaction progress upon substrate addition can be monitored.

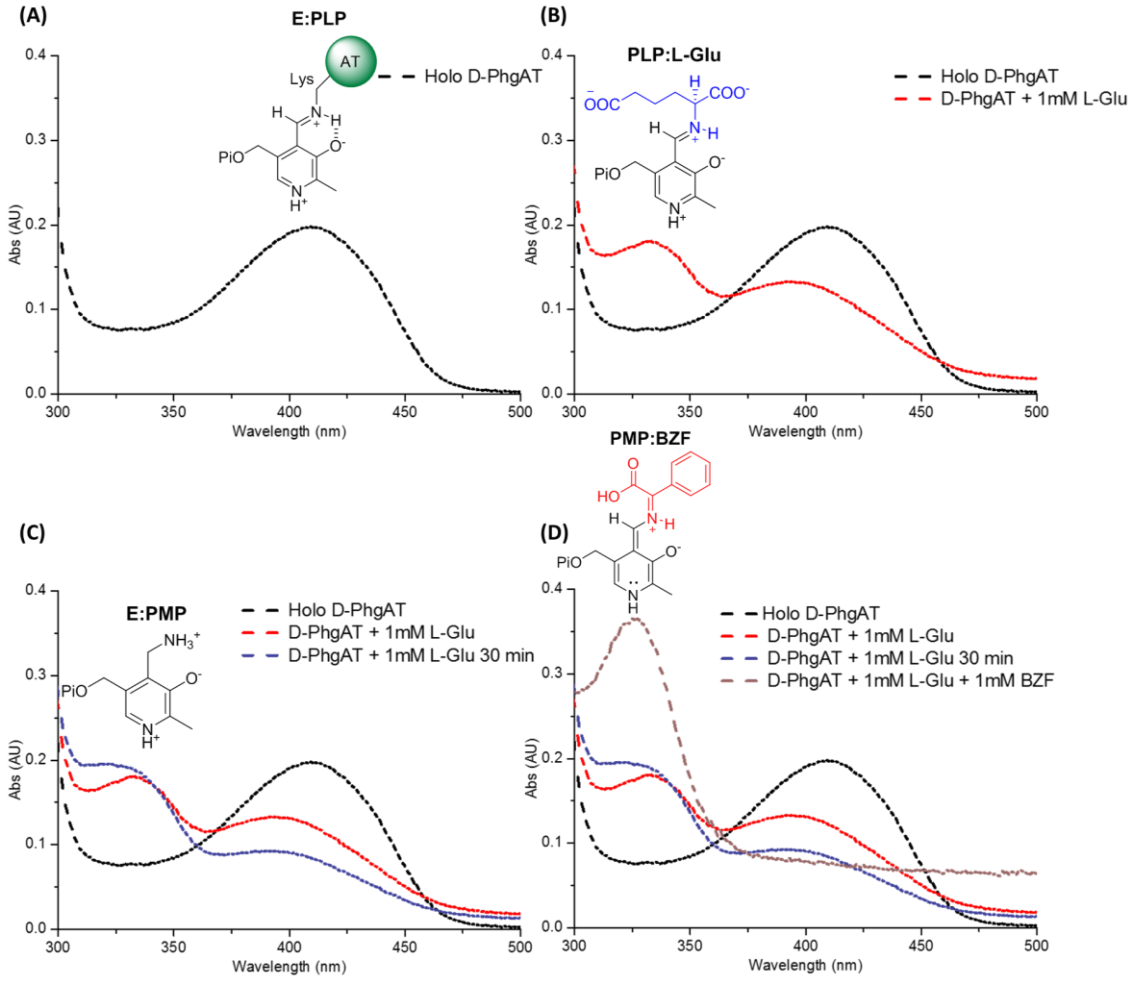


Figure 3.4 (A) UV-Vis scans of D-PhgAT internal aldimine (E:PLP). (**B)** The PLP external aldimine profile after the addition of L-Glu (1mM). The 411 nm peak, corresponding to the internal aldimine, decreases and a new peak at 330 nm, corresponding to the external aldimine (PLP: L-Glu) is observed. **(C)** Over time a further shift towards 314 nm suggests the production of PMP (E:PMP). **(D)** Upon addition of BZF (1 mM) a shift back towards the 330 nm region is observed suggesting the formation of the external aldimine and the consumption of PMP (PLP:BZF).

Using this approach, binding of different substrates, such as D-Phg and other amino donors/acids, can be followed. A general overview of the PLP profile upon amino donor binding is shown in Figure 3.5. Due to the fast enzyme turnover after substrate binding, calculation of the dissociation constant K_d (measurement of dissociation of the substrate from the ES complex) could not be achieved. This is because the fast shift towards the shorter wavelength would make the K_d calculation at 330 nm unreliable.

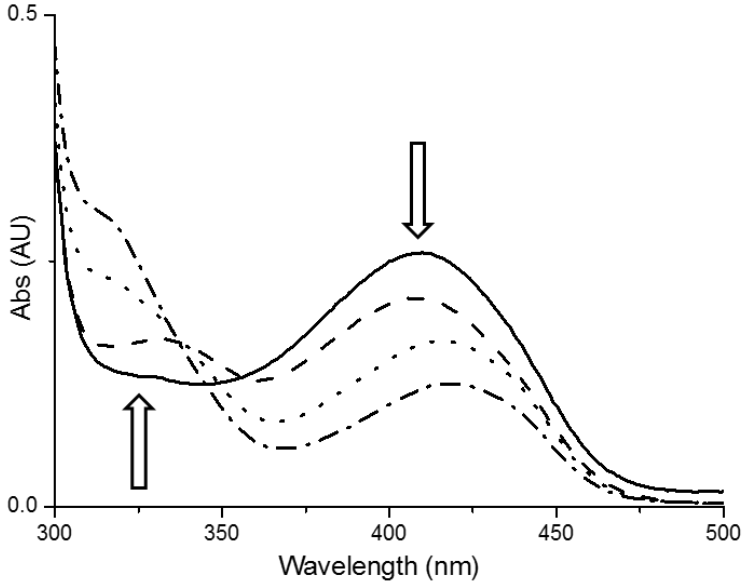


Figure 3.5 Generic UV-Vis profile of D-PhgAT upon amino acid addition over time. The D-PhgAT internal aldimine displays an absorbance maximum at 411 nm. Upon amino donor addition a first shift towards external aldimine (330 nm) is observed. Over time (depending on the amino donor and the concentration used) a further shift, observed towards PMP at ~ 311 nm is observed, suggesting that the enzyme is turning over and the exchange of the amino group from the amino donor to the PLP cofactor.

3.3 Enzyme assay

Two different approaches were developed with the purpose of checking the enzyme activity and to determine the enantioselectivity of the enzyme. A spectrophotometric coupled assay, using a commercially available α-ketoglutarate dehydrogenase (AKGDH) from porcine heart, was developed as a quick method of validating enzyme activity. Additionally, a high-performance liquid chromatography (HPLC) based method was also developed. This was more time consuming but was necessary to determine the enzyme enantioselectivity.

3.3.1 Coupled AKGDH assay (D-PhgAT/AKGDH)

A high-throughput assay for D-PhgAT activity was developed by coupling the enzyme with AKGDH. AKGDH is a key enzyme in the citric acid cycle, crucial for human health as decreased activity can lead to neurodegenerative diseases such as Alzheimer's disease.[146] AKGDH converts AKG into succinyl-CoA in presence of CoASH and NAD^+. Appearance of NADH can be readily monitored at 340 nm (ε_{340} = 6220 M^{-1} cm^{-1}), providing a fast route to measuring D-PhgAT activity. A scheme of the assay is show in Figure 3.6.

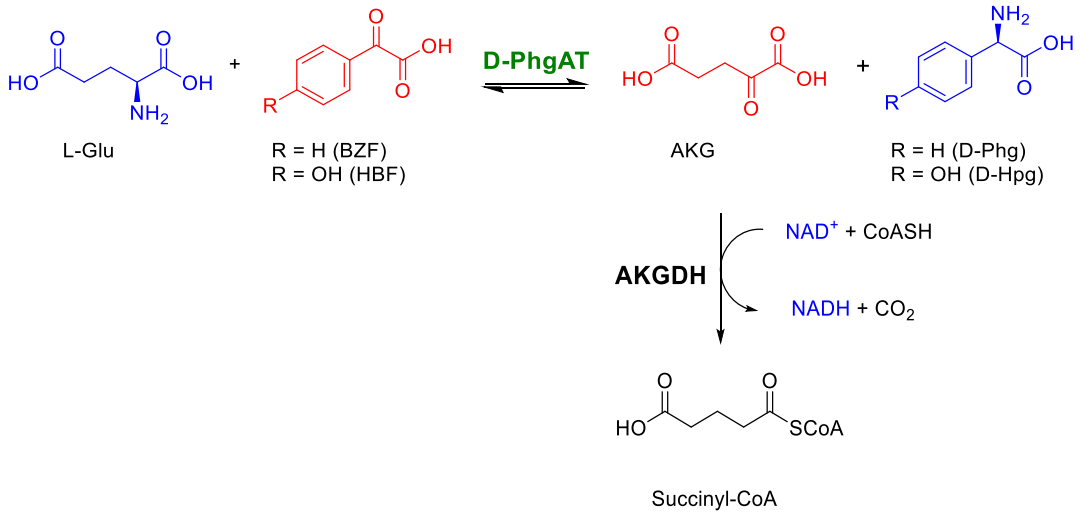


Figure 3.6 Reaction scheme of the D-PhgAT/AKGDH coupled assay. The AKG, product of the first half reaction, together with NAD^+ and CoASH, are substrates of the AKGDH. The conversion of NAD^+ to NADH can be easily monitored spectrophotometrically at 340 nm.

A typical reaction profile of the NADH appearance at different L-Glu and BZF concentrations is shown in Figure 3.7, where a steeper gradient is observed at increasing substrate concentrations (Figure 3.7A and B). The initial rate (over the first 15 min) was used to determine kinetic parameters and obtain Michaelis-Menten plots (Figure 3.8, Table 3.1). Due to high background absorbance of HBF at 340 nm, the kinetic characterisation for the substrate was not achievable using the D-PhgAT/AKGDH assay.

D-PhgAT exhibits a higher affinity towards BZF with a K_M 5-fold lower than the K_M for L-Glu (1.81 mM for BZF vs. 9.85 mM for L-Glu). Reported apparent K_M values for D-Phg and AKG are 1.1 mM and 2.4 mM, but no kinetic parameters have been reported for the other half of the reaction, so no comparison is available for BZF and D-Phg. These are the first reported kinetic parameters for this reaction.[127] Interestingly, the D-PhgAT enzyme

shows a high catalytic efficiency (k_{cat}/K_M) towards the substrates and a fast turnover with a k_{cat} in the order of sec. Thus, the enzyme has potential application in industrial biocatalysis.

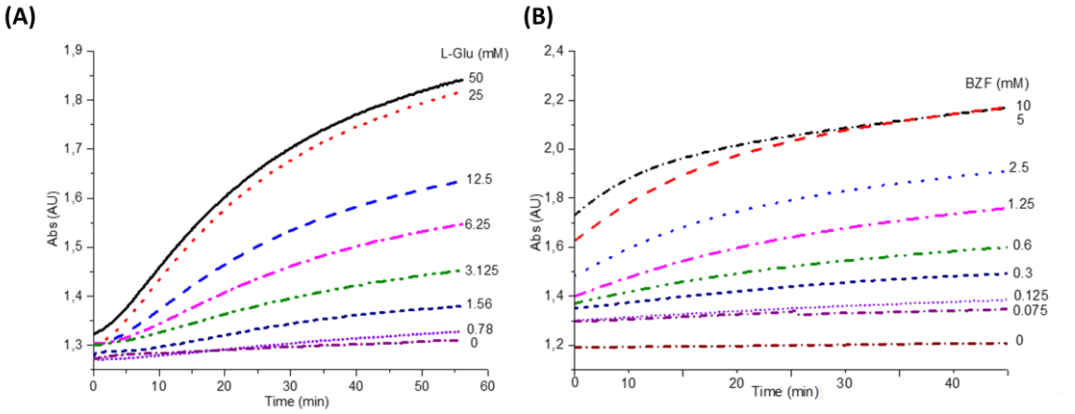


Figure 3.7 Typical raw data profile of NADH appearance at 340 nm obtained varying the **(A)** L-Glu concentration and **(B)** BZF concentrations. The initial rate (15 min) was used to calculate kinetic parameters using Graphpad Prism software.

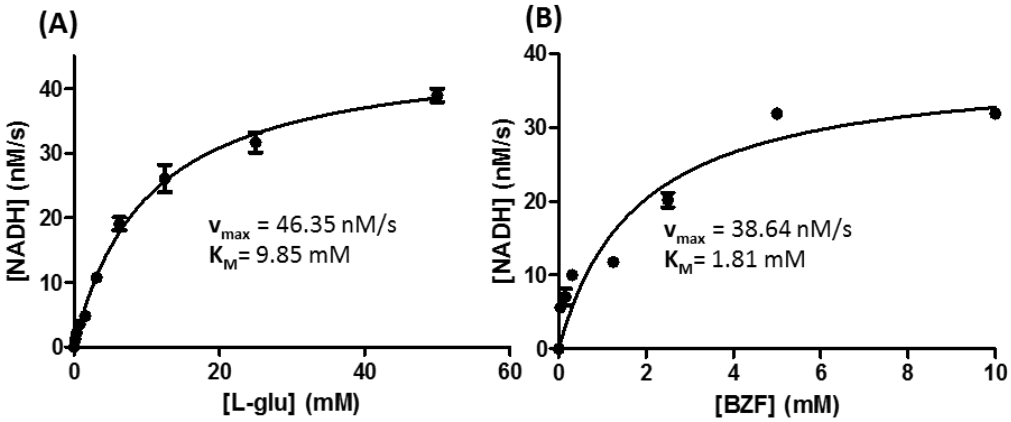


Figure 3.8 Michaelis–Menten plot of D-PhgAT for **(A)** the amino donor L-Glu and **(B)** the amino acceptor BZF. Data was analysed by non-linear regression using Graphpad Prism software.

Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
BZF	1.81 ± 0.57	0.39± 0.072	213.72± 0.49
HBF	NA	NA	NA
L-Glu	9.85 ± 0.62	0.46 ±0.011	46.65 ±0.32

Table 3.1. Summary of the kinetic parameters obtained using the D-PhgAT/AKG assay. k_{cat} values are in the order of sec. The same assay could not be applied (NA) for HBF due to high absorbance background at 340 nm.

3.3.2 Chiral HPLC method development (cHPLC)

The D-PhgAT/AKGDH assay presents many advantages, allowing a quick monitoring of kinetic parameters and enzyme activity. However, it cannot be used to fully characterise the enzyme. This is because the D-PhgAT/AKGDH assay directly monitors only the first half of the AT reaction and only indirectly D-Phg production. Furthermore, the high background absorbance does not allow the calculation of kinetic parameters for other substrates, such as HBF.

One of the major advantages of employing ATs is the possibility to produce chiral products. The D-PhgAT/AKGDH assay method does not allow the determination of enzyme enantioselectivity. As such, a stereospecific HPLC method was therefore developed to determine the chirality of the D-PhgAT reaction product. Initially, the protocol of Wiyakrutta S. *et al* was attempted, using a C18 column and a chiral mobile phase consisting of 2.4 mM copper acetate, 5 mM L-Pro and 3% Methanol (MeOH).[127] However, this did not give the expected results as the D- and L-Phg were found to co-elute (data not shown).

An alternative method, using a Chirobiotic T (Teicoplanin) column, was developed. This column had previously been used in our lab to resolve other amino acids.[148] Using a mobile phase of 0.25% triethylammonium acetate (TEAA): MeOH (50:50) and a flow rate of 1 mL min^{-1}, the D- and L-Phg enantiomers were found to show distinct retention times (4.3 min vs. 8.3 min). Using the developed method, the production of the only D-Phg enantiomer can be observed in > 99% enantiomeric excess (*ee*) for the D-PhgAT catalysed reaction (Figure 3.9), where the *ee* is defined as the absolute difference between the mole fractions of two enantiomers as described in 5.9.1.[149]

The exquisite D-PhgAT enantioselectivity is shown in Figure 3.9 were the HPLC trace for the D/L-Phg standards are overlaid with the HPLC trace for the D-PhgAT catalysed reaction. The specific formation of D-Phg for the D-PhgAT catalysed reaction can clearly be observed (Figure 3.9). Aliquots of single enantiomers of each amino acid (0-400 mM) were used to build a calibration curve to relate the area under the curve (AUC) with the amount of product formed over time (Appendix 8.6). Different enzyme concentrations (100 nM, 1 μM, and 10 μM) were tested at different time points in saturating conditions in order to

determine the range of linearity of the reaction and identify the ideal reaction conditions (Figure 3.10).

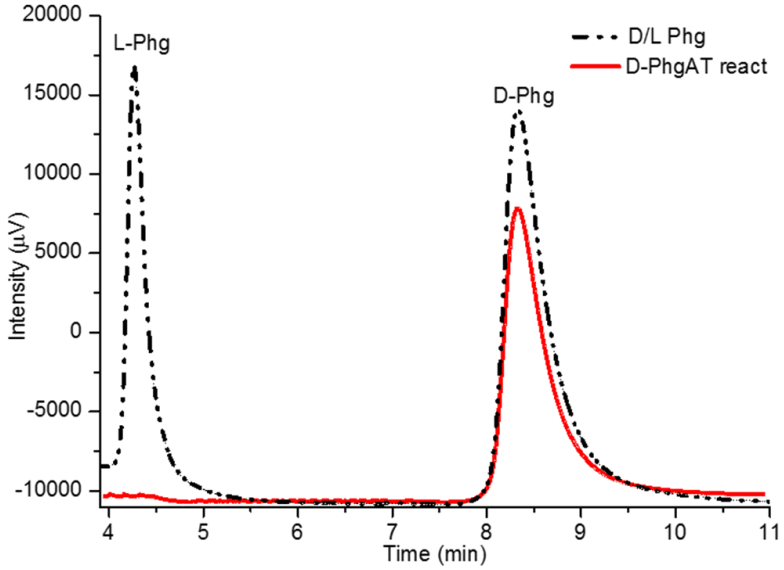


Figure 3.9 Chiral HPLC (cHPLC) analysis on a calibrated Chirobiotic T column at λ = 205 nm. The L- and D-Phg standards (10 mM, dash black line) are well resolved and the AT reaction carried out for 18 h at 37 °C (10 mM L-Glu, 10 mM BZF) shows the formation of only the D-Phg enantiomer (red line). The retention times are 4.3 min for L-Phg and 8.3 min for D-Phg respectively.

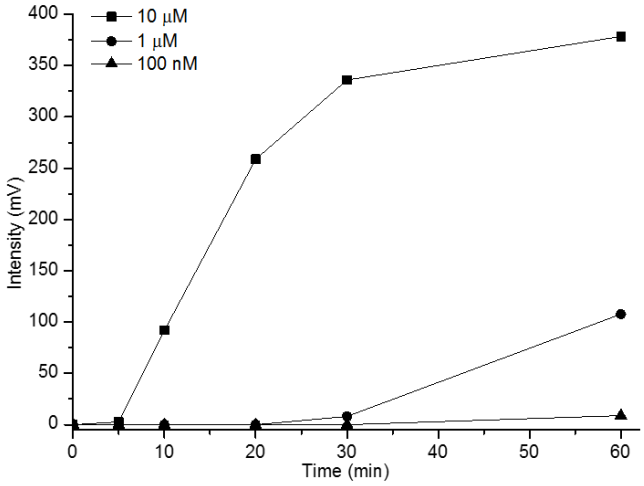


Figure 3.10 Reaction profile of three different enzyme concentrations (100 nM, 1 µM and 10 µM) at different time points at 37 °C under saturating conditions. Using D-PhgAT concentrations lower than 10 µM, the reaction is too slow to be monitored. Therefore 10 µM was selected as the enzyme concentration for kinetic analysis. As 15 min is in the linear range of the plot it was chosen as the reaction time.

Following the screening of these conditions, subsequent kinetic analyses were conducted with 10 μM of D-PhgAT at 37 °C for 15 min after which the reaction was quenched by diluting the reaction 40-fold in the chiral mobile phase. A typical chiral HPLC (cHPLC) profile of a kinetic experiment is shown in Figure 3.11 where a larger peak could be observed at increasing BZF concentrations. By relating the AUC with the amount of D-Phg/D-Hpg produced per unit of time, kinetic parameters could be calculated for each substrate, including HBF (Figure 3.11 B, Table 3.2).

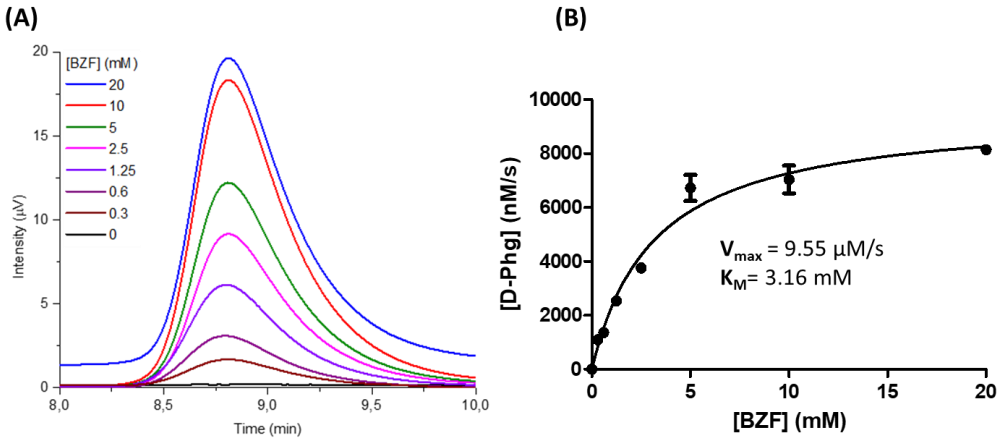


Figure 3.11 Representative chromatogram of D-PhgAT reactions at different BZF concentrations analysed by cHPLC. A bigger peak, corresponding to a larger amount of D-Phg produced is observed at increasing BZF concentrations. **(B)** Resulting Michaelis–Menten plot of D-PhgAT for BZF. Data was analysed by non-linear regression using Graphpad Prism software.

As observed with the D-PhgAT/AKGDH assay, the enzyme shows a high turnover number confirming its industrial potential. The enzyme catalytic efficiency (k_{cat}/K_M) of HBF is ~ 3-fold greater than the BZF suggesting a preference of D-PhgAT for the 4'-hydroxy derivative (HBF) (899 vs. 302.80 M^{-1} s^{-1}).

Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
BZF	3.16 ± 0.46	0.95 ± 0.056	302.80 ± 0.10
HBF	1.01 ± 0.07	0.91 ± 0.022	899.15 ± 0.27
L-Glu	26.17 ± 3.63	1.65 ± 0.12	63.21 ± 0.78

Table 3.2 Summary of the D-PhgAT kinetic parameters calculated using the Chirobiotic HPLC method.

3.3.3 Comparison of D-PhgAT assays

A comparison of the kinetic parameters calculated using both the D-PhgAT/AKGDH assay and the cHPLC method is shown in Table 3.3. All the calculated values are in the same order of magnitude. However, the values of the apparent K_M determined using the cHPLC method are higher than those calculated with the spectrophotometric approach for both substrates.

The apparent K_M for BZF of the cHPLC method (3.16 mM) is 1.7-fold higher than that determined using the coupled assay (1.81 mM). A larger difference is observed for the kinetic parameters calculated for L-Glu with a 3-fold difference observed for the K_M (9.85 mM with AKG assay vs. 26.17 mM with cHPLC).

AKG assay				**Chiral HPLC**		
K_M (mM)	**k_{cat} (s^{-1})**	**k_{cat}/K_M ($M^{-1}s^{-1}$)**	**Substrate**	**K_M (mM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_M (M^{-1} s^{-1})**
1.81 ± 0.57	0.39 ± 0.072	213.72 ± 0.49	**BZF**	3.16 ± 0.46	0.95 ± 0.056	302.80 ± 0.10
NA	NA	NA	**HBF**	1.01 ± 0.07	0.91 ± 0.022	899.15 ± 0.27
9.85 ± 0.62	0.46 ± 0.011	46.65 ± 0.32	**L-Glu**	26.17 ± 3.63	1.65 ± 0.12	63.21 ± 0.78

Table 3.3. Summary of D-PhgAT kinetic parameters towards natural substrate calculated with both spectrophotometric and HPLC approaches. All kinetic parameters have been calculated using Graphpad Prism software.

The AKGDH assay relies on the consumption of the transamination product of L-Glu, AKG, removing the ketone formed *via* a second enzyme and pushing the equilibrium towards product formation. This likely results in the determination of an apparent K_M lower than that calculated using the cHPLC approach, where the equilibrium is not altered by an additional enzyme. However, the parameters calculated with both approaches though different are comparable and for substrate comparison they could both be used.

3.3.4 Exploring amino donor promiscuity

In order to further characterise the D-PhgAT, substrate (BZF) conversion was assayed in the presence of all L- and D-amino acids and some industrially used amino donors (e.g. isopropylamine (IPA)). These were also tested to determine if a better amino donor could

be found to drive the AT reaction to completion. The amino donors were tested with BZF as amino acceptor as the product formation (D-Phg) can be easily monitored by cHPLC and the enantiopurity of the product can be assessed. Parallel reactions with 10 mM BZF and 10 mM amino donor were incubated for 18 h at 37 °C. Reactions were then quenched by diluting 40-fold into the mobile phase and analysed by cHPLC. According to the literature, no amino donors other than L-Glu are accepted by D-PhgAT, including L-Ala which has been shown to be accepted by the homologous D-PhgAT from *P. putida* (82% sequence identity).[127] The different substrate scopes of these two homologous enzymes cannot be explained by sequence alignment alone.[130] In contrast to reports in the literature,[127] the D-PhgAT used here displayed a much broader substrate scope, showing ability to use several L- and D-amino acids as substrates. The substrate scope is summarized in Figure 3.12, where BZF conversion (%) is normalized to L-Glu conversion which is the best amino donor.

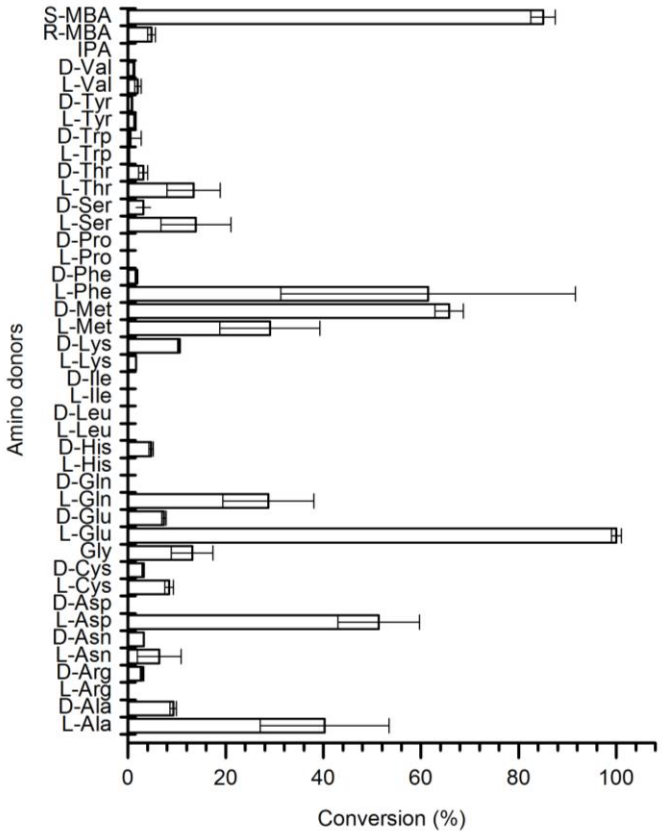


Figure 3.12 BZF conversion (%) when all L/D-amino acids are tested. All the reactions were carried out for 18 h in parallel. The percentage is an average of four experiments. The results are normalized by L-Glu conversion (70%).

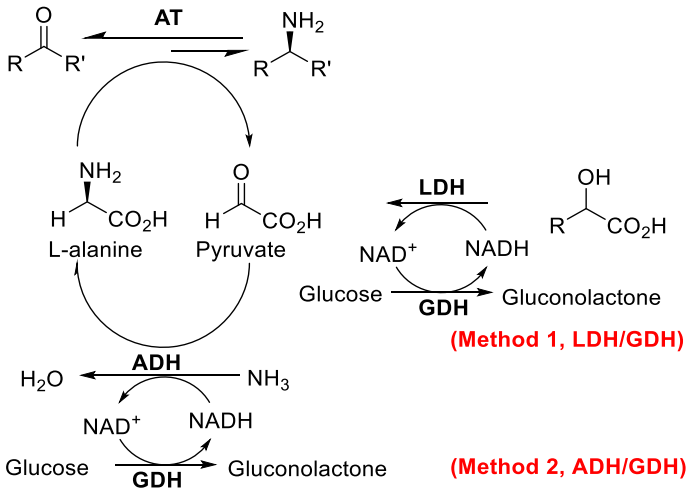
Regardless of whether L- or D- amino acids were used, the chirality of the final product remained unaffected, and the only detectable product was D-Phg. Interestingly, conversion was observed when D-Glu (7% conversion) was used as amino donor and with several D-amino acids such as D-Met (75%), D-His (4.7%) and D- aromatic amino acids such as D-Phe, D Tyr and D-Trp (1.7%, 0.46% and 0.8% respectively). The second best amino donor after L-Glu is (*S*)-α-methylbenzylamine (MBA) with 90% of conversion with respect to L-Glu conversion. The (*R*)-MBA enantiomer is still accepted by the AT, but with a much lower conversion (4.8%). D-PhgAT shows high activity towards L-Asp with a substrate conversion of 45%. Sulfur containing amino acids are accepted by D-PhgAT with 40% conversion when L- /D-Met are used as amino donors and 10% with L-Cys.

3.3.4.1 Non-natural amino donors

The application of ATs for the industrial synthesis of optically pure chiral amines still presents a number of challenges. One of the major drawbacks which has hampered the widespread use of ATs is the unfavourable reaction equilibrium. However, use of a large excess of the amino donor (if inexpensive) or removing the ketone by-products can be sufficient to shift the equilibrium towards product formation and achieve a satisfactory yield.[102,150] The use of a large excess of amino donor combined with the by-product removal is often applied in small/medium scale reactions. The most widely employed system is the lactate dehydrogenase (LDH)/ glucose dehydrogenase (GDH) system (Figure 3.13, method 1) which relies on the removal of the pyruvate by-product by using LDH, an enzyme that requires the expensive NADH cofactor, recycled using GDH. An alternative method, the alanine dehydrogenase (ADH)/GDH system (Figure 3.13, method 2) is used to both recycle the amino donor and shift the equilibrium towards product formation. However, both approaches require the use of expensive cofactor-dependent enzymes which are not sustainable for large scale reactions. Furthermore, the requirement of several enzymes could result in compromised conditions that are not optimal either for the AT or the enzyme(s) applied in the secondary irreversible reactions, resulting in sub-optimal enzyme performance in terms of activity and enantioselectivity. For this reason, industrially, the preferred amino donor is isopropylamine (IPA) since its keto co-product, acetone, is highly

volatile and can be easily removed by in *situ*-evaporation (Figure 3.13, method 3).[101,151] A summary of these approaches is shown in Figure 3.13.

(A) Favored for small/medium-scale reactions



(B) Favored for large-scale processes

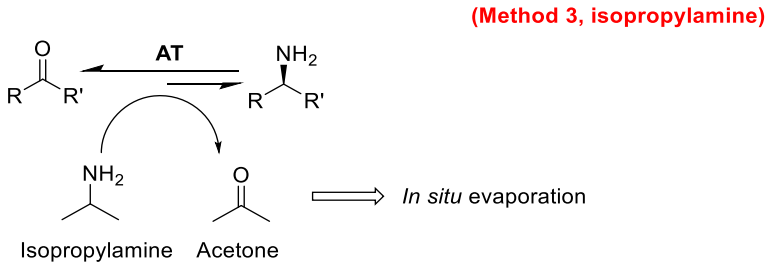


Figure 3.13 Most common approaches applied to overcome the unfavourable reaction equilibrium in TA applications. **(A)** Commonly, L-Ala is used as the amino donor as its keto by-product, pyruvate, can be removed from the reaction by applying the LDH. An additional enzyme, the GDH, is used in order to recycle the expensive NADH cofactor and therefore lower the amount of NADH required in the reaction. Alternatively the use of ADH has the double effect of driving the reaction towards product formation and recycling the L-Ala substrate. **(B)** In large scale processes the IPA method is preferred as the amino donor as its keto by-product acetone can be easily removed by in *situ*-evaporation. Adapted from Gomm *et al.* and Truppo *et al.*[101,151]

All these approaches have their limitations as they either involve additional enzymes, a costly cofactor or challenging by-product removal. The IPA method (Figure 3.13, method 3) has not been applied successfully for conversion of ketones with particularly challenging positions of equilibria.

3.3.4.2 *Ortho*-xylylenediamine (OXD)

In the last decade there has been significant interest in finding amino donors that could solve the issues related to unfavourable equilibrium position. This has led to the development of the so called 'smart' amine donors. The term 'smart' amine donor refers to amine donors whose keto by-products spontaneously and irreversibly transform into a different, more stable, compounds, shifting the equilibrium towards product formation.

Amongst the 'smart' amine donors, the commercially available *ortho*-xylylene diamine (OXD) developed by Prof. N. Turner (University of Manchester) is of great interest as its cyclic imine by-product spontaneously converts into the more stable isoindole, removing the component from the system.[152] The spontaneous polymerization of the isoindole causes the production of an insoluble purple derivative. The formation of a purple polymer has the double effect of displacing unfavourable reaction equilibria towards product formation whilst providing a substrate-independent high-throughput (HTP) screening method (Figure 3.14). Turner and co-worker successfully applied the OXD with a wide range of commercially available (*R*)- and (*S*)-ATs (Codexis) and amino acceptors suggesting the wide applicability of the 'smart' amine donor.[152] However, higher conversions have been observed with (*S*)-AT rather than (*R*)-AT, suggesting a preference for (*S*)-ATs.

Changes in the UV-Vis spectrum were observed when OXD was tested as amino donor for D-PhgAT, with a first shift towards the external aldimine and then to the PMP form (Figure 3.15). Overtime, an extra peak around ~ 550 nm was observed, suggesting the formation of the more stable isoindole product.

The OXD amino donor was also successfully used in a solid phase assay. The formation of a purple/black precipitate was observed only in cells expressing the D-PhgAT (Figure 3.16). However, colonies turned black in a short period of time (a few hours), suggesting that it could not be used for screening in a directed evolution approach. This is thought to be due to high background combined with the high sensitivity which precludes its use as a quantitative assay.

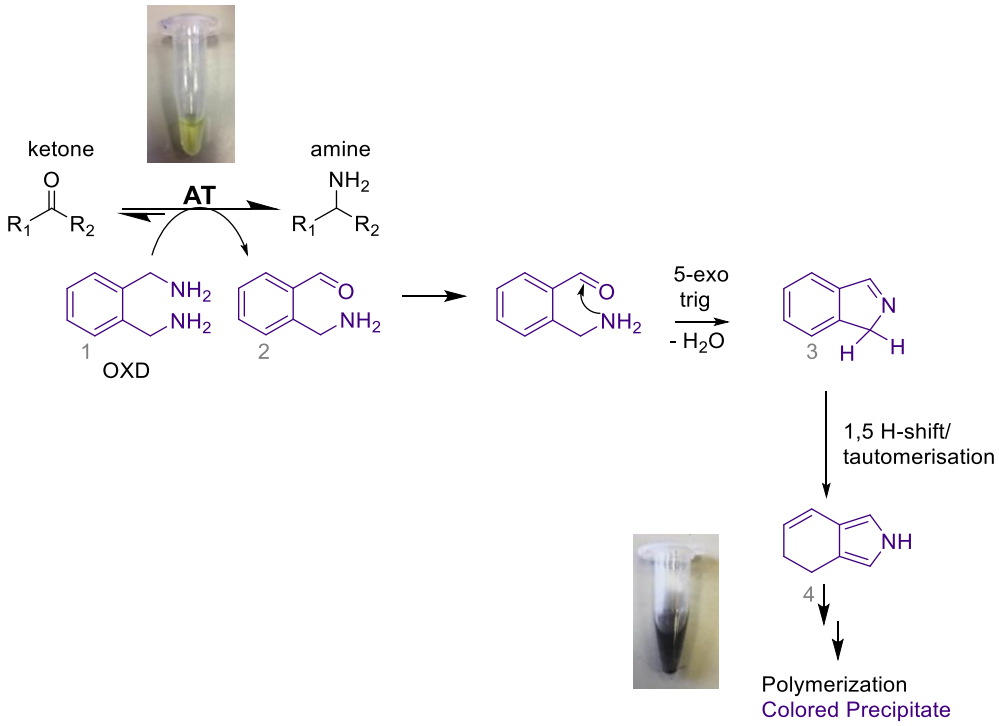


Figure 3.14 General transamination reaction when OXD (1) is used as amino donor mode of action. The OXD keto by-product (2) spontaneously cyclizes into (3) that tautomerizes into the more stable isoindole (4). Compound (4) spontaneously polymerizes causing the production of a purple precipitate.

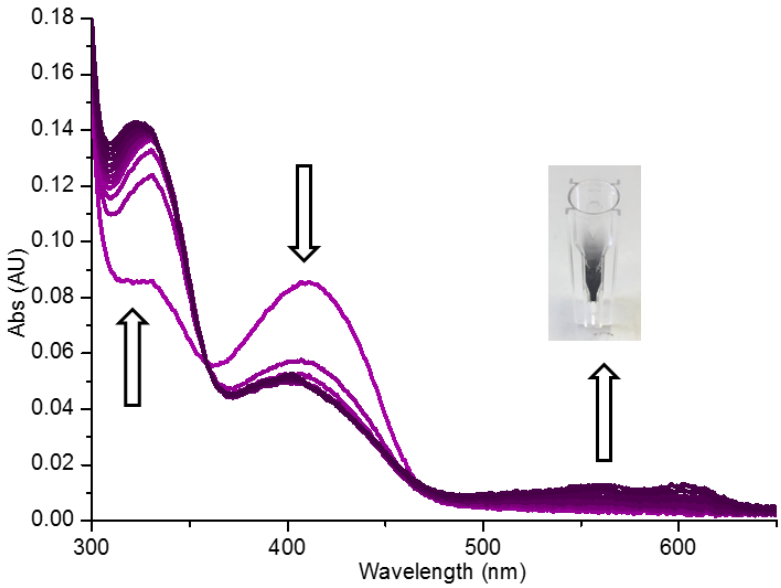


Figure 3.15 D-PhgAT UV-Vis studies with OXD as amino donor at 1 mM. Over time the internal aldimine peak at 411 nm disappears and the PMP formation at ~ 311 nm is observed. An additional peak at ~ 550 nm corresponding to the isoindole formation is observed suggesting the progression of the AT reaction. The formation of a purple precipitate was observed after ~ 1 hr of reaction.

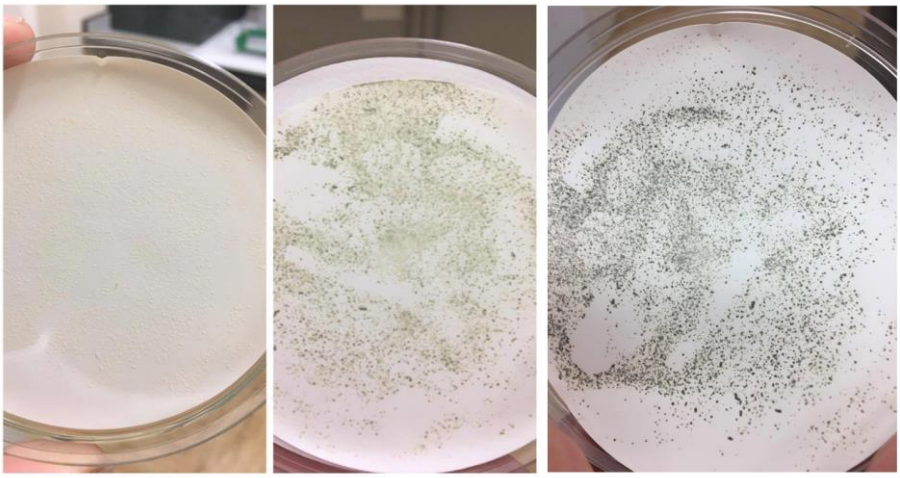


Figure 3.16 Colony-based screen with OXD. Cells expressing the *dpgA* gene turn dark in colour after ~ 1 hr (middle). Cells lacking the *dpgA* gene remain colourless (left). After leaving the reaction ON the purple colour becomes more intense (right).

A conversion comparison using the natural (L-Glu) and 'smart' amine donor (OXD) was carried out with 10 mM BZF over time (Figure 3.17) and analysed by cHPLC. Surprisingly L-Glu showed a much higher conversion overtime (93% for L-Glu vs. 30% for OXD). Thus, the 'smart' amino donor does not present any beneficial effect in either reaction time or percentage conversion for the D-PhgAT reaction. The 'poor' conversion observed with OXD as amino donor could be due to the reported preference of OXD for (*S*)-ATs as D-PhgAT is an unusual 'stereoinverting' AT.

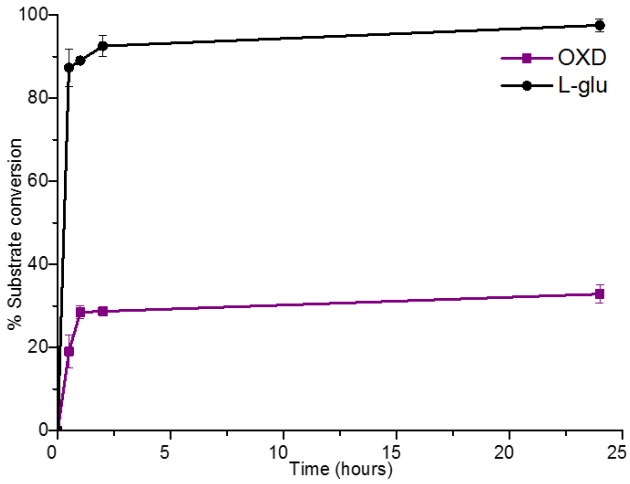


Figure 3.17 Time course comparison when OXD (purple line) and L-Glu (black line) are used as amino donor. The substrate conversion (%) was determined by cHPLC. OXD and L-Glu were used at a final concentration of 100 mM.

3.3.4.3 Amino benzenes as 'smart' amino donors

Recently, Prof. Helen Hailes (UCL) developed a range of 'smart' amine donors with a different mode of action compared to OXD.[153] An example is shown in Figure 3.18, with 2-(4-nitrophenyl) ethan-1-amine (4-NPEA) (3). The aldehyde, (4), generated by the 4-NPEA (3) transamination reaction, reacts with the amine donor (3) forming the imine (5), which, after tautomerization, gives the conjugated red precipitate (6). The amine produced by the TA reaction (2) can also react with the amine donor 3 contributing to the observed colour. The use of a high excess of the amine donor should minimize this background colour.

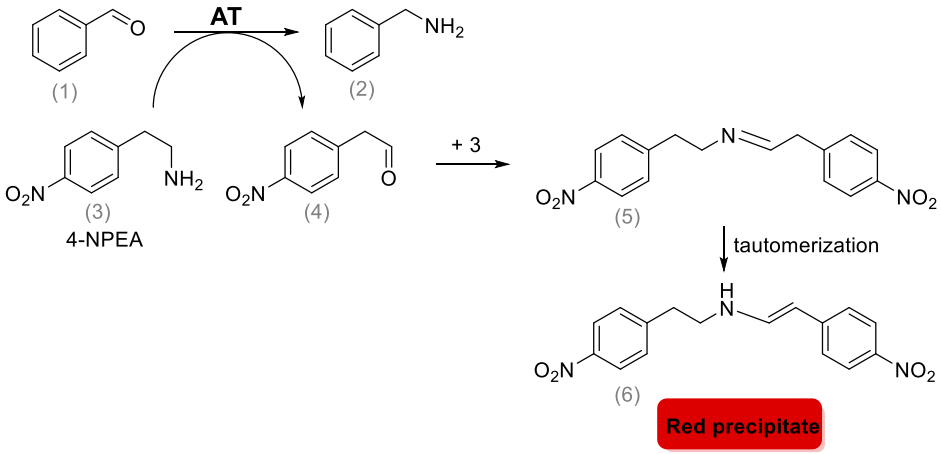


Figure 3.18 Mechanism of action of the 4-NPEA (3) amine donor. The 4-NPEA (3) conversion into aldehyde (4) and basic work up and deprotonation yields a highly conjugated structure (6) with absorbance in the UV-region.

Prof. Helen Hailes and co-workers reported conversion using well characterised (*S*)-ATs such as *Chromobacterium violaceum* AT (Cv-AT) and *Pseudomonas putida* AT (Pp-TA) and (*R*)-AT such as *Klebsiella pneumoniae* KPN_00799 (Kp-TAm) showing the applicability of the assay to both (*R*)- and (*S*)-selective ATs.

As the presence of the 4-nitroaryl electron withdrawing group (EWG) enhances the tendency to form the enamine, other commercially available amine donors with EWG were explored such as 4-(2-aminoethyl) benzonitrile (4-AEB) and 5-nitro-2, 3-dihydro-1H-inden-2-amine (2,3-DHIA). Thus, 4-NPAE (3) ,4-AEB (7) and 2,3-DHIA (8) were tested with different amino acceptors such as BZF, AKG, benzaldehyde (BZALD) and n-butyraldehyde (NBA) as reported in the literature (Figure 3.19). The formation of a very pale orange/red precipitated could be observed with amine donors (3) and (8) with BZF as amino acceptor in

comparison to the control (A3, A4; A5, A6) (Figure 3.19). However none of the amino donor screened seems suitable for the D-PhgAT reaction as no colour formation was observed.

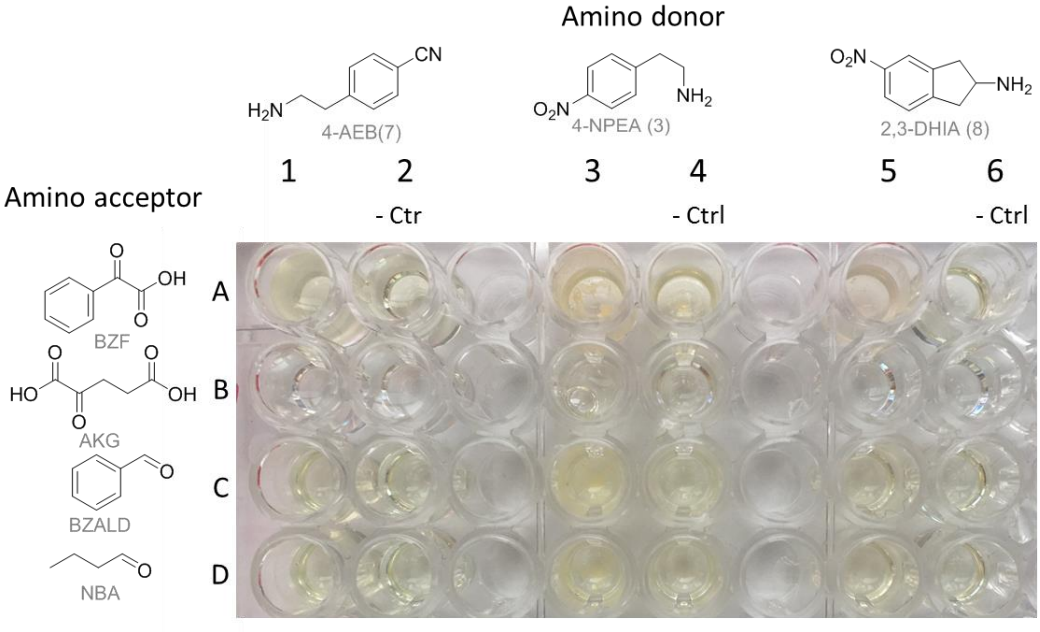


Figure 3.19 Screening of the amine donors 4-AEB (7) (lane 1, 2), 4-NAE (3) (lane 3, 4) and 2, 3-DHIA (8) (lane 5, 6) with different amino acceptors such as BZF (lane A), AKG (lane B), BZALD (lane C), NBA (lane D) with D-PhgAT (0.5 mg mL^{-1}). Amine donor were used at a final concentrations of 25 mM with amino acceptors at 10 mM, as suggested in the literature.[153] For each reaction a control with no enzyme was performed (lane 2, 4, 6). Reactions were incubated for 18 h at 37 °C. A very pale red colour is observed for lane A3 and A5 with 4-NAE and 2, 3 DHIA and BZF.

3.4 Biotransformation optimization/development

Amongst all the tested amino donors, the natural substrate, L-Glu, was found to be the best amino donor and therefore was used in the following studies. The biocatalytic reaction was scaled up to a gram scale with BZF (1 g) in order to prove the applicability of the enzyme at an industrial level. A large excess of L-Glu (500 mM) was used in the reaction to shift the equilibrium towards product formation. As the high pH (9.5) is crucial for the enzyme activity the highest L-Glu concentration achievable was 500 mM. Therefore 1 g of BZF (~ 133 mM) and 500 mM L-Glu were added to the reaction mixture (50 mL) and the pH was adjusted before addition of the catalyst.

The catalyst concentration is a crucial factor for the biotransformation, therefore the reaction was carried out at 3 different D-PhgAT concentrations (0.2 mg mL^{-1}, 0.4 mg mL^{-1} and 1 mg mL^{-1}) to identify the optimal conditions. Samples were quenched at different

time points from 15 min to 24 h. Reactions were quenched by diluting 40-fold into the cHPLC mobile phase and analysed by cHPLC. The results are summarized in Figure 3.20 showing the substrate conversion at different time points for each enzyme concentration.

A 93% BZF conversion is observed after just 3 h of reaction with an enzyme loading of 1 mg mL^{-1}. A 60% BZF conversion and 50% BZF conversion is observed at 0.4 mg mL^{-1} and 0.2 mg mL^{-1} D-PhgAT, respectively. This yield is better than the previously reported 83% conversion with isolated enzyme.[127]

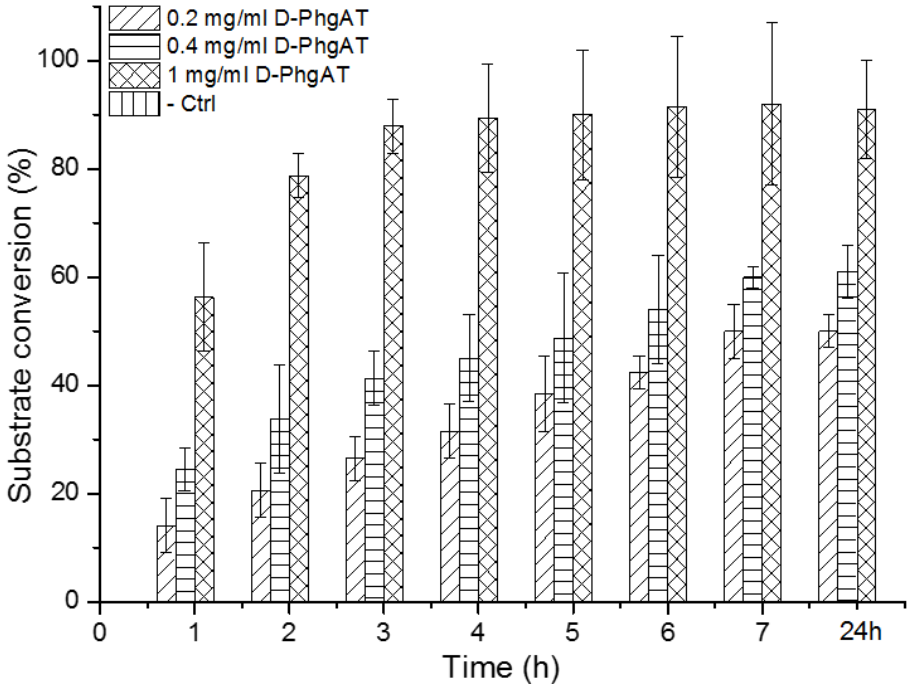


Figure 3.20 Large scale biotransformations at different D-PhgAT concentrations (0.2, 0.4 and 1 mg mL^{-1}). A control with all the reactants apart from the enzyme was carried out at each time point in order to prove that the reaction is enzyme catalysed. Reactions were quenched by diluting the sample (40-fold) into the mobile phase (50:50 MeOH: 0.025% TEAA). Samples were analysed by cHPLC at each time point in triplicate.

Zhou Y. *et al.* [154] recently described the engineering of three recombinant *E. coli* strains that co-expressed four, seven and nine enzymes including *P. stutzeri* D-PhgAT as the final step. Thanks to these whole cell biotransformation approaches, the production of a wide range of D-Phg derivatives such as 4-bromo-, 4-fluoro-D-Phg is achieved in one pot from racemic mandelic acid, styrene and L-Phe starting materials. Using this approach, Zhou Y. *et al.* reported a yield of 93% for D-Phg which is consistent with our result.

This reaction scale-up proves that D-PhgAT can be applied to a large-scale biosynthesis of D-Phg without requiring any additional engineering for the synthesis of D-Phg. An excess of amino donor L-Glu (~ 20-fold K_M) is enough to drive the reaction towards product formation.

3.5 Exploring amino acceptor promiscuity

Having established the best operative conditions for the D-PhgAT bioconversions in terms of both amino donor and enzyme concentrations, and demonstrated its efficiency for the natural catalysed reaction, screening of amino acceptors was carried out. This was done to identify other high value products that could be potentially synthesized using this promising biocatalyst.

Non-proteinogenic D-amino acids such as D-phenylalanine (D-Phe), D-tyrosine (D-Tyr) and D-tryptophan (D-Trp) are high value compound whose efficient synthesis would be extremely advantageous.[155] Current biosynthetic methods still present several limitations such as either the requirement of multiple enzyme (e.g. L-amino acid ATs and a racemase or hydantoinase–carbamoylase system) or an expensive starting material (D-amino acids for D-AT).[125,155]

D-PhgAT was therefore tested for its ability to use BZF, HBF, phenylpyruvic acid (PPA), 4-hydroxyphenylpyruvic acid (HPPA), and indole-3 pyruvic acid (IPyA) as amino acceptors in order to yield D-Phg, D-Hpg, D-Phenylalanine (D-Phe), D-Tryptophan (D-Trp) and D-Tyrosine (D-Tyr) respectively. A scheme of the tested reactions is shown in Figure 3.21. Surprisingly D-PhgAT was able to use all the aromatic compounds tested as amino acceptors, producing enantiopure high value D-amino acids. A summary of the cHPLC traces of the enantiopure synthesized D-amino acids is shown in Figure 3.22. These D-aromatic amino acids have a range of applications. D-Phe, for instance, has anti-depressant and analgesic activities.[156] Furthermore D-Phe derivatives are high value building blocks for a wide range of pharmaceuticals from antibiotics to antidiabetics.[156] D-Trp has been shown to be a product of probiotic bacteria, influencing the gut microbiome and chronic immune diseases.[157] D-Tyr is a chiral precursor of potent inhibitors of human non-pancreatic secretory phospholipase A2 (IIa) with anti-inflammatory activity.[158]

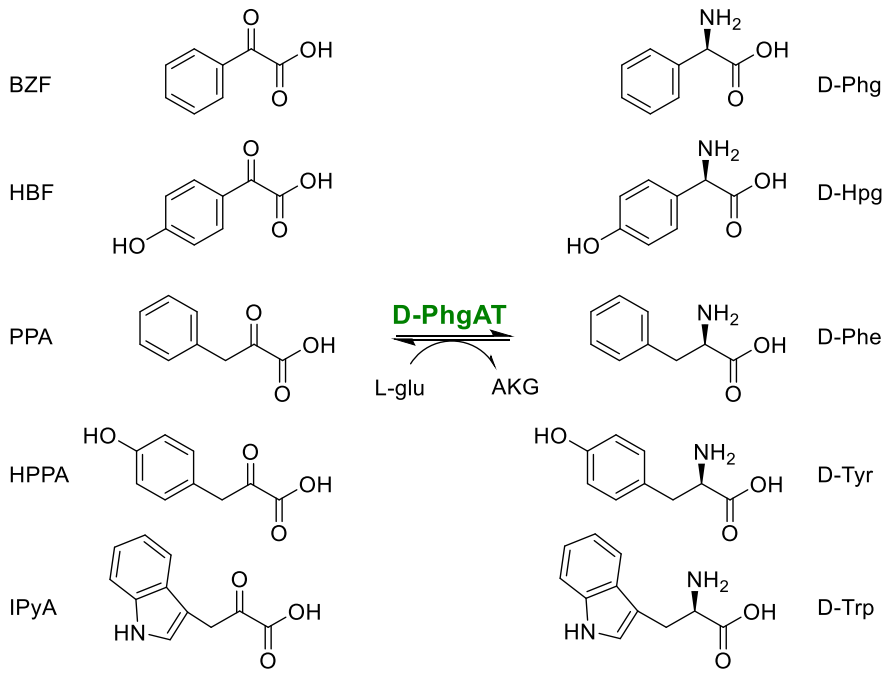


Figure 3.21 D-PhgAT reactions for the synthesis of aromatic D-amino acids. Reactions were analysed by cHPLC which was calibrated with a single enantiomer of each amino acid and used to build calibration curves.

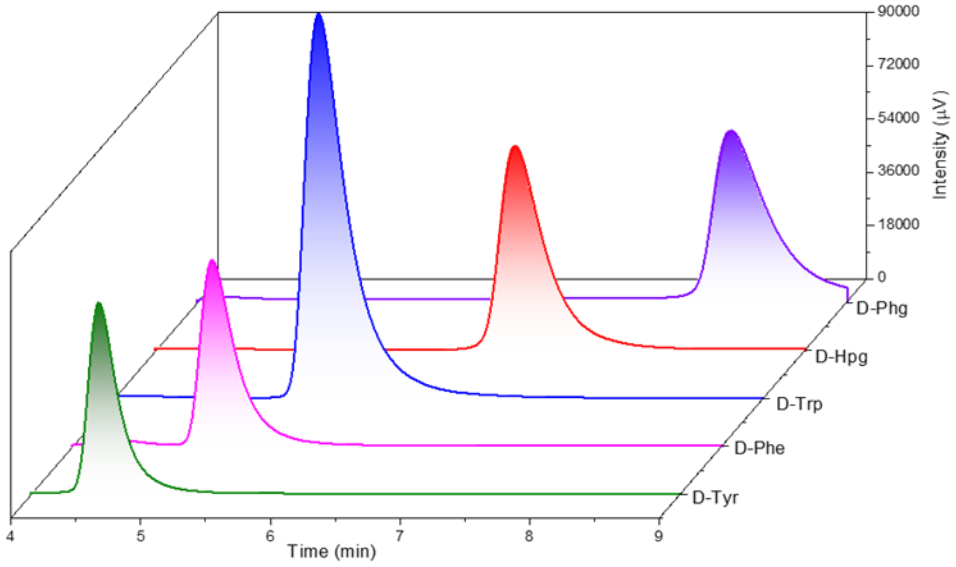


Figure 3.22 cHPLC traces of D-PhgAT production of D-Phg, D-Hpg, D-Trp, D-Phe and D-Tyr at λ = 205 nm. Reactions were carried out for 15 min and quenched by diluting the sample into the mobile phase (50:50 MeOH: 0.025% TEAA). Retention times are 4.6 min (D- Tyr), 5.4 min (D-Phe), 6.1 (D-Trp), 7.6 min (D-Hpg) and 8.3 min (D-Phg).

Kinetic parameters and % conversion were calculated for all amino acceptors up to 100 mg scale (Table 3.4).

Substrate	Product	Conv /%	*ee* /%	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
BZF	D-Phg	93	>98	3.16 ± 0.46	0.95 ± 5.6 *10^{-2}	302.80 ± 0.10
HBF	D-Hpg	50	>98	1.01 ± 0.07	0.91 ± 2.2 *10^{-2}	899.15 ± 0.27
PPA	D-Phe	57	>98	6.91± 0.99	0.09 ± 8.5 *10^{-3}	14.49± 0.819
HPPA	D-Tyr	15	>98	9.24 ± 2.4	5.95*10^{-4}± 2.34 *10^{-4}	0.0616 ± 0.011
IPyA	D-Trp	30	>98	0.39 ±0.19	0.10 ±1.4 *10^{-2}	262.12 ± 0.13

Table 3.4 Summary of product formation (%) and kinetic parameters of D-PhgAT for all the amino acceptor obtained using the cHPLC method. Kinetic parameters were calculated at 100 mM L-Glu at different amino acceptor concentrations (0-40 mM) with a reaction time of 15 min except for HPPA whose reaction was quenched after 30 min. Reactions were quenched by diluting the sample (40-fold) into the mobile phase (50:50 MeOH: 0.025% TEAA). For the percentage of conversions, reactions were carried out at 100 mg scale as described in Material and Methods. The L-Glu concentration used was 500 mM to drive the reaction towards completion.

The highest substrate conversion was observed with BZF as substrate (93%). However, good conversion was observed for PPA, with a 57% production of D-Phe, achieved after just 3 h of reaction. The lowest substrate conversion was observed for HPPA (15%). This is consistent with the observed catalytic efficiency (k_{cat}/K_M) of D-PhgAT for HPPA, which is dramatically reduced compared to the other substrates (~ 5000-fold lower than BZF).

According to the calculated kinetic parameters the best substrate (k_{cat}/K_M) is HBF, followed by BZF, IpyA, PPA and finally HPPA. Unsurprisingly, a decrease in the reaction rate (k_{cat}) is observed for the non-natural amino acceptors. Thus, in order to apply these bioconversions on an industrial scale, approaches such as directed evolution/site directed mutagenesis will be necessary. Alternatively, approaches such as enzyme immobilization or biphasic systems could be applied to overcome problems such as substrate inhibition.[159] As the position of the equilibrium is particularly challenging, especially for the 4-hydroxylated derivative of PPA, a range of 'smart' amine donors could also be tested to check improvement in percentage conversion.

3.6 High resolution structural studies of D-PhgAT

Determination of the 3D structure of D-PhgAT is fundamental for allowing an understanding of the molecular basis of the enzyme enantioselectivity and identifying residues crucial for substrate catalysis and conversion. These studies underpin future directed evolution/site-directed mutagenesis studies and optimisation of D-PhgAT as a biocatalyst. To this end, it was decided to determine the 3D structure of D-PhgAT from *P. stutzeri* ST-201 by X-ray crystallography. Previous attempts to solve the D-PhgAT structure are reported in the literature; a PDB file deposited under 2CY8 is available. However, this model is of 'poor' quality as it does not contain the PLP-cofactor and residues around the active site are not modelled, presumable due to their flexibility in the absence of the cofactor and ligand.[131]

In collaboration with Dr Jon Marles-Wright (Newcastle University), crystal trials with purified D-PhgAT (See paragraph 5.13.1) were performed using a broad range of conditions from commercially available screens to identify hit conditions (PEG/ION, MIDAS, STRUCTURE, JCSG+ etc.). These screens were performed *via* vapour diffusion in 96 well plates format prepared with a Gryphon robot. Crystals were observed under a handful of different conditions (C3, H10 PEG/Ion, B8 JCSG$^+$, H4 MIDAS). These specific conditions were taken forward for further optimization.

Following optimisation, suitable conditions were found to give monoclinic crystals with the ideal morphology. These were generated by vapour diffusion *via* the hanging drop method at room temperature (25°C) using 0.2 M ammonium citrate, polyethylene glycol 15% (w/v) (PEG) 3350 (PEG/ION C3), pH 7 and a protein (9.1 mg mL^{-1}): precipitant ratio of 1:1 with the D-PhgAT in 50 mM Tris, 150 mM NaCl, 50 µM PLP, pH 9.5. These crystals were fished, frozen in cryoprotectant and sent to the Diamond Light synchrotron, Harwell for X-ray diffraction analysis.

Several diffraction-quality datasets were collected with the highest resolution being 2.0 Å. Solution of the D-PhgAT structure from the electron density data was carried out by Dr Jon Marles-Wright (Newcastle University) by molecular replacement using the incomplete D-PhgAT structure with PDB code 2CY8. Unfortunately, the structure lacked the PLP cofactor in the active site Figure 3.24). The lack of PLP in both structures (2CY8 and

ours) is probably caused by the buffer used (Tris) that, interacting with the cofactor, removes the PLP from the protein (Figure 3.23).

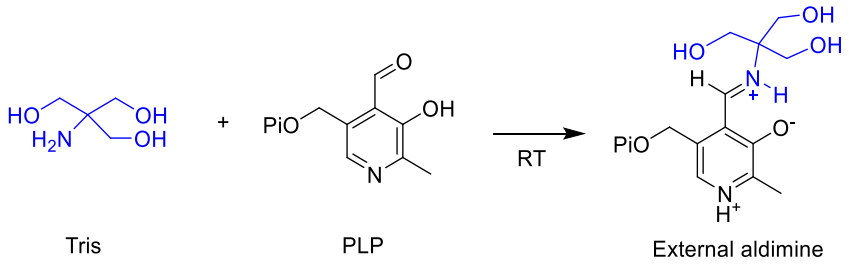


Figure 3.23 Mechanism of interaction between PLP and Tris. The primary amine of Tris reacts with the cofactor forming an external aldimine, removing the PLP from the enzyme which results in an *apo*-structure.

A second screen was performed, changing the protein buffer to 0.1 M CAPS, 150 mM NaCl, 50 µM PLP, pH 9.5. Yellow crystals were produced by adding 1 µL of 9.1 mg mL^{-1} D-PhgAT (in 0.1 M CAPS, 150 mM NaCl, 50 µM PLP, pH 9.5) to 1 µL of 0.1 M Tris, 0.2 M $MgCl_2$, 10% (w/v) PEG 8000, pH 7.5 (Figure 3.25). Crystals were cryoprotected with a solution containing 0.1 M Tris, 0.2 M $MgCl_2$, 10% (w/v) PEG 8000, pH 7.5, 20% (v/v) PEG 200 and then frozen by immersion in liquid nitrogen. Several diffraction-quality datasets were collected (Diamond Light) with the highest resolution of 2.25 Å undergoing full refinement to generate structural information (Appendix 8.7).

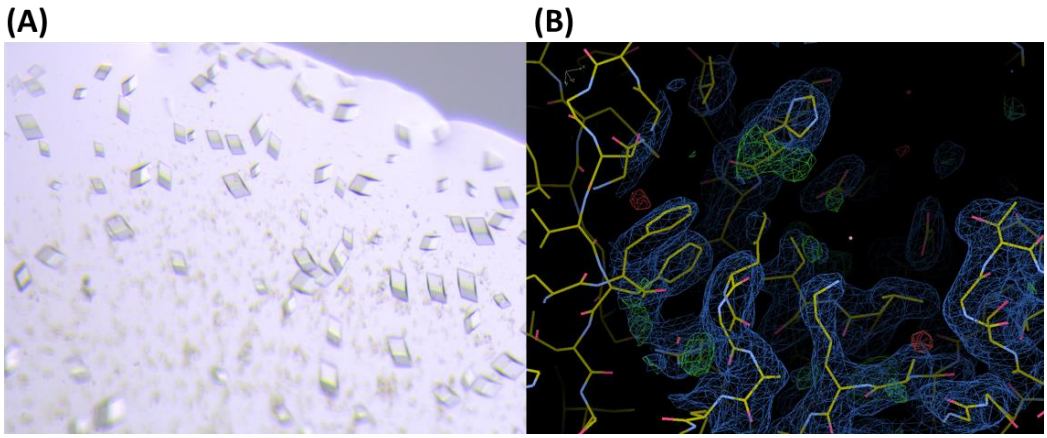


Figure 3.24 (A) Crystals morphology of the first attempt at obtaining diffraction quality D-PhgAT crystals. Conditions: 0.2 M ammonium citrate, 15%PEG 3350, pH 7 (Peg ION C3) and a protein (9.1 mg mL^{-1}): precipitant ratio of 1:1 with the D-PhgAT in 50 mM Tris, 150 mM NaCl, 50 µM PLP, pH 9.5. **(B)** Electron density map of the active site lacking the PLP cofactor contoured at 1.5 σ.

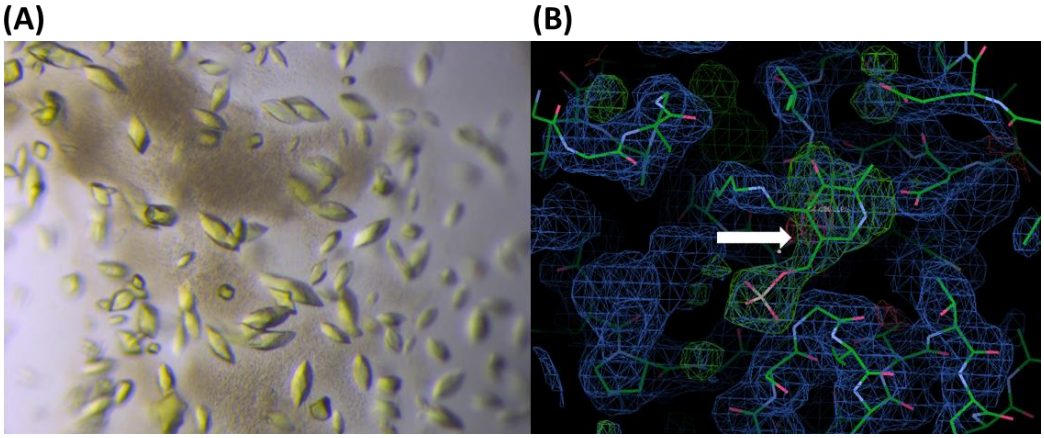


Figure 3.25 (A) Crystal morphology of the second attempt of obtaining diffraction quality D-PhgAT crystals by changing the D-PhgAT purification buffer. Yellow crystals were produced by adding 1 µL of 9.1 mg mL^{-1} D-PhgAT in 0.1 M CAPS, 150 mM NaCl, 50 µM PLP, pH 9.5to 1 µL of 0.1 M Tris, 0.2 M $MgCl_2$, 10% (w/v) PEG 8000, pH 7.5. **(B)** Electron density map of the active site containing the PLP cofactor contoured at 1.5 σ.

Solving of the D-PhgAT structure was performed by Dr. Jon Marles-Wright (Newcastle University) using the same procedure describe above, using the 2CY8 structure as a template. The refined protein structure is deposited under the PDB code 6G1F. The final refined model contains three homodimers of D-PhgAT with the PLP cofactor present as an internal aldimine with residue Lys269 in each chain (Figure 3.26).

Overall, our D-PhgAT structure conforms to the type I PLP-dependent fold seen in the group III of the aspartate AT (AspAT) family (Figure 3.27 A). The monomeric D-PhgAT structure can be subdivided into two distinct domains: a small discontinuous domain comprising the residues 1-72 and 336-453 (N- and C-terminal part of the polypeptide chain); and a large domain formed by residues 73 to 335 (Figure 3.27 B). The N-terminal portion of the small domain comprises a kinked α-helix followed by a three-stranded antiparallel β-sheet, while the C-terminal part consists of an α-helix followed by two antiparallel β-strands, an extended loop interspersed with a short α-helix that continues into a longer α-helix. This is followed by a β-strand that extends the N-terminal β-sheet and an additional β-strand which extends the C-terminal β-sheet, the structure finishes with a α-helix opposed between to two other major helices in this domain (Figure 3.27 B).

The large domain consists of a central seven-stranded β-sheet, with five parallel strands and two in an anti-parallel orientation. The β-sheet is connected by α-helices, which harbour the cofactor binding site and main dimerization interface (Figure 3.27 B).

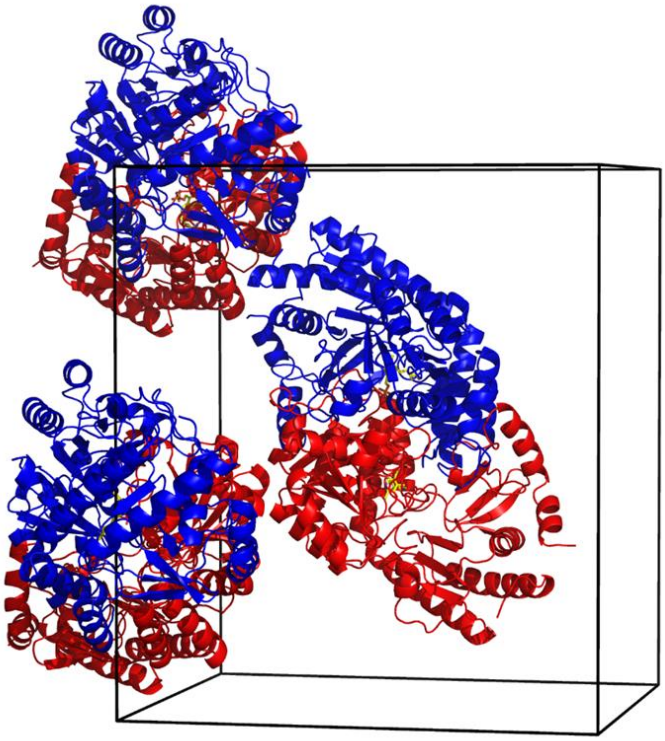


Figure 3.26 Final refined D-PhgAT model containing three homodimers of D-PhgAT with the PLP cofactor present as an internal aldimine with residue K279 in each chain. PLP is shown as sticks and highlighted in yellow. Monomer A is highlighted in red while monomer B is in blue.

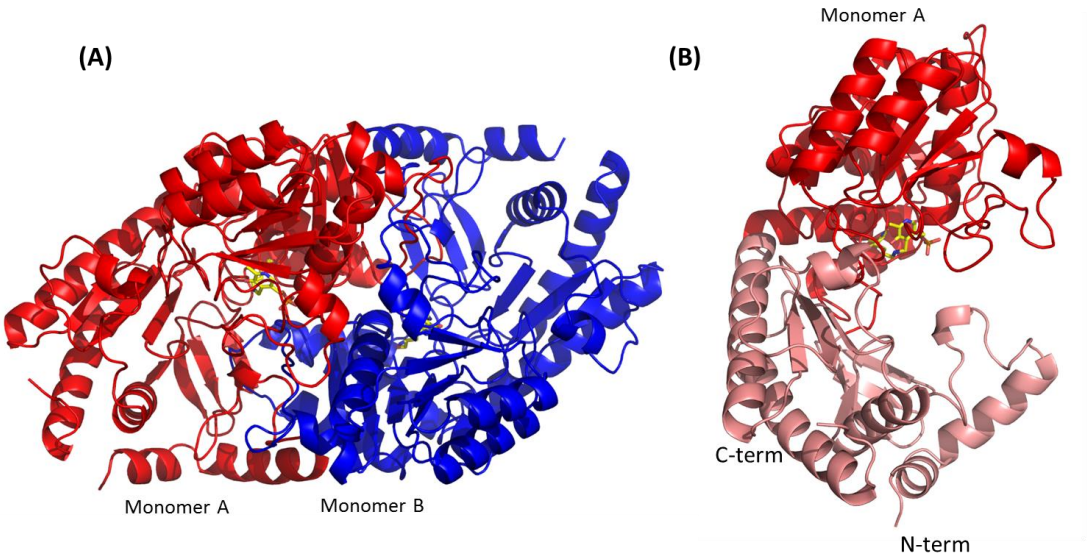


Figure 3.27 (A) Crystal structure of the D-PhgAT internal aldimine at 2.25 Å resolution. Monomer A is shown in red, monomer B in blue. **(B)** Large and small domain of monomer A are highlighted in bright and pale red. PLP is shown as sticks and highlighted in yellow.

The PLP-binding site and the active site of D-PhgAT is located at the dimer interface with the PLP binding lysine (Lys269) located in the loop connecting strands β9 and β10. The aromatic ring of PLP is sandwiched between Val243 and Tyr149. The phosphate group of the cofactor is anchored by a series of hydrogen bonds with Ser121, Gly122, and Thr123 from one monomer and Ser121, Glu124, Phe304 and Thr313 from the other, directly or through water molecules.[118] Residues involved in cofactor binding/stabilization are highlighted in Figure 3.28.

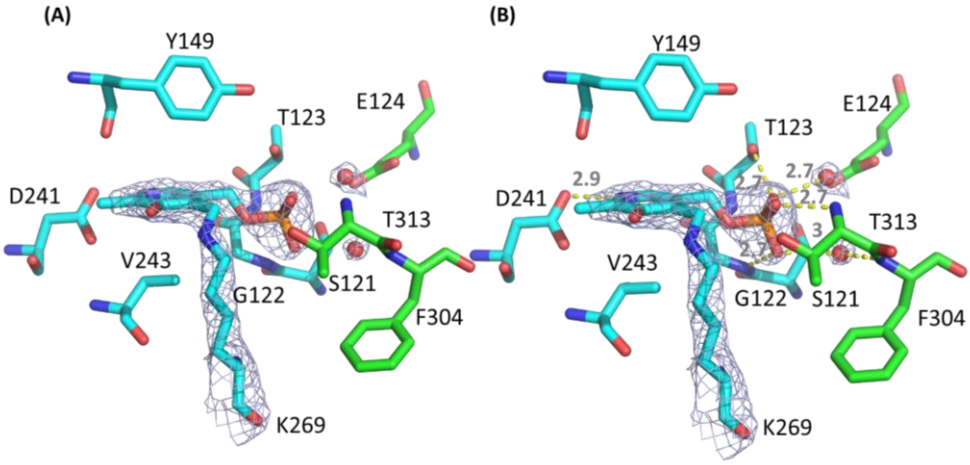


Figure 3.28 (A) D-PhgAT PLP internal aldimine. Residues involved in cofactor stabilization are shown as stick representations, with carbon atoms shown in cyan and green for the two monomers in the active dimer, oxygens are shown in red, nitrogens in blue and phosphorous in orange. Coordinated water molecules are shown as red spheres. The aromatic ring of PLP is sandwiched between Val243 and Tyr149. The nitrogen of the pyridine ring is fixed by hydrogen bond coordination to Asp241 as is common in ATs. The phosphate group is stabilized by a network of interactions with Ser121, Gly122, and Thr123 from Monomer A; and Glu124, Phe304 and Thr313 from monomer B. A water molecule forms a bridge between one of the phosphate oxygens and the side chain carboxylic acid of Glu124; and another water molecule bridges a phosphate oxygen and the peptide nitrogen of Phe304. **(B)** Hydrogen bonds, direct or through water molecules, are shown as dotted lines and distances (Å) in grey.

The large active site cavity is crucial to accommodate bulkier substrates such as 4-HBF and explains the reported ability of D-PhgAT to accept bulkier halogenated substituents like bromine and fluorine (Figure 3.29 A).[154] The cavity extends from the surface of the protein and has a wide mouth with a constriction near the active site formed by Thr44, Phe74, His224, and Phe422 (Figure 3.29 B). This constriction accommodates hydrophobic side chains through π-stacking interactions, while orienting the keto substrate to accept the amine group from the PMP intermediate.

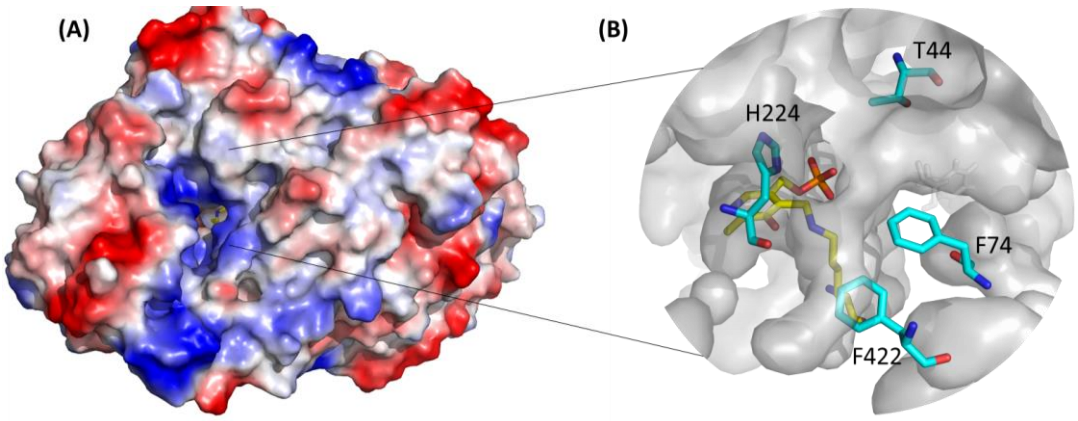


Figure 3.29 (A) D-PhgAT surface coloured by electrostatic potential. Around the active site cavity a predominance of positively charged residues can be observed. **(B)** The protein displays a wide cleft to accept a wide range of amino acid substrates. The cavity extends from the surface of the protein and has a wide mouth with a constriction near the active site formed by Thr44, Phe74, His224, and Phe422, highlighted as sticks. PLP is shown as yellow sticks.

Comparison of the *holo*- and *apo*-forms of the *P. stutzeri* D-PhgAT highlights a number of key differences around the PLP-binding pocket. The two structures align with an overall RMSD Cα of 1.29 Å, revealing three loop regions absent in the *apo*-structure and ordered in the *holo*-structure, due to PLP binding in the active site (Figure 3.30). A loop between residues 149 and 174 directly contacts the pyridine ring, with Tyr149 opposed edge-on against the ring, with the hydroxyl group pointing towards the phosphate group. This interaction orders a stretch of residues from Tyr149 – Trp152 which provides a template to order the remainder of this loop. The phosphate group interacts closely to form a hydrogen bond with Thr313 of the partner chain to order a loop between residues 291 and 306, which stabilises the loop between residues 28-35 of the primary chain (Figure 3.30 B).

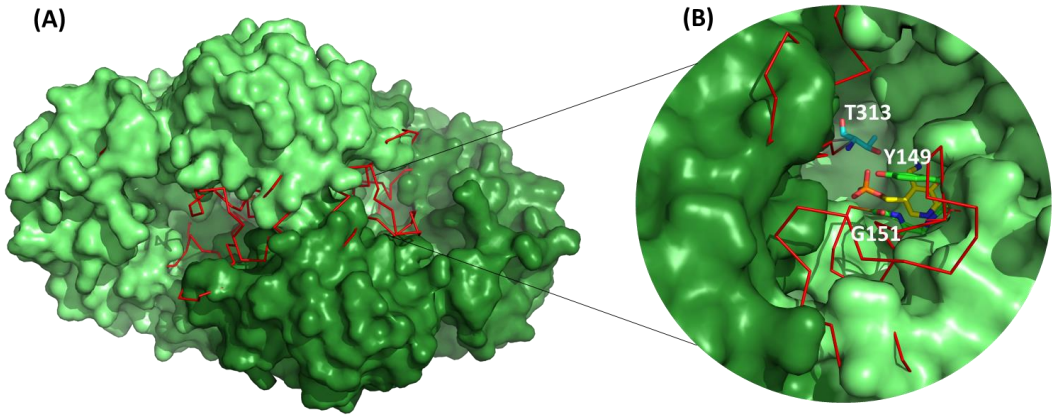


Figure 3.30 (A) Comparison of the *apo*- (PDB code: 2CY8) and *holo*-form (PDB code: 6G1F) of the *P. stutzeri* D-PhgAT. The *apo*-structure is shown as a surface while the *holo*-structure is shown as

ribbon. **(B)** Zoom into the active site reveals some of the key residues absents in the *apo* structure. Amongst them, the conserved Try149 involved in PLP stabilisation and Thr303 involved in the phosphate stabilisation.

The highlighted residues missing in the *apo*-structure are conserved amongst the AT family as they are all involved in cofactor stabilisation. Amongst them the tyrosine residue, sandwiching the cofactor which is typically present in the class III ATs.

In order to gain further insight into residues involved in substrate binding, co-crystallization experiments and substrate soaking have been performed with several substrates, to attempt to trap an external aldimine structure. As the enzyme has a fast turnover rate, a D-PhgAT K269A variant was also produced to stop the enzyme from turning over and trap the external aldimine. Natural substrates (L-Glu and D-Phg), non-natural 'smart' substrate (OXD) and substrate mimics, reported to inhibit ATs in the literature (such as L-Tyr hydrazide and L-Phe hydrazide and cycloserine) were also tried in co-crystallisation experiments. However, none of the approaches were successful and it was not possible to obtain the x-ray structure of an external aldimine of D-PhgAT.

3.7 Characterisation of enzyme variants

Concurrent with the determination of the D-PhgAT 3D structure, site-directed mutagenesis (SDM) experiments were undertaken in order to gain more insight into the residues involved in substrate binding and recognition. Looking at the residues around the active site, Gln300 was identified as a potential residue involved in the recognition of the carboxylic group of the BZF, crucial for the enzyme enantioselectivity (Figure 3.31A). This was identified by bioinformatic analysis as is not conserved amongst the other class III AT as highlighted in Figure 3.39. In order to confirm its role, SDM studies were undertaken, by mutating the residue to an asparagine (shortening the side chain length) and also to an alanine (completely disrupting the function).

As Arginine residues are crucial for the dual substrate recognition displayed by ATs, a detailed analysis of Arg residues around the active site and the surface pocket was carried out.[113] This analysis has led to the identification of Arg406 (Figure 3.31 B). This differs from the Arg348 suggested for the same role by Bornscheuer and co-workers which, in our structure, is very far from the active site pocket and surface.[108]

Mutagenesis primers were designed to prepare D-PhgAT Q300N, Q300A and R406A *via* the method of Naismith and Liu.[160] Using the original pET15b-dpgA plasmid as a template, the entire vector was PCR amplified with the relevant primer pairs incorporating the mutated residues into the gene sequence. Subsequent sequencing confirmed the mutations were present and that gene fidelity was otherwise maintained.

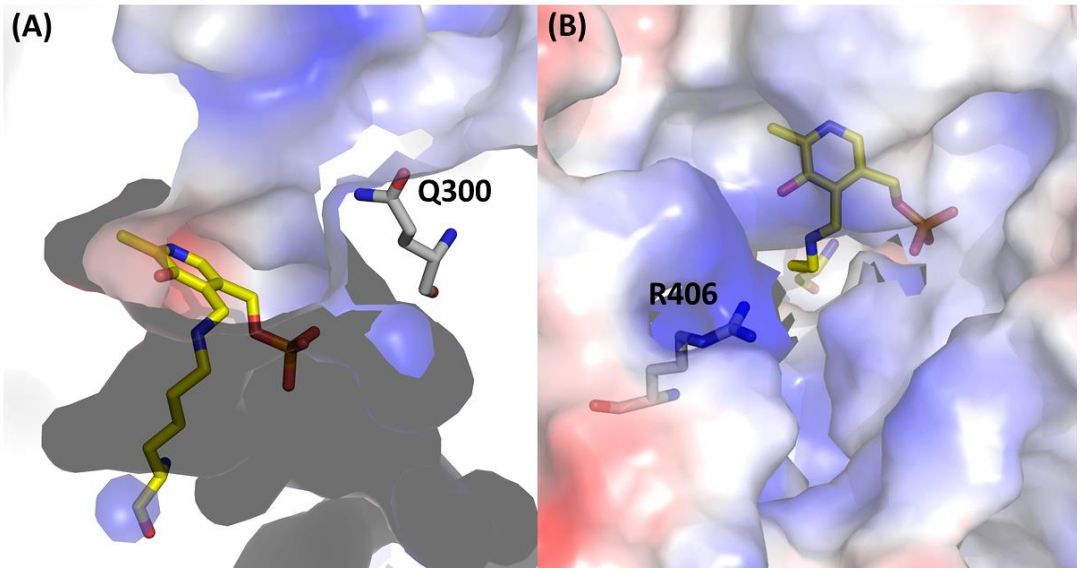


Figure 3.31 Residues identified in substrate recognition. **(A)** Amongst the residues around the active site the Gln300 has been identified as the potential residue involved in the interaction with BZF. **(B)** On the enzyme surface an arginine residue, Arg406 is suggested to be important for substrate recognition.

3.7.1 D-PhgAT Q300N and Q300A characterisation

D-PhgAT mutants Q300N and Q300A were expressed and purified according to the established procedures. Both mutants were soluble and expressed in yield similar to the wild type (WT) enzyme (~ 70 mg per litre of culture). The SEC showed a profile comparable to the WT enzyme with an elution volume of 78.0 mL (Figure 3.32 A and Figure 3.33 B) corresponding to a homodimer in solution. SDS-PAGE analysis showed the pure protein with a monomeric mass of ~ 51 kDa (Figure 3.32 A and Figure 3.33 B).

The presence of the designed mutation was further conformed by LC-MS as previously described and was is in keeping with the expected mass based on the predicted recombinant protein sequence minus the initial methionine(Figure 3.32 C and Figure 3.33 C).

As observed for the WT enzyme, multiple forms of the enzyme can be observed in the LC ESI-MS spectrum, including the loss of the methionine and gluconoylation of the histidine tag (+ 178 Da).

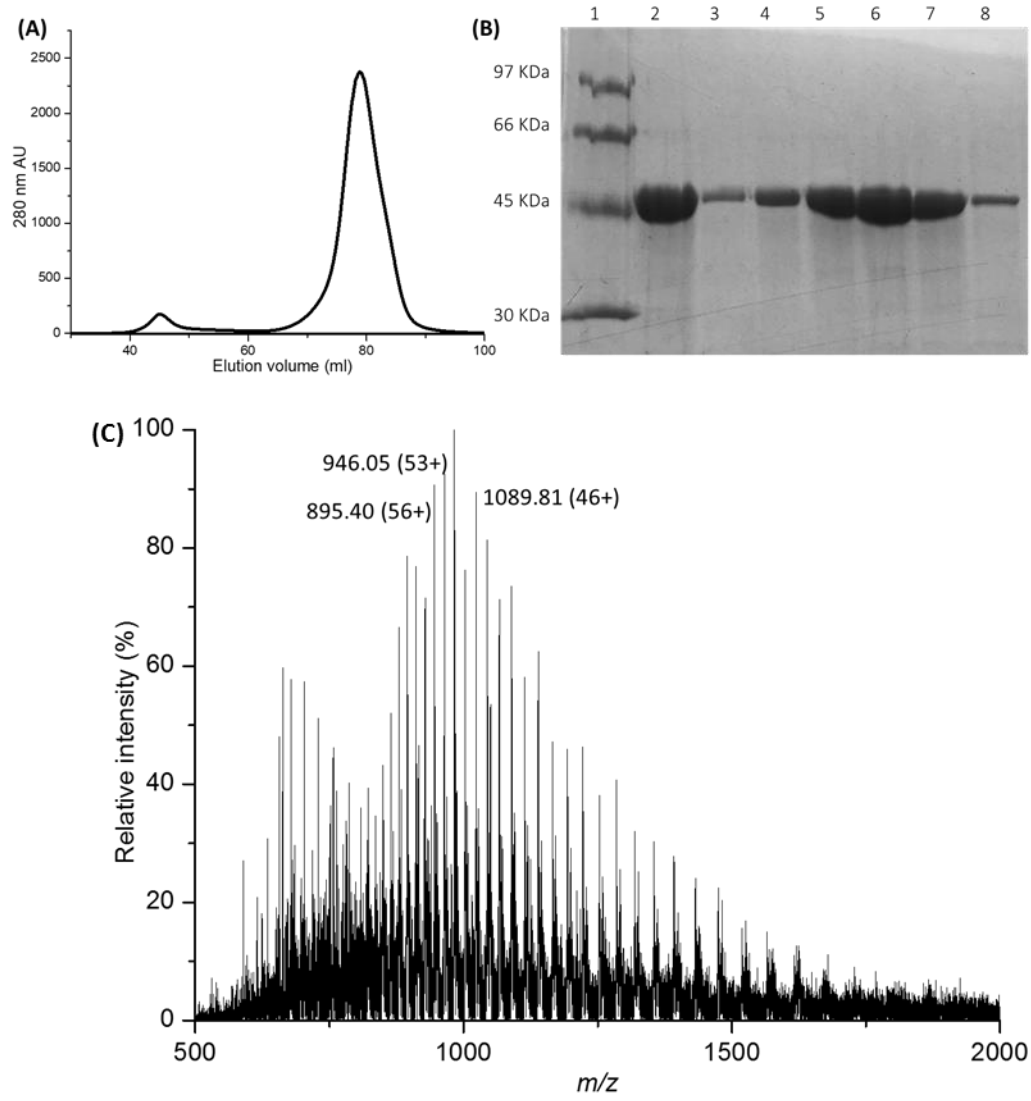


Figure 3.32 (A) Chromatogram from Superdex 16/60 S200 gel filtration chromatography of D-PhgAT Q300N **(B)** 12% SDS-PAGE gel of D-PhgAT purification steps. Lane 1: LMW marker (GE), Lane 2: D-PhgAT Q300N after nickel affinity purification, Lanes 3-8: elution fractions from SEC (3 mL fractions from 69-87 mL)**. (C)** Denaturing LC ESI-MS spectrum of recombinant D-PhgAT Q300N (20 µM) with a deconvoluted mass C of 50096.22 ±0.44 Da which agrees with the expected one for the recombinant protein of 50095.57 Da. The values are *m/z* with the charge states given in brackets.

D-PhgAT Q300N and Q300A were successfully characterised by cHPLC, according to the established procedures. The Michaelis-Menten plots obtained for both mutants are shown in Figure 3.34. Since the Q300 residue was hypothesized to be involved in BZF recognition, a

change in the enzyme enantioselectivity was expected by mutating this residue. However, the enzyme enantioselectivity was not affected by the mutation and both Q300N and Q300A mutants showed the production of only the D-enantiomer of Phg in >99% *ee*.

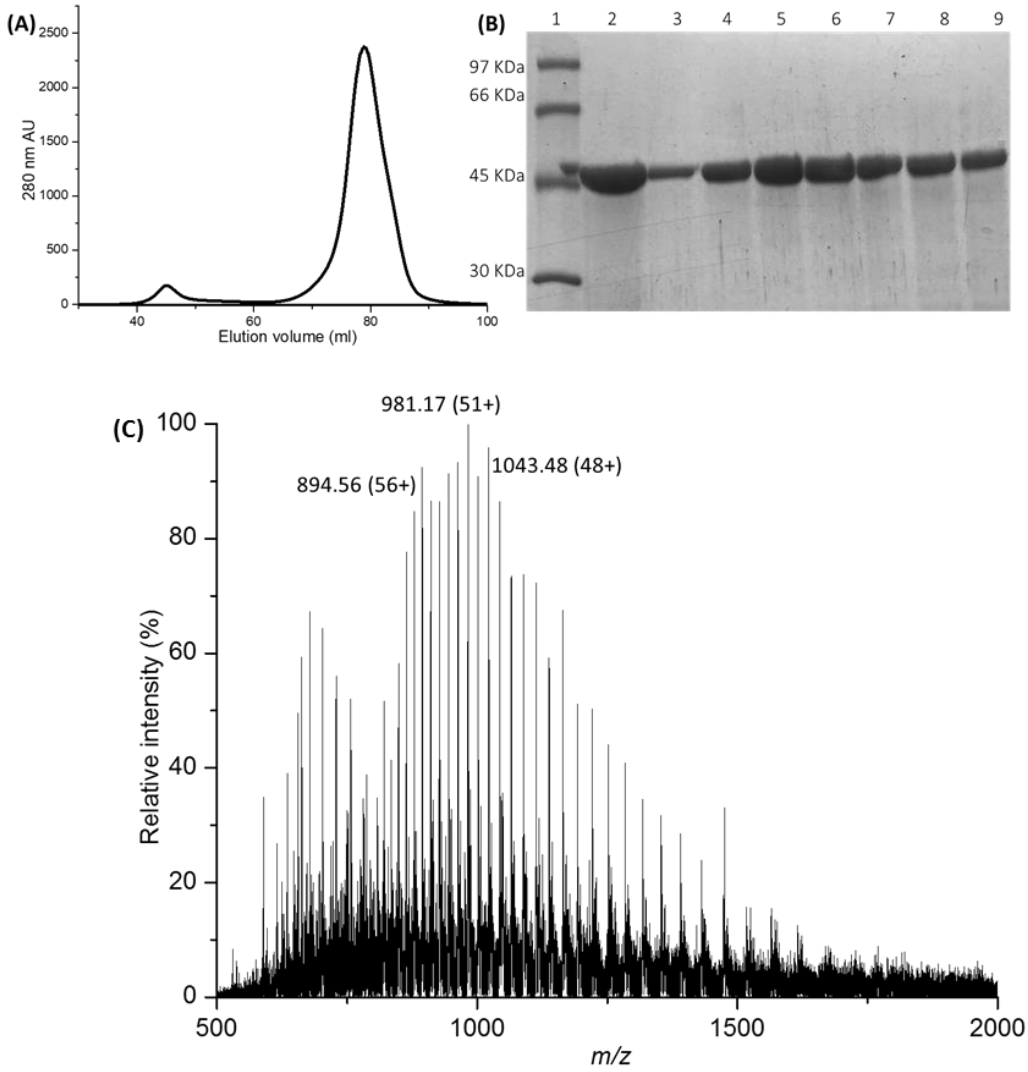


Figure 3.33 Chromatogram from Superdex 16/60 S200 gel filtration chromatography of D-PhgAT Q300A. D-PhgAT elution volume corresponds to a dimer. An aggregate is also visible at ~ 40 mL. **(B)** 12% SDS-PAGE gel of D-PhgAT purification steps. Lane 1: LMW marker (GE), Lane 2: D-PhgAT Q300A after nickel affinity purification, Lanes 3-8: elution fractions from SEC (3 mL fractions from 69-87 mL)**. (C)** Denaturing LC ESI-MS spectrum of recombinant D-PhgAT Q300A (20 µM) with a deconvoluted mass of 50053.15 ±0.31 Da which agrees with the expected one for the recombinant protein of 50052.55 Da. The values are *m/z* with the charge states given in brackets.

Furthermore, the K_M constant of the D-PhgAT Q300N for both L-Glu and BZF did not show any significant difference compared to the WT enzyme (for BZF 3.54 mM vs. 3.16 mM and

for L-Glu 43.93 mM vs. 26.17 mM for Q300N and WT respectively). In contrast to the Q300N mutant, a significant increase (by 3-fold) in the K_M towards the BZF was observed for the D-PhgAT Q300A. This result suggests that Gln300 may play a role in BZF binding as the K_M for L-Glu was not dramatically affected. A comparison between the WT and the two mutants is shown in Table 3.5.

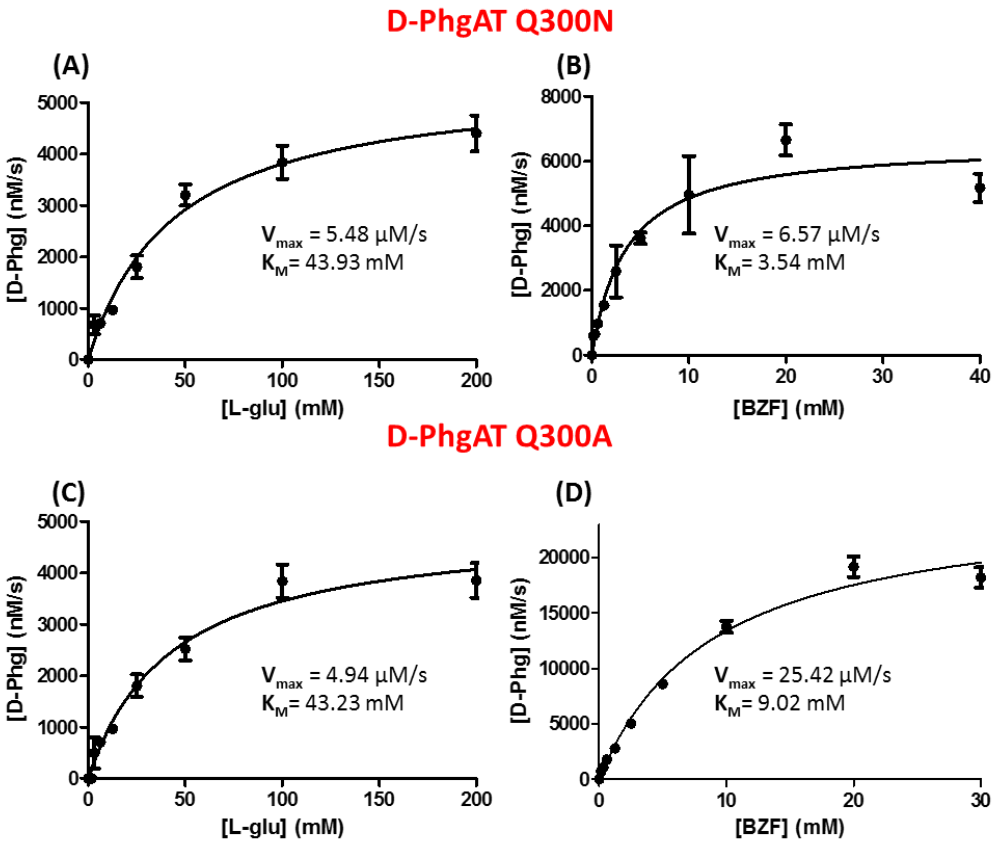


Figure 3.34 Michaelis–Menten plot of D-PhgAT Q300N (A and B) and Q300A (C and D) for L-Glu and BZF. Data was analysed by non-linear regression using Graphpad Prism software.

Enzyme Variant	K_M L-glu (mM)	K_M BZF (mM)
D-PhgAT wt	26.17 ± 3.63	3.16 ± 0.46
D-PhgAT Q300N	43.93 ± 7.98	3.54 ± 0.92
D-PhgAT Q300A	43.23± 7.90	9.02± 0.95

Table 3.5 K_M comparison between the D-PhgAT WT and Q300 mutants towards the two substrate L-Glu and BZF. No significant changes were observed for L-Glu while an effect toward BZF was observed with Q300A.

In conclusion, the Gln300 may play a role into the BZF binding, however is not crucial for the enzyme activity and/or enzyme enantioselectivity.

3.7.2 D-PhgAT R406 mutant characterisation

As for the other mutants, the D-PhgAT R406A was successfully expressed and purified according to the established procedures with yields similar to the WT enzyme. No significant changes were observed in the SEC profile and the pure protein was analysed by SDS-PAGE as shown in Figure 3.35. LC ESI-MS of D-PhgAT confirmed the presence of the mutation (observed mass: 50024.87 Da) in the translated sequence and further kinetic characterisation was therefore undertaken.

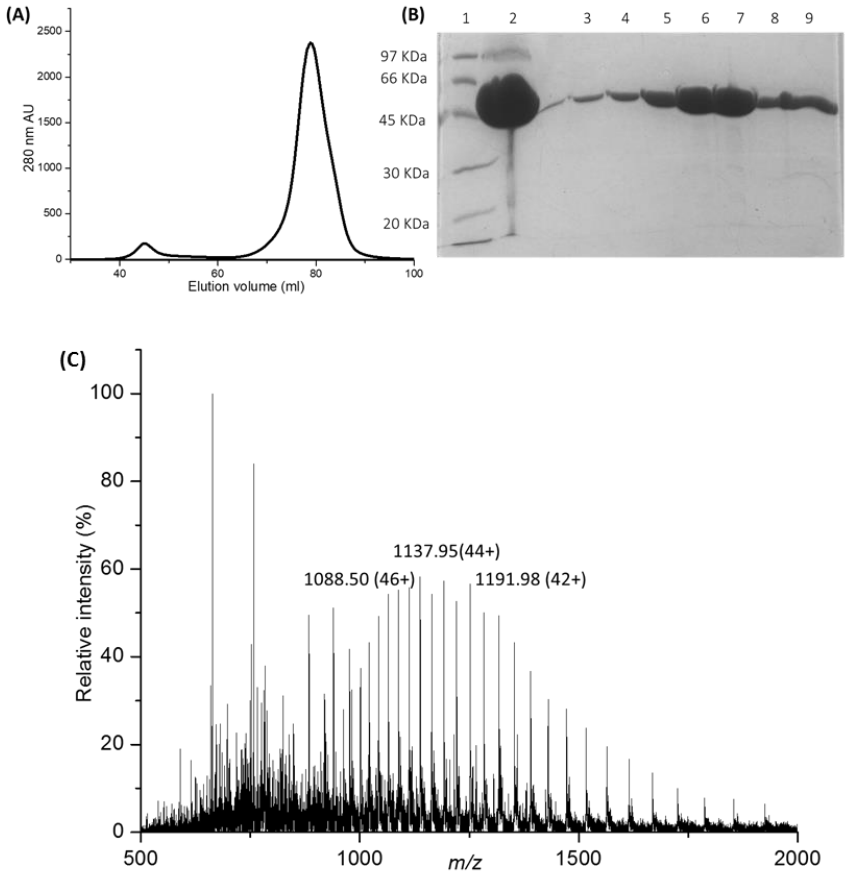


Figure 3.35 A Superdex 16/60 S200 chromatogram of the D-PhgAT R406A variant. D-PhgAT elution volume corresponds to a dimer. **(B)** 12% SDS-PAGE gel of D-PhgAT purification steps. Lane 1: LMW marker (GE), Lane 2: D-PhgAT Q300A after nickel affinity purification, Lanes 3-8: elution fractions from SEC (3 mL fractions from 69-87 mL)**. (C)** Denaturing ESI-MS spectrum of recombinant of D-PhgAT R406A (20 µM) with a deconvoluted mass of 50024.87 ±0.21 Da which agrees with the

expected one for the recombinant protein of 50024.49 Da. The values are *m/z* with the charge states given in brackets.

The D-PhgAT R406A mutant was kinetically characterised by cHPLC, according to the established procedures (Figure 3.36). The mutation did not affect the enzyme enantioselectivity as the production of only the D-enantiomer was observed. A full comparison between the kinetic parameters generated in this study is show in Table 3.6.

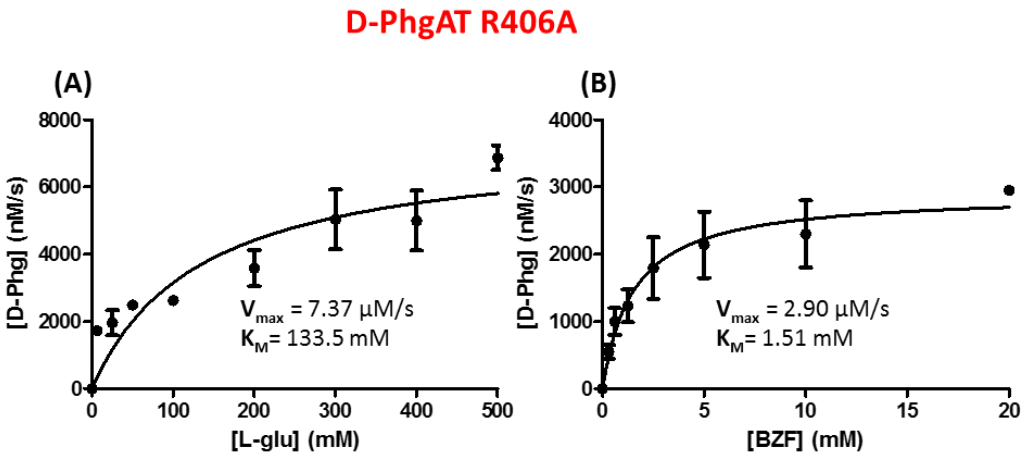


Figure 3.36 Michaelis–Menten plot of D-PhgAT R406A for L-Glu **(A)** and BZF **(B)**. Data was analysed by non-linear regression using Graphpad Prism software

Enzyme Variant	K_M L-glu (mM)	K_M BZF (mM)
D-PhgAT wt	26.17 ± 3.63	3.16 ± 0.46
D-PhgAT Q300N	43.93 ± 7.98	3.54 ± 0.92
D-PhgAT Q300A	43.23 ± 7.90	9.02± 0.95
D-PhgAT R406A	133.5 ± 30.5	1.51 ±0.45

Table 3.6 K_M comparison between the D-PhgAT WT and the mutants generated in this study towards the two substrate L-Glu and BZF.

The D-PhgAT R406 displays a K_M constant 6-fold higher than the WT enzyme (133.5mM vs. 26.17 mM). Affinity towards BZF is not dramatically affected with a K_M ~ 2-fold lower compared to the WT enzyme (1.51 vs. 3.16 mM). These results suggest that the hypothesized Arg406 plays a fundamental role in L-Glu recognition as the substrate affinity is dramatically affected when mutated.

3.8 D-PhgAT-D-Phg model

After several failed attempts to trap the D-PhgAT external aldimine and inconclusive SDM results, substrate docking was carried out to gain insight into the R-selectivity of the D-PhgAT. Dr Jon Marles-Wright (Newcastle University) modelled the PLP-D-Phg external aldimine in the active site based on the (*R*)-MBA bound structure of the (*S*)-selective, ω-AT from *Bacillus megaterium* (Bm-AT, PDB code: 5G09) (Figure 3.37).[161]

The obtained model suggests that the amine group of the PMP intermediate is delivered to the keto substrate from only one face, since His66, His213, and Thr303 (from the adjacent monomer) align the substrate so that the reaction can only take place in the correct orientation (Figure 3.37). This suggests that the unique stereo-inverting feature of the enzyme is due to the enzyme architecture, that allows BZF substrate to be aminated from just one face and not another. The carboxylic acid group of the acceptor substrate is within hydrogen bonding distance of His66 and Thr303, and a positively charged 'appendix' extends to Arg34, which could accommodate the carboxylic acid group of both the donor and the acceptor substrates. This is in contrast to the Bm-AT, which has a tryptophan in the position of Arg45 and thus does not possess this structural feature. A comparison of the different active site architecture of the two ATs is shown in Figure 3.38.

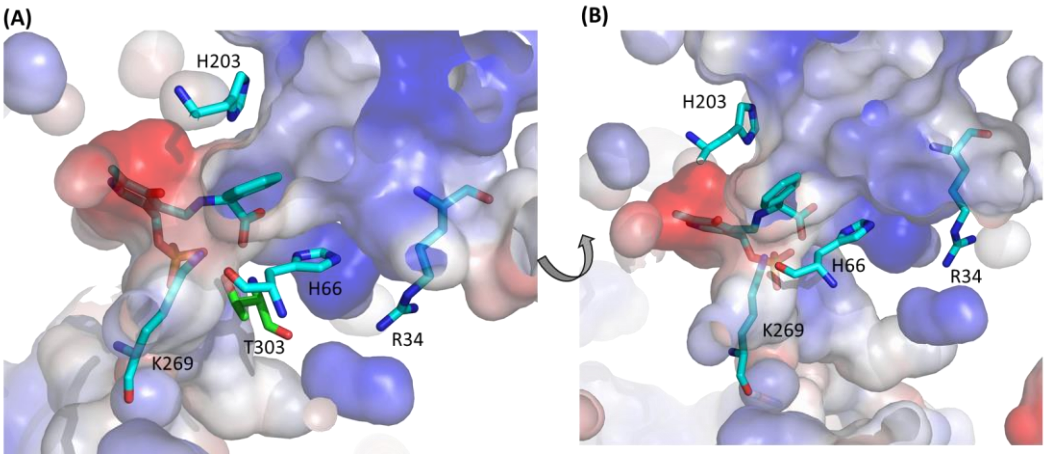


Figure 3.37 Model of D-Phg PLP external aldimine in active site tunnel of D-PhgAT. The D-Phg was modelled using the Bm-AT:(*R*)-MBA external aldimine (PDB code: 5G09) to place the substrate. D-Phg shown as sticks with cyan carbon atoms; His66 and His213, along with Thr303 from the partner chain are shown with cyan and green sticks respectively. Solvent accessible surface of D-Phg shown in grey.

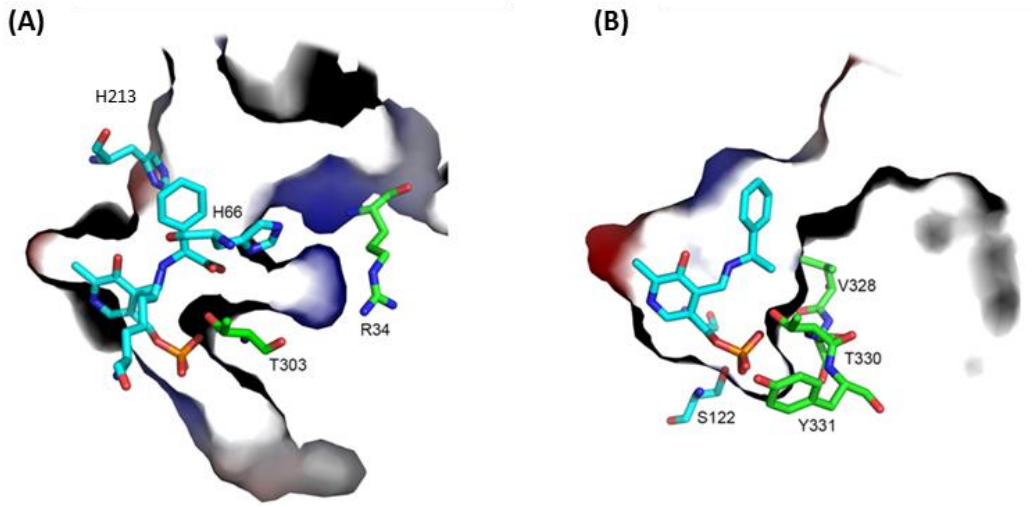


Figure 3.38 Comparison of the substrate-binding tunnel of D-PhgAT:D-Phg and BM (*S*)-AT: (*R*)-MBA. **(A)** Solvent accessible surface representation of the D-PhgAT substrate-binding tunnel shown colored by electrostatic charge. The modelled D-Phg external aldimine is shown as cyan sticks with residues contributing to the substrate 'appendix' highlighted, with the pocket lined with His66 and Thr303 and terminating at Arg34. This pocket could accommodate the carboxylic acid groups of both the donor and acceptor substrates. **(B)** Solvent accessible surface representation of the *B. megaterium* AT (PDB code: 5G09) substrate-binding tunnel colored as in (A). The (*R*)-MBA ligand occupies a pocket lined with residues with hydrophobic side-chains, with no charged appendix capable of accepting acidic groups**.**

Sequence analysis of some of the best characterised ATs belonging to the same fold (Cv-AT, Vf-AT and Bm-AT) revealed that the His66 and His213 residues hypothesized to be crucial for the correct substrate orientation are not present in other members of the class III AT family, confirming the uniqueness of the residues for the D-PhgAT (Figure 3.39).[162] The His66 in *P. stutzeri* D-PhgAT corresponds to a conserved Leu residue in Cv-AT, Vf-AT and Bm-AT while the His213 corresponds to a conserve Ala residue in Cv-AT, Vf-AT and Bm-AT. Furthermore, the *P. putida* D-PhgAT homologous, displaying 82% sequence identity with the *P. stutzeri* D-PhgAT, shows the presence of the same His residues supporting implication of these residues in the 'stereo-inversion' activity of the enzyme.[130] The Thr303 residue corresponds to a serine residue in Cv-AT and Vf-AT, suggesting the importance of a polar residue in that position.

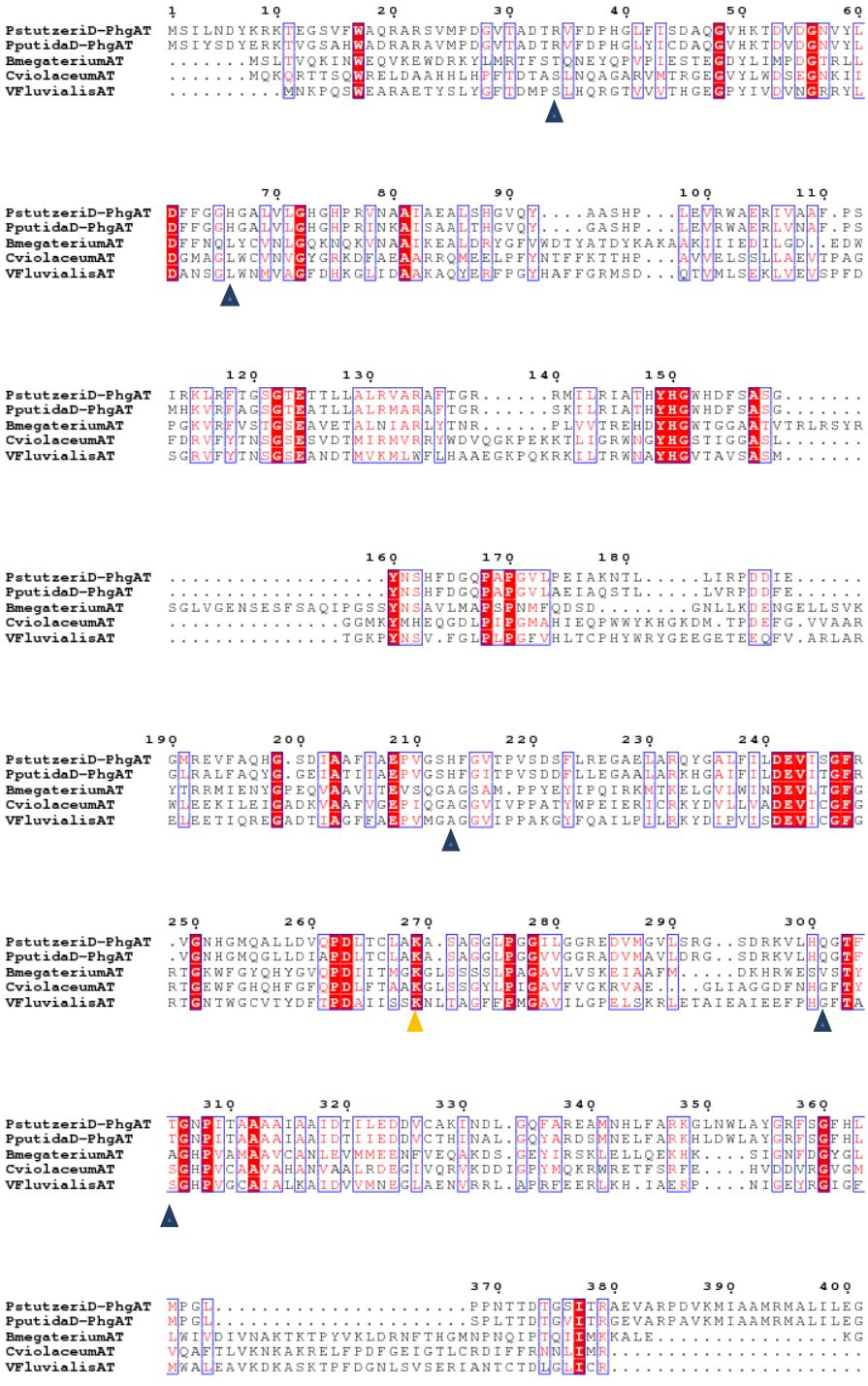


Figure 3.39 ESPript[162] alignment of *P. stutzeri* D-PhgAT (accession code: AAQ82900) and *P. putita* D-PhgAT (AX467211) with related, well characterised (*S*)-selective Class III ATs. These are *Bacillus megaterium* AT (A0A1C7D190), *Vibrio fluvialis* AT (AEA39183), *Chromobacterium violaceum* AT (Q7NWG4).The D-PhgAT specific residues Arg34, His66,His213, Gln300 and Thr303 proposed to be involved in substrate orientation and D-amino acid product formation are marked with a blue triangle. The conserved Lys269 residue is highlighted in yellow.

The Arg34 residue, suggested to be involved in the substrate recognition, is also unique to the stereo inverting D-PhgAT enzymes and shared with the *P. putida* D-PhgAT homologue. Based on the D-PhgAT:D-Phg model and the SDM results, a model of the dual substrate recognition achieved by D-PhgAT was built (Figure 3.40). According to the model the Arg406 residue, which was a target of SDM and Arg34, identified by substrate docking, recognize the carboxylic acid group of both amino donor and amino acceptor. A schematic representation of the substrate recognition achieved by the combination of these two residues is shown in Figure 3.40.

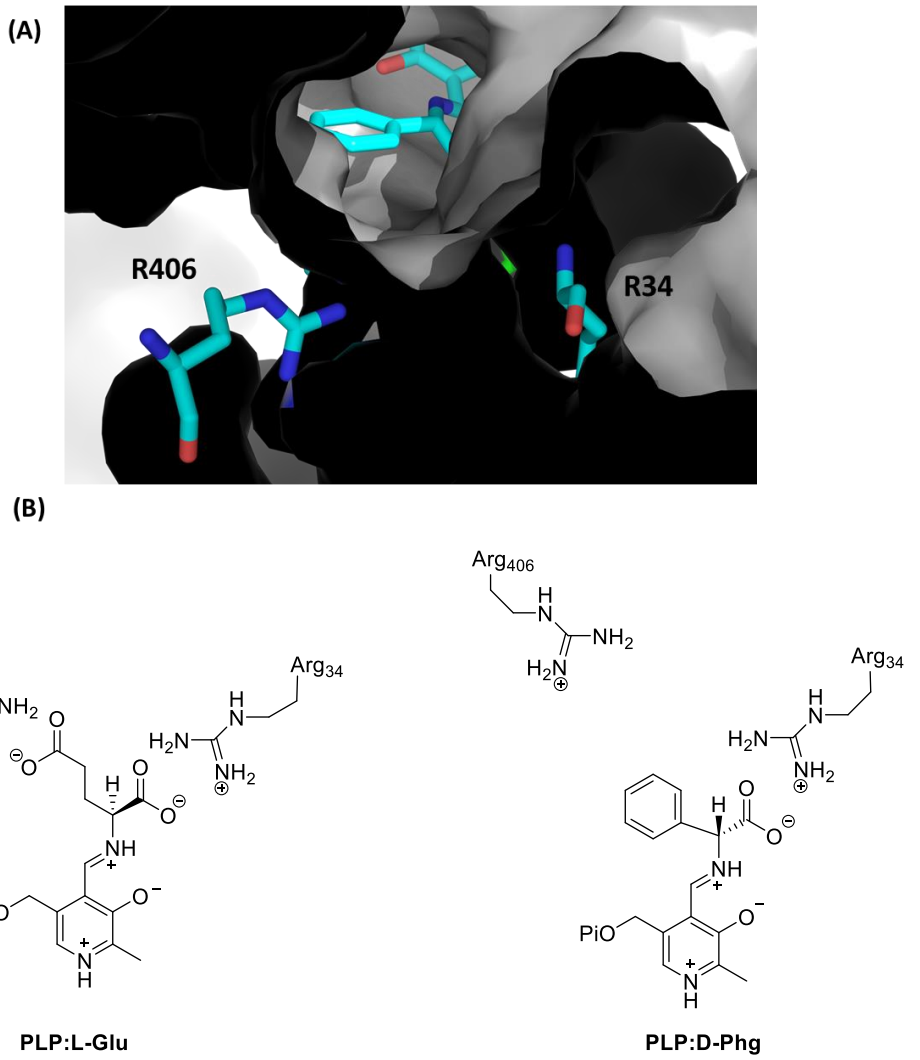


Figure 3.40 (A) Surface view of the D-PhgAT:D-Phg model. The Arg406 and Arg34 are located at the entrance of the active site **(B)** Schematic representation of the dual substrate recognition mechanisms of D-PhgAT. The 'flipping' Arg406 is involved in the recognition of the γ carboxylic group of L-Glu or moves out of the active site to accommodate the benzyl ring of D-Phg. Adapted from Eliot *et al.*[69]

According to this model, the 'flipping' Arg406 could either form a salt bridge to the glutamate γ carboxylate group or alternatively can move out of the active site to accommodate the benzyl ring of BZF/D-Phg. This phenomenon has been described in the literature as arginine 'switch' and was introduced in section 1.4.2.[111] The other arginine residue, Arg34 is thought to be involved in the recognition of the α-carboxylic group of L-Glu and BZF.[113]

These comprehensive analyses suggest that the D-PhgAT unique enantioselectivity yielding solely (*R*)-products has arisen from using the Type III TA fold and incorporating three key amino acids: His66, His213 and Arg34. A second arginine residue, Arg406, has been identified as responsible for the dual substrate recognition achieved by D-PhgAT. Future mutagenesis of these residues will provide further insight into their role in the mechanism.

3.9 D-PhgAT reaction mechanism

A detailed reaction mechanism for the D-PhgAT catalysed reaction is shown in **Figure 3.41**.

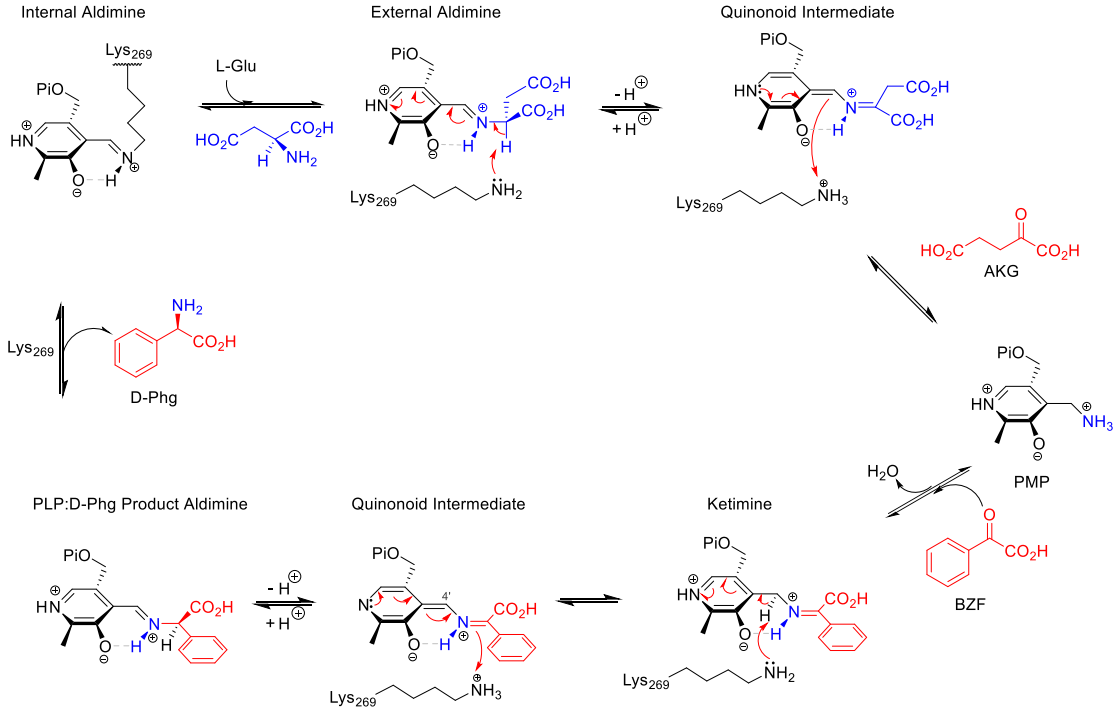


Figure 3.41 Reaction mechanism for the D-PhgAT reaction using L-Glu as amino donor and BZF as amino acceptor. The L-Glu amino donor is shown in blue while the BZF is shown in red. The key step determining the enzyme enantioselectivity is the C4'-hydrogen transfer from the lysine to the quinonoid intermediate.

The stereochemistry of the C-4' hydrogen transfer to the PLP cofactor is the key step of the AT mechanism, determining the stereospecificity of the AT catalysed reaction. As the stereocentre produced by the protonation from the catalytic lysine can occur on either the *si*- or the *re*-face, the position of the lysine residue in comparison to PLP determines the chirality of the final product. As discussed in 1.4.1, (*S*)-ATs are reported to act from below the quinonoid on the *si*-face, while (*R*)-ATs act on the *re*-face above the quinonoid intermediate (Figure 1.20).[106]

Because of the 'stereo-inverting' nature of the D-PhgAT the stereochemistry could not be predicted. However, reports in the literature have determined that the C-4' hydrogen transfer in D-PhgAT seems to occur on the *si*-face. This is consistent with our substrate model, which sees the Lys269 residue on the *si*-face of the cofactor and is shown in the proposed reaction mechanism (Figure 3.41).[163]

3.10 Synthesis of ^{15}N labelled D-Phg

Glycopeptide antibiotics such as vancomycin contain several types of D-Phg derivatives including the D-Hpg and the D-3, 5-dihydroxy-phenylglycine (D-Dpg) as shown in Figure 3.42.

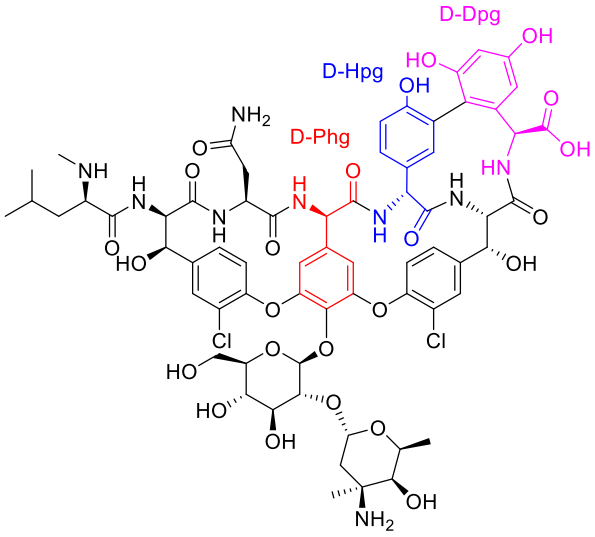


Figure 3.42 Vancomycin structure. The presence of different D-Phg (red) derivatives such as D-Hpg (in blue) and D-Dpg (in pink) can be observed.

Heavy atom labelling is a useful tool in natural product biosynthesis studies, as it allows the position of specific precursors to be identified in complex metabolites.[164,165] To this end, Dr Max Cryle (Monash University, Australia), interested in vancomycin biosynthesis, is actively searching for ^{15}N D-Phg for feeding studies.[119] D-PhgAT was therefore applied to the synthesis of ^{15}N labelled D-Phg and D-Hpg by feeding commercially available ^{15}N labelled L-Glu as amino donor (£80 for 100 mg). The strategy used is summarized in Figure 3.43.

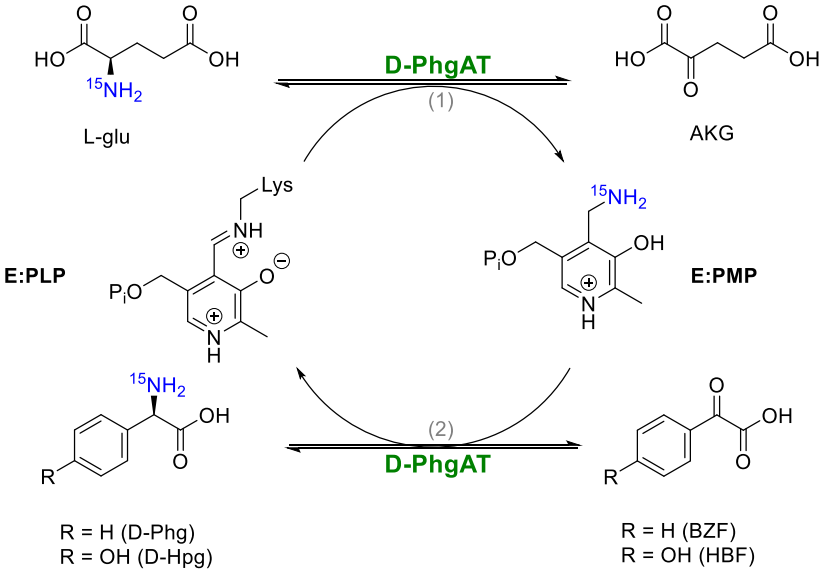


Figure 3.43 Experimental strategy for the preparation of ^{15}N labelled D-Phg and D-Hpg using ^{15}N labelled L-Glu as the amino donor and D-PhgAT as the catalyst.

This work was carried out with Bsc project student Fengxin Li. An analytical method for analysis of D-Phg/D-Hpg production was developed using reversed phase HPLC and a C18 column (mobile phase of 5% to 95% acetonitrile (ACN) with 0.1% Trifluoroacetic acid (TFA) (v/v)/ water 0.1% TFA (5.12.1)). This was done as no preparative chiral columns were available in the department. Single enantiomers of each amino acid (0-400 mM) were used to build a calibration curve to relate the AUC with the amount of product formed over time (Appendix 8.8).

As the AT unfavourable reaction equilibrium is usually overcome by using a large excess of the inexpensive unlabelled L-Glu, a compromise between conversion and usage of the expensive ^{15}N labelled L-Glu was needed to obtain ^{15}N D-Phg and ^{15}N D-Hpg. Thus, parallel reactions with different concentrations of unlabelled L-Glu were carried out to find the best ratio of conversion to amino donor concentration at constant BZF concentration (~

66 mM, 50 mg of BZF in 5 mL reaction mixture). Samples were withdrawn at different time points up to 8 h and analysed by HPLC after dilution by 40-fold in 5% ACN 0.1% TFA. A summary of the results is shown in Figure 3.44. Using the lowest L-Glu concentration (136 mM) a 30% BZF conversion was observed after 8 h. A 57% BZF conversion was observed when using 340 mM L-Glu after 8 h of reaction. Finally, a 62% conversion was observed when using the highest L-Glu concentration (680 mM) after 1 h of reaction. As the final yield in this model reaction at 340 mM L-Glu (57%) at 8 h was very similar to one the obtained with 680 mM L-Glu (62%) after 1 h, large scale synthesis of ^{15}N D-Phg was carried out using 340 mM of ^{15}N L-Glu for 8 h. Using the optimised conditions, a final 59% BZF conversion was observed. A control with unlabelled L-Glu was carried out for each biotransformation.

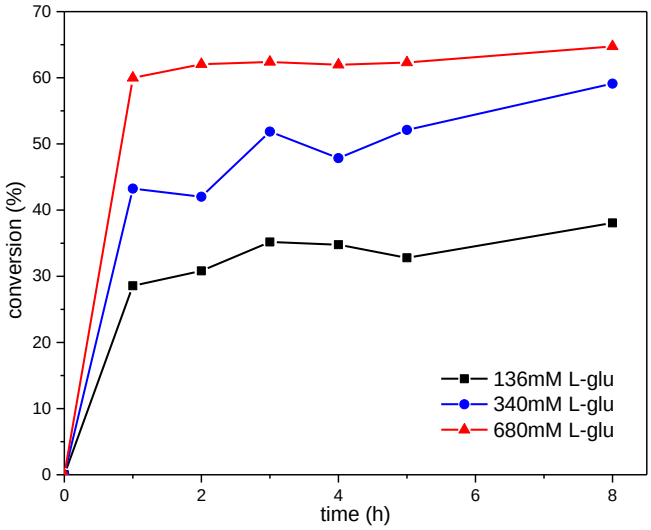


Figure 3.44 BZF conversion over time at different L-Glu concentrations. A compromise between L-Glu concentration and percentage conversion was needed because of the high cost of the ^{15}N L-Glu. A conversion of 30%, 57% and 62% was observed at 136 mM, 340 mM and 640 mM L-Glu concentration, respectively. BZF concentration was 66 mM.

Product purification was carried out using a preparative reverse phase HPLC on a Waters 600 with a Phenomenex Luna C18 column (5 µm, 250 × 21.2 mm) at a flow rate of 18.0 mL min^{-1}. The HPLC profile of the ^{15}N D-Phg purification is shown Figure 3.45.

The desired product (8.9 min) was collected and freeze-dried and a white powder was obtained. The purity of the obtained ^{15}N labelled D-Phg was confirmed *via* NMR spectroscopy, matching reported spectral data (section 5.12.4).[166] Fourier Transform mass

spectrometry (FT-MS) analysis, performed by Dr Joanna Simpson (University of Edinburgh), confirmed the total incorporation of ^{15}N into the molecule (Figure 3.46, Table 3.7). An *m/z* value of 152.0712 was observed for the unlabelled compound which is consistent with the expected value of 152.0706. The ^{15}N D-Phg displayed an *m/z* value of 153.0675 value matching the expected value of 153.0676. A summary of all the expected and observed *m/z* values is shown in Table 3.7. The 1 Da shift mass of the labelled D-Phg compared to the unlabelled confirmed the successful labelling.

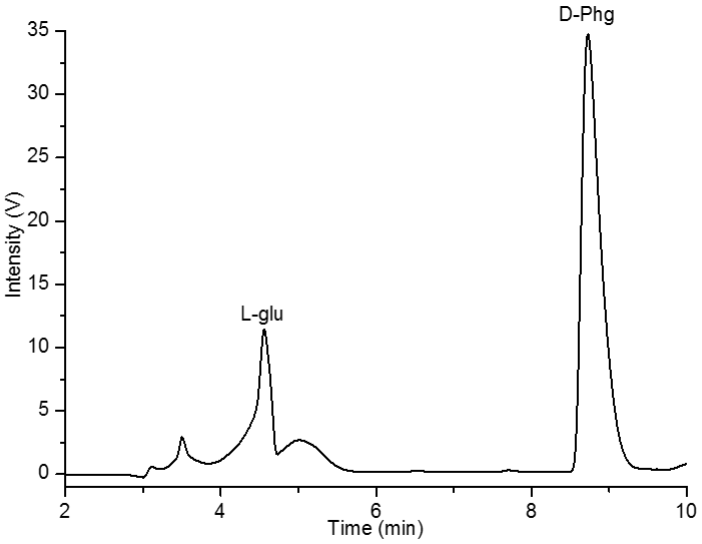


Figure 3.45 Preparative HPLC chromatogram of ^{15}N D-Phg purification at 205 nm. The D-Phg elutes at 8.9 min. The peak was collected, freeze-dried and analysed by NMR and mass spectrometry.

Compound	Expected *m/z*	Observed *m/z*
D-Phg	152.0706	152.0712
^{15}N D-Phg	153.0676	153.0675
D-Hpg	166.0504	166.0505
^{15}N D-Hpg	167.0477	167.0476

Table 3.7 Observed and expected *m/z* values for the ^{15}N labelled and unlabelled D-Phg/D-Hpg.

Due to the higher cost of the HBF substrate, the optimization of the reaction conditions with unlabelled L-Glu was not carried out. Thus, the same reaction conditions established for BZF were applied for the synthesis of ^{15}N D-Hpg. A final HBF conversion of 20% was observed, ~ 3-fold lower than the conversion observed with BZF. As D-PhgAT displays high

catalytic efficiency (k_{cat}/K_M) towards HBF (899.15 ± 0.27 M^{-1} s^{-1}), the reduced yield could be due to inhibition effects, typically observed in ATs.

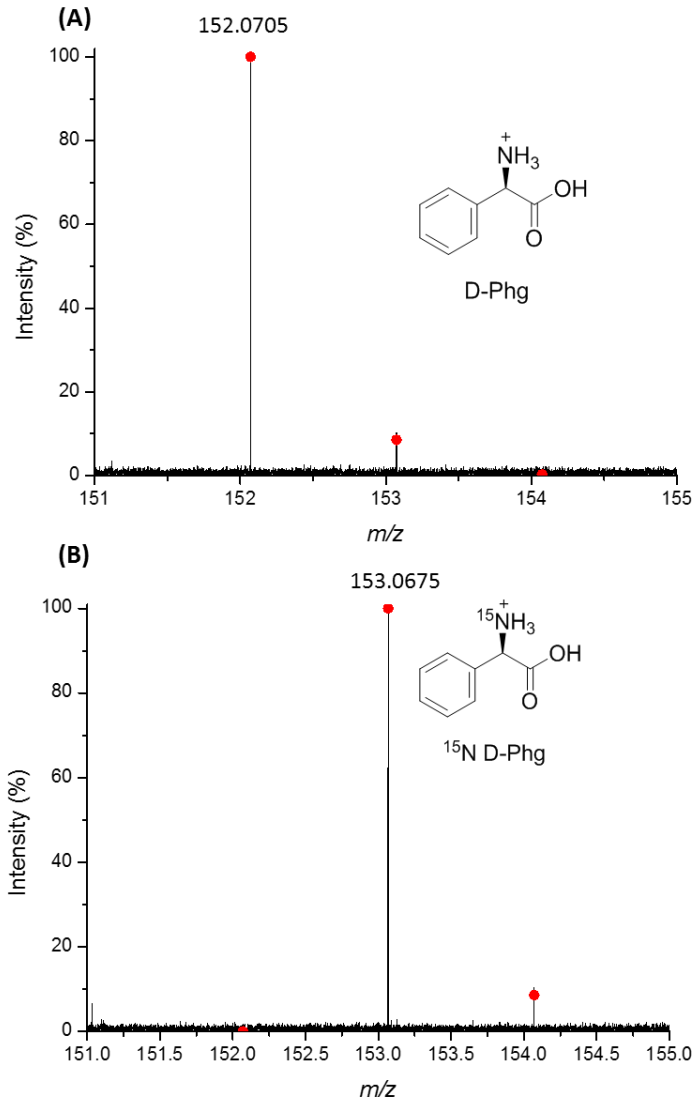


Figure 3.46 FT-MS of **(A)** the unlabelled and **(B)** the purified ^{15}N D-Phg. The red dots are the predicted isotopic distribution for each compound. The change in the isotopic distribution of the ^{15}N D-Phg substrate confirms the successful synthesis of the ^{15}N derivative of D-Phg. Analysis was carried out in positive mode.

The increased hydrophilicity of D-Hpg compared to D-Phg, caused problems in product purification. Using the C18 optimized method, the L-Glu substrate and the D-Phg product

co-eluted. By lowering the ACN percentage in the mobile phase down to 2% a small separation could be observed (4.4 min vs. 4.8 min) (Figure 3.47).

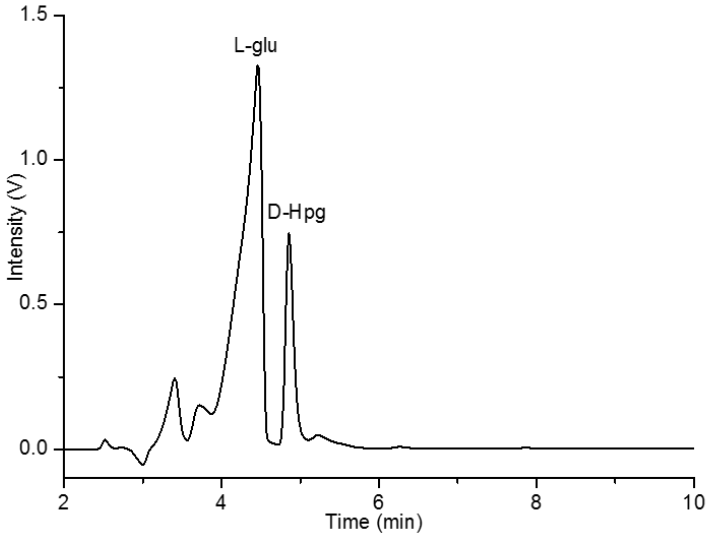


Figure 3.47 Preparative HPLC chromatogram of ^{15}N D-Hpg purification at 205 nm. L-Glu shows a retention time of 4.4 min while D-Hpg of 4.8 min. The peak corresponding to D-Hpg was collected, freeze-dried and analysed by NMR and mass spectrometry.

Fractions corresponding to the D-Hpg peak were collected and freeze dried. In this case a yellowish oil was obtained that was analysed *via* NMR and MS analysis. FT-MS analysis, carried out in negative mode, showed a perfect incorporation of the ^{15}N into the D-Hpg molecule (Figure 3.48, Table 3.7). An *m/z* value of 166.0505 was observed for the unlabelled compound which is consistent with the expected value of 166.0504. The ^{15}N D-Hpg displayed an *m/z* value of 167.0476 value matching the expected value of 167.0477. A summary of all the expected and observed *m/z* values is shown in Table 3.7. However, NMR analysis indicated the presence of contaminants in the sample, corresponding to the substrate L-Glu (section 5.12.4).[166,167] In order to obtain a pure product approaches such as flash chromatography may neee to be carried out to separate the two amino acids. Alternative purification procedures, such as using a different HPLC columns (such as the chirobiotic T column), may be necessary. The developed method for the synthesis of ^{15}N D-Phg derivatives represents a useful tool for drug feeding studies.

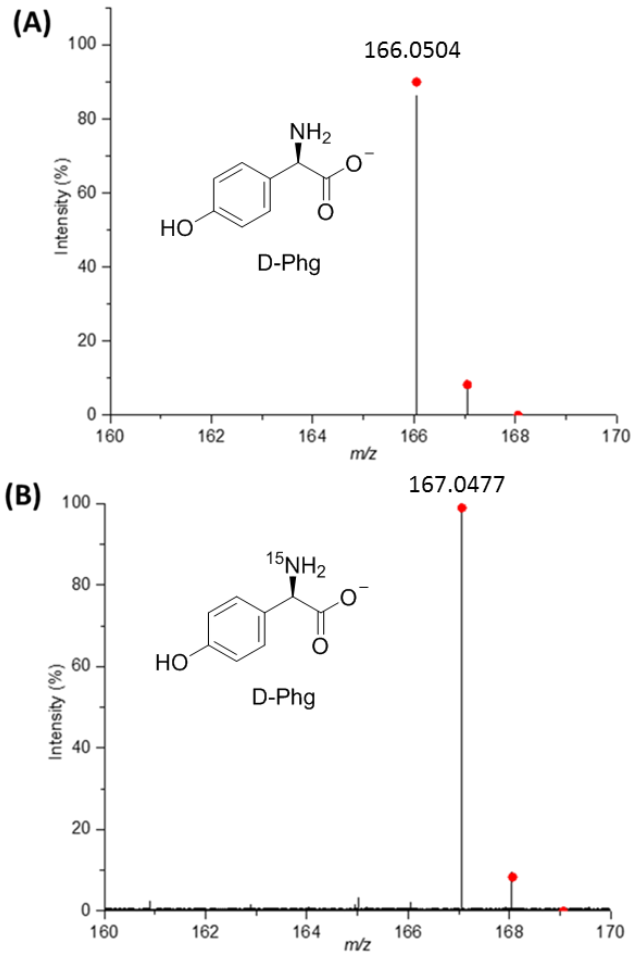


Figure 3.48 FT-MS of **(A)** the unlabelled and **(B)** the purified ^{15}N D-Hpg. Red dots are the predicted isotopic distribution for each compound.

3.11 Enzyme immobilization

The use of biocatalysis is increasing industrially, but is still limited to a restricted number of reaction types, such as the production of enantiopure alcohols, amine and esters.[7,168,169] Despite well-established enzymatic procedures, there remain many potential cases in which enzymes are not used. An estimate by Codexis indicates that ~ 30% of the small molecule compounds handled by pharmaceutical companies could benefit from the inclusion of a biocatalyst in the production of either the API or the final drug, but the actual usage of enzymes is much lower.[170] One of the principle barriers for the widespread application of enzymes in a wider range of chemical transformation and to a larger number of products is the high development and applications costs of enzymes compared to traditional chemical

methods. A strategy to overcome the high costs associated with enzymes consists of immobilizing the enzyme on a solid support, increasing the half-life of the enzyme and decreasing the reaction costs as the enzyme can be reused multiple times without additional isolation and purification. Furthermore, immobilization allows for easy recovery of the product, continuous operation of the enzymatic processes and quick termination of reactions.[171]

Immobilization techniques can be grouped into three main categories: 1) encapsulation/ entrapment of the enzyme in a permeable material, 2) cross-linking of the intact enzyme, and 3) binding to a suitable (inert) carrier.[170] Approaches such as encapsulation and cross-linking alter the direct environment of the enzyme, affecting the natural structure of the enzyme. This is why the most common used method is the use of a suitable carrier, typically a porous plastic bead or a silica based material that binds the enzyme. However, despite the effort in trying to develop established immobilization procedures that could be generally applicable, there is still no universal immobilization method and screening is still required for each new process. To this end, EnginZyme has developed an immobilization method, the EziG carrier, which allows the immobilization of potentially any enzyme with the commonly used His6-tag without requirement of having to drastically change immobilization conditions (http://enginzyme.com/).

Selective binding of target protein with an affinity tag is a standard procedure in recombinant protein purifications such as IMAC. The same principle is applied for the EziG carrier, which binds the enzyme *via* the His6-tag using iron in ionic form (+3). Iron is used due to its low toxicity and stable binding as no leaching occurs from the resin. This immobilization method maintains the enzyme in its natural form without altering the direct environment and without altering the enzymes structural mobility. Because of the specificity of the affinity tag, immobilization can be carried out without requirement of prior purification. EziG is a carrier based on a specific controlled pore glass (CPG) material that has interconnecting pores, allowing efficient mass transfer (Figure 3.49). Three different EziG carriers are commercially available, all of them based on CPG but with different properties with regards to surface, pore diameter and bulk density: EziG-1 Opal, EziG-2 Coral and EziG-3 Amber.

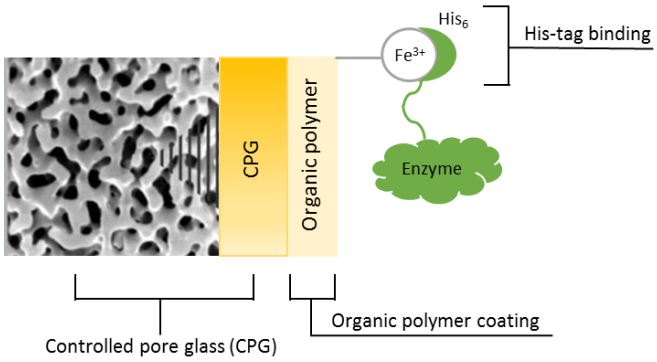


Figure 3.49 Schematic picture of the EziG enzyme carrier. The external and internal porous surface of CPG is covered with an organic polymer layer derivatized to chelate iron (Fe (III)) so that can bind the enzyme His-tag. Adapted from Cassimjee et *al.*[170]

In collaboration with Maria Abagnale, a visiting PhD student from the Universitá degli studi di Napoli Federico II (Naples, Italy), EziG was tested as a suitable support to immobilize the D-PhgAT. All resins were tested in order to identify the most suitable carrier for D-PhgAT immobilization. Aiming for 10% loading (9 mg of EziG *per* 1 mg of purified D-PhgAT), as suggested by the manufacturer, the three EziG carriers were tested at different imidazole concentrations, a crucial factor for successful immobilization, and for different times (30 and 60 min). Immobilization efficiency was evaluated by Bradford assay,[172] a standard protein quantification method, carried out on the supernatant after immobilization. A summary of the results is shown in Table 3.8. This experiment allowed the identification of the best immobilization conditions and the best carrier which was found to be EziG-2 with 75 mM imidazole. All subsequent studies were conducted under the aforementioned conditions.

[0] Imidazole (mM)	30 min	60 min	[25] Imidazole (mM)	30 min	60 min	[75] Imidazole (mM)	30 min	60 min
EziG^1	1.12	0.84	EziG-1	1.13	1.38	EziG^1	0.98	1.22
EziG^2	0.54	0.09	EziG-2	0.73	0.29	EziG^2	0.05	0.04
EziG^3	1.32	0.43	EziG-3	0.99	0.38	EziG^3	0.26	0.09

Table 3.8 Bradford results on supernatant after immobilization. If the immobilization is successful no remaining protein would be present on the supernatant. The numbers refer to the D-PhgAT concentration in mg.

Once the enzyme was successfully immobilized, BZF (10 mM) and L-Glu (300 mM) were added to the reaction mixture and gently shaken on an orbital shaker for 1 h at 40 rpm. The

supernatant was quenched by dilution into the cHPLC mobile phase and analysed by cHPLC. Using this method, the enzyme activity was evaluated at different temperatures (15 °C, 37 °C, 50 °C and 60 °C) and pH (5, 7.5, 9.5 and 11) in order to identify any difference in the activity of the free and immobilised forms of the enzyme. Furthermore, as EziG allows the direct immobilization of the cell free extract, activity was also evaluated in the free and immobilized cell free extract (Figure 3.50). However, as shown in Figure 3.50, there are no differences between the optimum pH and temperature of the free and the immobilized enzyme, with the highest activity observed at pH 9.5 and at 37 °C.

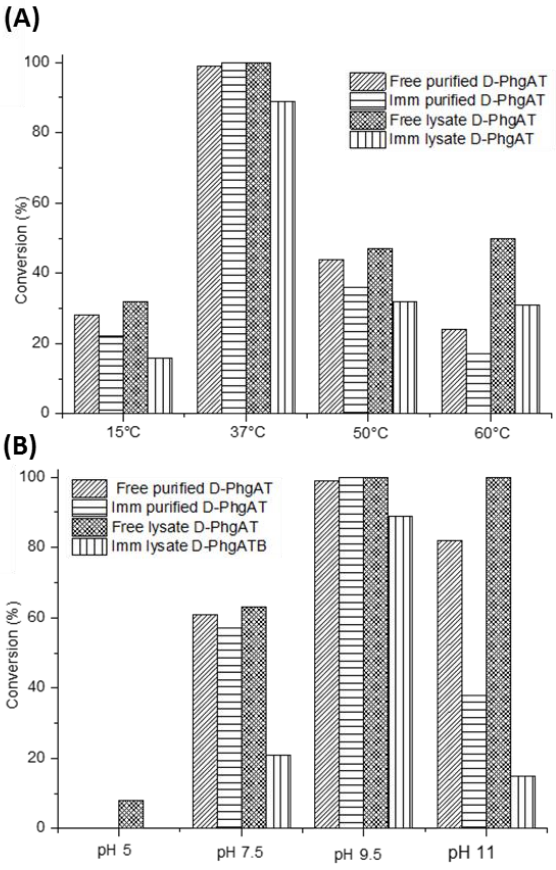


Figure 3.50 D-PhgAT activity of free and immobilized, pure and cell free extract D-PhgAT activity at **(A)** different temperatures (15 °C, 37 °C, 50 °C and 60 °C) and **(B)** different pHs (5, 7.5, 9.5 and 11).

Recycling tests were performed at pH 9.5 and 37 °C to investigate the possibility of re-using the enzyme multiple times. Reactions with purified and cell free extract immobilized D-PhgAT were carried out for 1 h (at 10 mM BZF and 300 mM L-Glu). The resin was then washed in buffer (0.1 M CAPS, 150 mM NaCl, 75 mM imidazole, 50 µM PLP) and stored at 4

°C. After 24 h the reaction was repeated under the same conditions and this same procedure was repeated 9 times. At each recycling point an aliquot was taken and analysed by HPLC (Figure 3.51).

Unlike the free D-PhgAT which can be used only once as it cannot be recovered from the reaction mixture, the immobilised enzyme is still active after 9 cycles. Furthermore the enzyme was immobilised directly from the cell free extract and showed conversion yields comparable to the purified enzyme suggesting no requirement for expensive purification procedures for successful catalysis prior to application in an industrial plant.

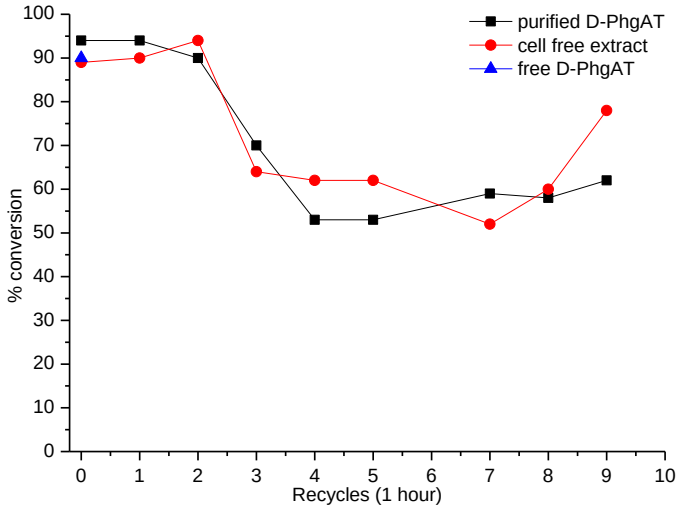


Figure 3.51 Recycling analysis of the free D-PhgAT, immobilized from cell-free extract and immobilized purified D-PhgAT. A decrease in substrate conversion is observed after 2 recycling but the conversion is stable at ~ 60% up to 9 recycles.

This preliminary immobilization study show that D-PhgAT can be successfully immobilized using the commercial carrier EziG. These results are promising and future studies on a larger scale with continuous reactors will be necessary to evaluate the reaction on an industrial scale.

3.12 Conclusions and future work

The initial aims of the project have been met. The *P. stutzeri* D-PhgAT has been fully characterised and successfully applied to the biosynthesis of high value compounds starting

from the natural D-Phg substrate to a wider range of products such as aromatic D-amino acids.

The widespread use of PLP-dependent ATs for industrial biotechnology has been hindered by the unfavourable reaction equilibrium towards product formation. Approaches to overcome these include the use of 'smart' amine donors, coupled recycling systems or removal of the keto co-product.[101–103] A wide range of amino donors was tested in order to identify the best amino donor, identified as L-Glu. Using a large excess of inexpensive L-glu (~ 20-fold K_M) the equilibrium was driven towards product formation. Substrate concentration, time and enzyme loading have been optimised, obtaining a final conversion of 93%, % higher than those previously reported. Moreover, D-PhgAT has shown the capability to yield a range of D-amino acids such as D-Phe, D-Tyr and D-Trp.

A recent report described the engineering of three recombinant *E. coli* strains that co-expressed four, seven and nine enzymes, including the *P. stutzeri* D-PhgAT as the final step to obtain D-Phg derivatives from racemic mandelic acid, styrene and L-Phe respectively.[173] The D-PhgAT substrate promiscuity shown in this study suggests that similar cascades could be constructed to allow conversion of simple building blocks to produce a variety of enantiopure aromatic D-amino acids.

The determination of the crystal structure of the D-PhgAT with its bound PLP cofactor has begun to shed light on the unique stereo-inverting and enantioselective properties of the enzyme. The broad substrate scope is explained by the large active site pocket and a combination of modelling and sequence analysis highlights three key active site residues that potentially control the exquisite (*R*)-selectivity of the enzyme, the His66, His213 and Arg34.

These residues have been shown to be unique to the stereo-inverting enzymes and are shared between the two D-PhgAT homologues from *P. stutzeri* and *P. putida.* Two Arg residues, Arg406 and Arg34, have been identified as responsible for the recognition of the carboxylic group of BZF and L-Glu. Future mutagenesis of these residues will provide further insight into their role in the mechanism.

The crystal structure of the PLP form of D-PhgAT paves the way for enzyme engineering in order to further expand the substrate scope of the enzyme. Enzyme

evolution will be necessary in order to improve conversion towards desirable non-natural substrates such as HPPA, for which a mere 15% conversion was observed.

Enzyme immobilization was carried out using the commercially available EziG resin in order to decrease the biotransformation costs, allowing reuse of the enzyme multiple times without the requirement to isolate and purify the enzyme. Immobilization conditions have been optimised and optimum pH and temperature of the immobilised enzyme has been determined. The enzyme has been recycled up to 9 times, displaying a stable conversion of 60%. These results are promising and future work will be carried out in order to perform in flow reactors, allowing a quick separation of the product and overcoming problems such as substrate inhibition.

Moreover, D-PhgAT has also been shown to be a useful tool for the production of ^{15}N D-Phg derivatives that will be used for feeding studies.

This study strengthens D-PhgAT's place in the biocatalytic tool box and also lays down the foundations for future applications in the production of enantiopure natural and unnatural D-amino acids.

4 The aminotransferase FumI

The amino-pentol AT FumI, isolated from phytopathogenic filamentous fungus *Sphingopyxis sp.* MTA144, is involved in the second step of the mycotoxin fumonisin B_1 (FB_1) catabolic pathway. [135,140,141,174] FB_1 is carcinogenic and is a common contaminant of corn and corn-based products known to cause severe effects after ingestion by humans and animals.[134,175] Studies of FB_1 degradation in bacteria has led to the discovery of a class III AT responsible for the deamination of the hydrolysed FB_1 (HFB_1, 2-amino-12,16-dimethylicosane-3,5,10,14,15-pentol) by transferring its 2-amino group to pyruvate (pyr), yielding 2-keto HFB_1 (3,5,10,14,15-pentahydroxy-12,16-dimethylicosane-2-one) and L-Ala (Figure 4.1).[142] The research organization BIOMIN (Austria) has investigated FumI as a possible food detoxification reagent and this application is covered by a patent (WO2004085624). However the application of FumI for biocatalysis is unexplored and its substrate scope in terms of both amino donor and acceptors is unknown.[135] The ability of FumI to accept long chain alkyl substrate (C_{20}) is a unique feature amongst the AT family and represents an appealing property for industrial biotechnology.

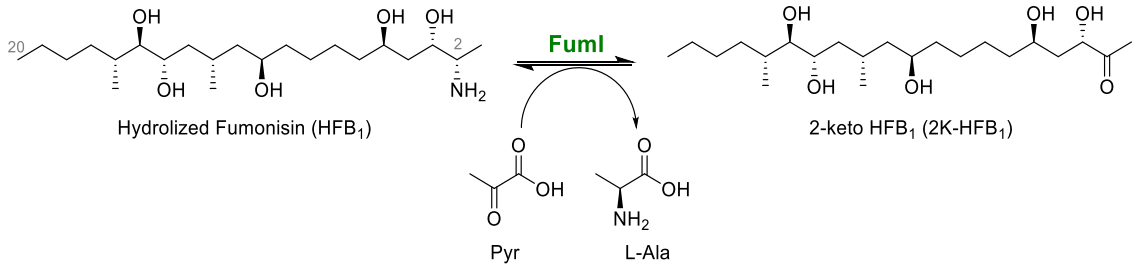


Figure 4.1 Reaction catalysed by the amino-pentol AT FumI in fungus *Sphingopyxis* sp. MTA144. The amino donor HFB_1 transfers the 2-amino group to pyruvate, yielding 2-keto HFB_1 and L-Ala.

Fatty amines, introduced in section 1.5.2, are long chain primary, secondary and tertiary amines from C_8 to C_{22} used in a wide variety of applications, from cosmetic formulations to fabric softeners.[132,133] As the global demand for these compounds is expected to grow finding a green biocatalytic route for their production would be of commercial interest. In this study FumI was explored as a biocatalytic tool for the production enantiopure alkyl amines from aldehydes and ketones of various chain lengths.

4.1.1 Aim

The aim of the project is to express, purify and characterise FumI, assessing its suitability for the scalable biocatalytic production of high value fatty amines. The substrate scope in terms of both amino donor and amino acceptor will be investigated. Crystallographic investigations aimed to shed light on the structural determinants of the enzyme's ability to accept a C_{20} substrate and to identify hot-spots for site directed mutagenesis.

4.2 FumI purification, characterisation

The *fumI* gene (*Sphingopyxis sp.* MTA144, UNIPROT code: D2D3B2), cloned into pET30a (pET-30a-AT144HIS) with NdeI/SacI restriction sites to give a C-terminal His6-tag, was received from Dr. Wulf-Dieter Moll (BIOMIN, Austria).[142] Due to the potential application of FumI in food detoxification, efforts have been made to optimise the expression of this protein. Hartinger D. *et al.* had developed a protocol to obtain soluble AT by expressing the gene in *E. coli* ArcticExpress (DE3) competent cells.[143] Initially, a 10 L grow up in shake flasks was attempted following the procedure described by Hartinger *et al*[143] by expressing in *E. coli* ArcticExpress (DE3) with 0.1 mM IPTG at 11 °C for 24 h. However, the yield of recovered FumI was less than optimal (0.8 mg L^{-1} of culture) and ~ 2-fold less than the reported optimised yield (1.45 mg L^{-1} of culture). Hence, a 12 L fermentation was performed by Dr John White (University of Edinburgh) in order to increase the levels of protein expression. An optimization of the cell lysis method was also carried out, and lysis by cell disruption was found to result in a higher yield of recovered protein when compared to cell sonication by ~ 2-fold. Using these optimized methods typical yields were in the order of ~ 3 mg per litre of culture, almost 2-fold higher than the reported yield of 1.45 mg per litre of culture.[143]

The protein was purified from the cell lysate using nickel immobilized metal affinity chromatography IMAC, where it eluted at a concentration of 150 mM imidazole. A second purification step, on a calibrated size exclusion chromatography (SEC), yielded a symmetrical peak at an elution volume of 75.2 mL (Figure 4.2A), corresponding to a dimer (Appendix 8.5). SDS-PAGE analysis showed the presence of a band at ~ 48 kDa, corresponding to the protein monomer (Figure 4.2B). After the two purification steps the

protein is ~ 90% pure and the presence of other contaminants, including the chaperon Cpn10 at ~ 10 kDa involved in helping the protein folding can be observed (Figure 4.2B).

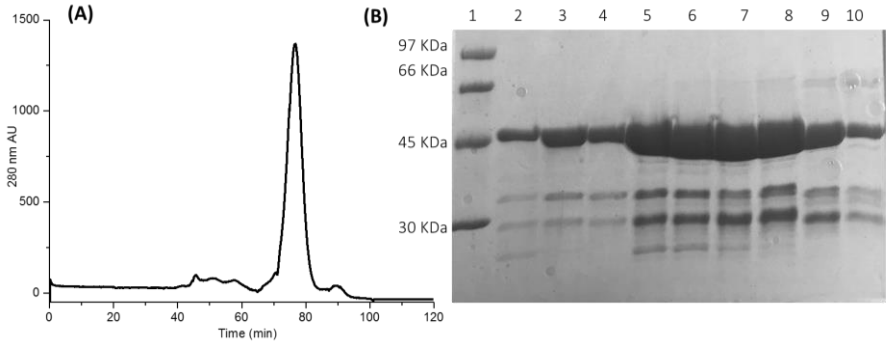


Figure 4.2 **(A)** Chromatogram from Superdex 16/60 S200.FumI elution volume corresponds to a dimer (75.2 mL). **(B)** 12% SDS-PAGE gel of FumI SEC purification. Lane 1: LMW marker (GE), Lane 2: Lanes 2-10: elution fractions from SEC (3 mL).

To ensure the correct protein was expressed and purified, the mass of purified FumI was obtained using denaturing LC ESI-MS. The resulting spectrum (Figure 4.3) showed the presence of several peaks, representing the many charge states of the protein. The observed deconvoluted mass of 48136.21 **±** 0.74 Da was in keeping with the expected mass based on the recombinant protein sequence with the loss of the starting methionine (48271.40 Da, Figure 4.3) , calculated from the expressed protein sequence using the ExPaSy ProtParam tool (https://web.expasy.org/protparam). The loss of the initial methionine (131 Da) is very common as a post translational modification.[146]

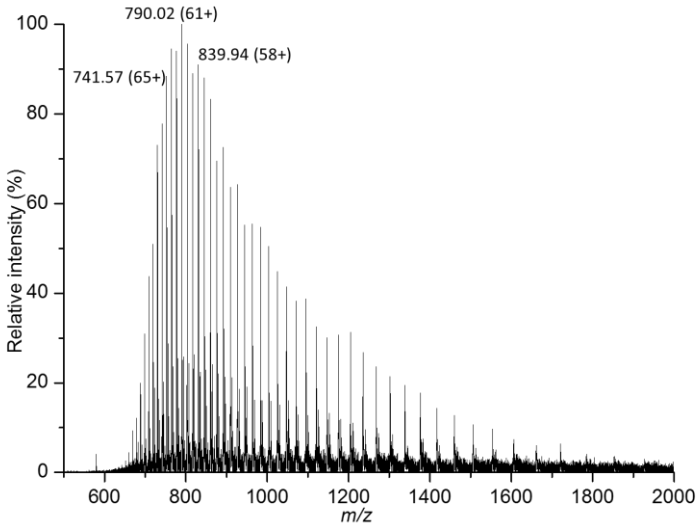


Figure 4.3 Denaturing LC ESI-MS of recombinant FumI (20 µM) with a deconvoluted mass of 48136.21 ± 0.74 Da which agrees with the expected mass based on the expressed protein sequence (48140.40 Da). The values are *m/z* with the charge states given in brackets.

4.3 UV-Vis studies

UV-Vis spectroscopy was used to follow binding of amino acid substrates to FumI. In order to generate the *holo*-form of the enzyme, FumI was dialyzed for two h against buffer containing PLP (20 mM potassium phosphate (KPhos), 150 mM NaCl, 50 µM PLP, pH 7.5). Excess PLP was then removed using a PD-10 desalting column and the resulting sample was analysed by UV-Vis spectroscopy.

The enzyme in its *holo*-form shows a typical PLP binding profile with a λ_{max} at 415 nm and a broad shoulder at 325 nm. When 1 mM L-Ala was added to FumI, changes in the UV/Vis spectrum were observed with a decrease in the absorbance maximum at 415 nm and concomitant increase in the 325 nm region, suggesting the formation of an external aldimine (Figure 4.4). By monitoring the changes in the UV-Vis spectrum at 325 nm (ΔA_{325}) at different L-Ala concentrations, the apparent dissociation constant K_d was calculated (Figure 4.5 A and B, section 5.5.4). The observed K_d for L-Ala (261 µM ± 0.69) is surprisingly low, suggesting a high affinity of FumI for the L-Ala substrate.

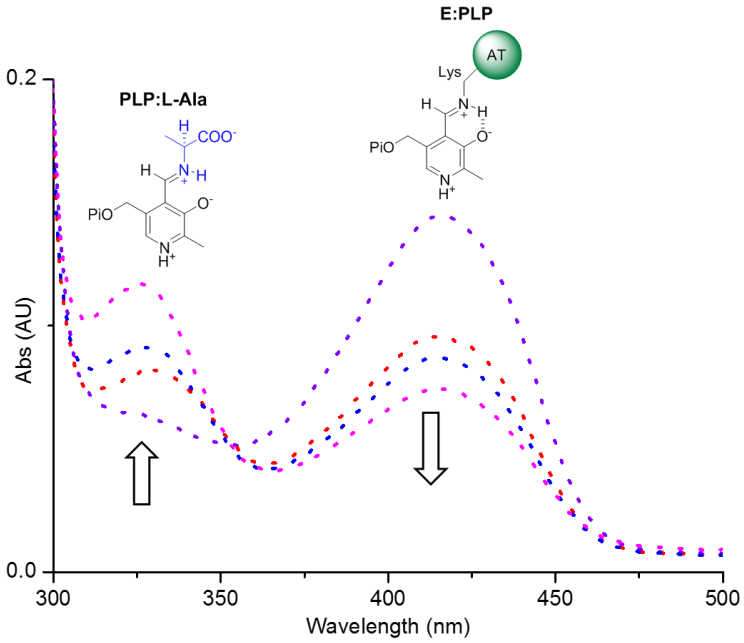


Figure 4.4 UV-Vis profile of FumI:PLP upon L-Ala (1mM) addition over time (2 h). FumI in its *holo*-form absorbs at 415 nm and 325 nm. Upon amino donor addition a decrease in the 425 nm peak and a concomitant increase in the 325 nm peak is observed.

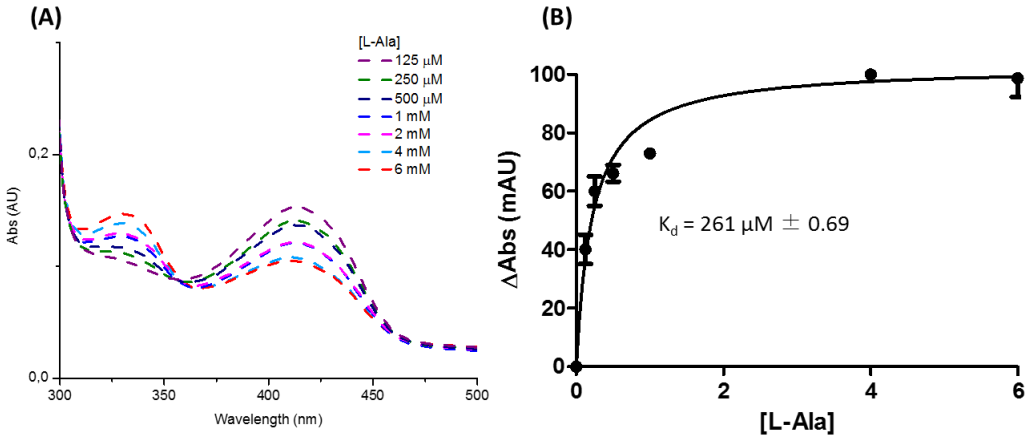


Figure 4.5 (A) UV-Vis spectrum of FumI with varying concentrations of L-Ala showing an increase in the 325 nm peak and a decrease in the 425 nm peak with increasing concentrations of L-Ala (**B**) Plot of the change in absorbance at 325 nm to calculate the K_d for L-Ala. Data was analysed by non-linear regression using Origin. Changes in absorbance at 325 nm were plotted against L-alanine (L-Ala) concentrations, and data points were fitted to a hyperbolic saturation curve.

As deamination of HFB_1 was reported not only in presence of pyruvate but also with α-ketoglutarate, oxaloacetate, glyoxylate, and α ketobutyrate as amino acceptor,[142] binding was tested with the corresponding amino acids (L-Glu, L-Asp, Gly and L-2 aminobutyric acid (L-Aba) respectively). FumI showed the ability to bind each of the amino acids tested. By varying the substrate concentrations and monitoring the ΔA_{325}, the apparent dissociation constant K_d was calculated for each of them (Table 4.1, Figure 4.6).

Amino donor	K_d (mM)
L-Ala	0.261 ± 0.069
L-Aba	0.352 ± 0.042
L-Asp	3.80 ± 1.00
Gly	0.739 ± 0.18
L-Glu	6.20 ± 1.60

Table 4.1 Summary of the dissociation constant K_d for all the tested amino donors. Values have been obtained by plotting the ΔA_{325} at different substrate concentrations.

According to the calculated dissociation constants, FumI displays the highest affinity towards L-Ala with a K_d of 261 µM but also showed tight binding to L-Aba and Gly with a K_d of 352 µM and 739 µM respectively. Weaker binding is observed with L-Glu and L-Asp with

a calculated K_d ~ 15- and ~ 25-fold higher than the K_d for L-Ala (3.80 mM and 6.20 mM respectively).

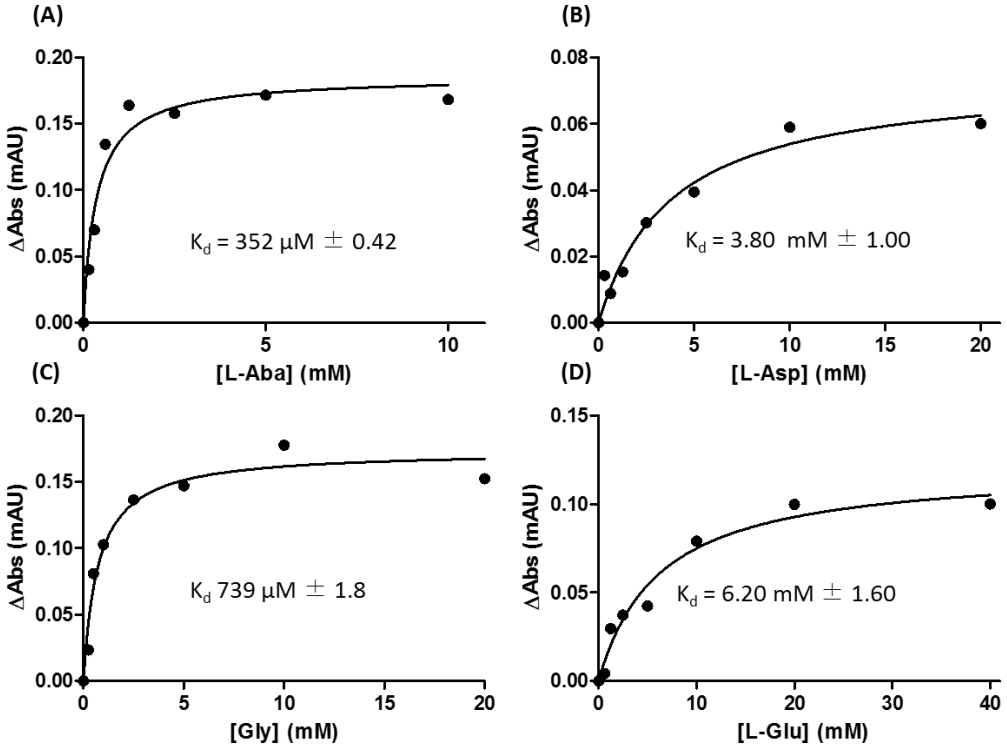


Figure 4.6 Plot of the change in absorbance at 325 nm to calculate the K_d for **(A)** L-Aba, **(B)** L-Asp, **(C)** Gly and **(D)** L-Glu. Data was analysed by non-linear regression using Origin. Changes in absorbance at 325 nm were plotted against amino acid concentration and analysed using Origin.

4.4 FumI/Pyruvate dehydrogenase assay (FumI/PDH)

Kinetic parameters for the 'forward' reaction catalysed by FumI, have been determined by Hartinger *et al.* by using a LC ESI-MS method.[142] This method relied on the peak area ratio of the FB_1 product and HFB_1 substrate in comparison with that of the internal standard ^{13}C-FB_1. However, this method is time consuming and doesn't allow the quick screening of a range of substrates. Thus, a spectrophotometric assay has been developed by coupling FumI with pyruvate dehydrogenase (PDH) from porcine heart, commercially available. PDH converts pyruvate into acetyl-CoA in the presence of CoASH and reducing NAD^+. The production of NADH can be easily monitored at 340 nm, providing a convenient method for assaying FumI transamination activity (ε_{340} = 6220 M^{-1} cm^{-1}).[176] A scheme of the assay is shown in Figure 4.7.

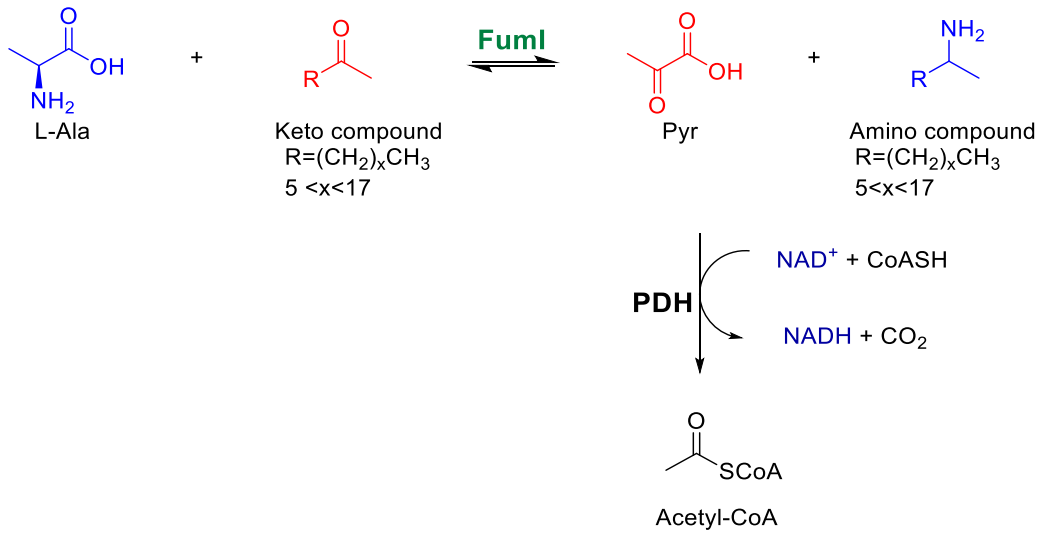


Figure 4.7 Reaction scheme of the FumI/PDH coupled assay. Pyr, product of the first half reaction, is the substrate for PDH. The conversion of NAD^+ and NADH can be easily monitored spectrophotometrically at 340 nm (ε_{340} = 6220 M^{-1} cm^{-1}).

In order to test FumI for the capability of yielding industrially relevant alkyl amines, several long chain substrates that resembled the 2-keto HFB_1 by-product were purchased from Sigma and tested as substrates. These included 2-decanone (C10-2K), 2-tridecanone (C13-2K), 2-pentadecanone (C15-2K) and 2-heptadecanone (C17-2K) tested for the ability to produce 2-decanamine (C10-2A), 2-tridecanamine (C13-2A), 2-pentadecanamine (C15-2A) and 2-heptadecanamine (C17-2A) respectively (**Figure 4.8**). The convention of calling the substrate according to the chain length and the position of the keto or amine group will be adopted in this chapter.

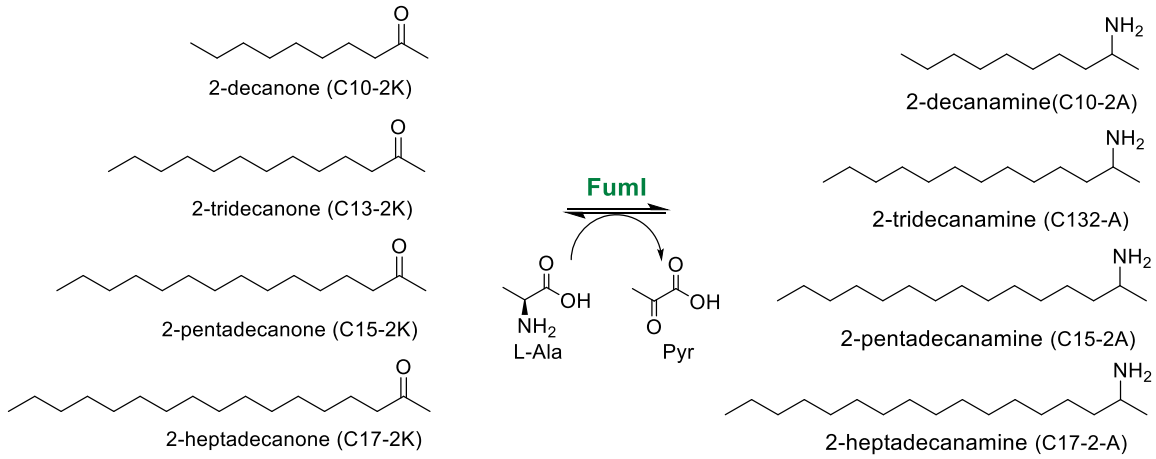


Figure 4.8 FumI reactions for the synthesis of long chain amines.

The keto substrates were tested in order to check differences in the reaction rates at 100µM substrate and 1 mM L-Ala (Figure 4.9). A faster rate is observed for C10-2K and C13-2K suggesting the enzyme preference towards shorter substrates.

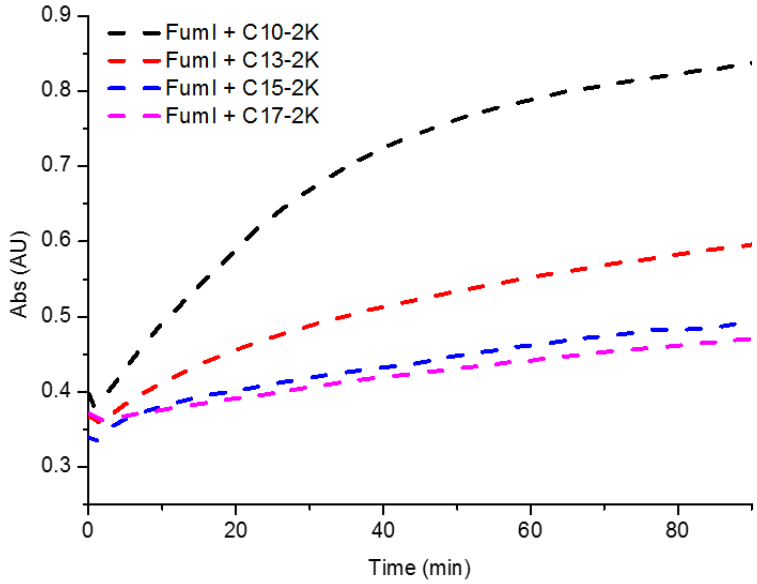


Figure 4.9 Raw data profile of NADH appearance at 340 nm obtained with C10-2K, C13-2K, C15-2K and C17-2K. A concentration of 100 µM amino acceptor and 1 mM amino donor (L-Ala) was used.

By varying the amino acceptor concentration and keeping L-Ala concentration constant (1 mM) it was possible to kinetically characterise the enzyme. The initial rate (over the first 15 min) was used to determine kinetic parameters (Table 4.3, Figure 4.10).

Substrate	K_M (µM)	k_{cat} (10^{-3}) (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
2-decanone (C10-2K)	22.06 ± 12.05	17.91 ± 5.6	812.14 ± 2.98
2-tridecanone (C13-2K)	1.61 ± 0.56	3.1 ± 0.36	1754.34 ± 1.20
2-pentadecanone (C15-2K)	9.97 ± 3.52	1.25 ± 0.30	372.67 ± 1.24
2-heptadecanone (C17-2K)	1.1 ± 0.051	0.45 ± 0.083	539.36 ± 2.17
L-Ala	143.9 ± 26.11	3.29 ± 0.23	22.87 ± 0.39

Table 4.2 Summary of the kinetic parameters obtained using the FumI/PDH assay. The enzyme shows higher affinity towards 2-heptadecanone with a K_M of 16.3 µM compared to shorter substrates.

FumI exhibits higher affinity towards longer substrates such as C13-2K and C17-2K with a K_M of 1.61 and 1.1 µM respectively, compared to C10-2K and C15-2K which show a K_M of 22.6 and 9.97 µM respectively. These value are within the same order of magnitude as the natural substrate HFB_1 that has a reported K_M of 1.1 µM.

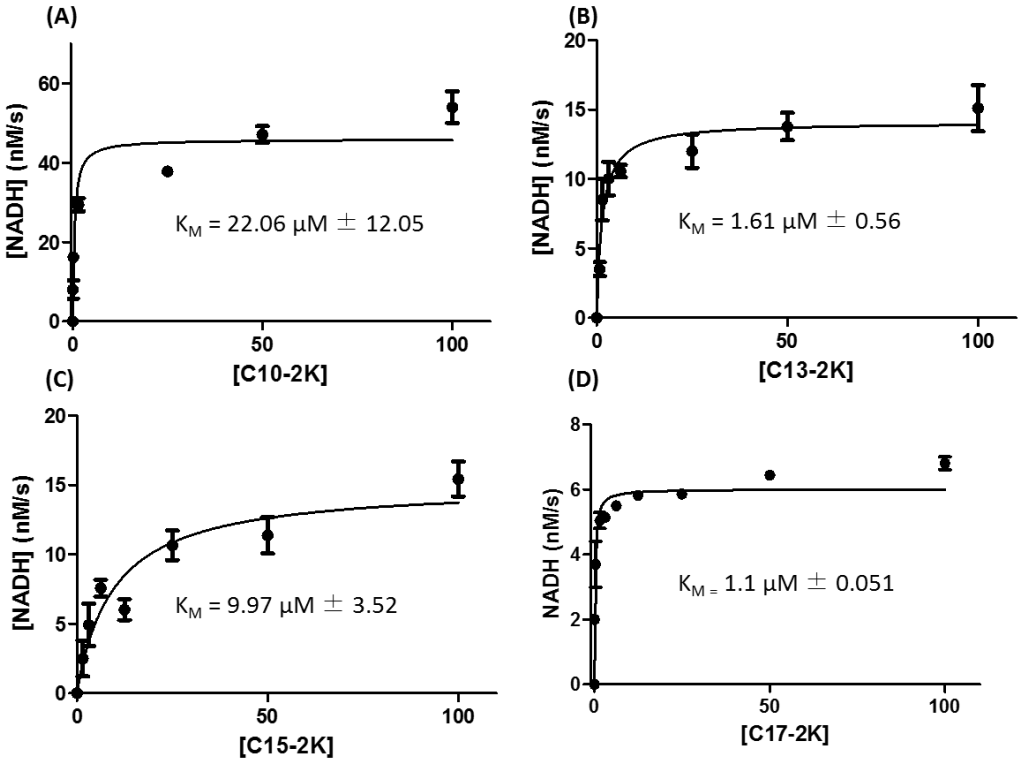


Figure 4.10 Michaelis–Menten plot of FumI for **(A)** C10-2K, **(B)** C13-2K, **(C)** C15-2K and **(D)** C17-2K. Data was analysed by non-linear regression using Graphpad Prism software.

According to the calculated kinetic parameters, the best keto substrate is C13-2K with a high catalytic efficiency (k_{cat}/K_M) of 1754.34 M^{-1} s^{-1}, followed by C10-2K, C17-2K and finally C15-2K. Significantly, the reaction occurs faster with shorter substrates, with a k_{cat} up to 40 times higher for C10-2K compared to C17-2K. An inverse correlation between the chain length of the substrate and the observed turnover rate k_{cat} was observed. A lower turnover rate was observed for longer alkyl chain substrates in comparison to shorter ones. As longer substrates more closely resemble the natural substrate HFB_1, these observations are somewhat counter-intuitive, but this inverse correlation between the length of substrate and catalytic turnover could mean that the enzyme has evolved a way of downregulating the degradation pathway in response to a downstream product.

Because of the low solubility of the keto substrates, 10x stock solutions of each amino acceptor were prepared in dimethyl sulfoxide (DMSO), resulting in a final concentration of 10% DMSO in the assay. Tolerance for DMSO in the reaction is an appealing quality for industrial application as high solvent concentrations are often used.

4.5 Mass Spectrometry analysis

4.5.1 LC ESI-MS analysis

The FumI/PDH coupled assay directly monitors only the first half of the AT reaction and only indirectly fatty amine production. In order to directly monitor the amine product formation, an LC ESI-MS method was used. Reactions were prepared by incubating FumI with 100 µM amino acceptors and 1 mM L-Ala at 35 °C for 18 h in 20 mM KPhos, 150 mM NaCl, 50 µM PLP at pH 7.5, shaking at 250 rpm. These reactions were quenched by dilution into ACN 0.01% TFA and then centrifuged to remove any trace of precipitated protein. The samples were analysed by LC ESI-MS in positive mode. The presence of a signal at an *m/z* matching the expected *m/z* values for the amine products was observed for all the tested reactions suggesting that the enzyme is able to catalyse the formation of C10-2A, C13-2A, C15-2A and C17-2A. The extracted ion chromatograms of the reactions are shown in Figure 4.11 and the expected and observed *m/z* values are summarised in Table 4.3.

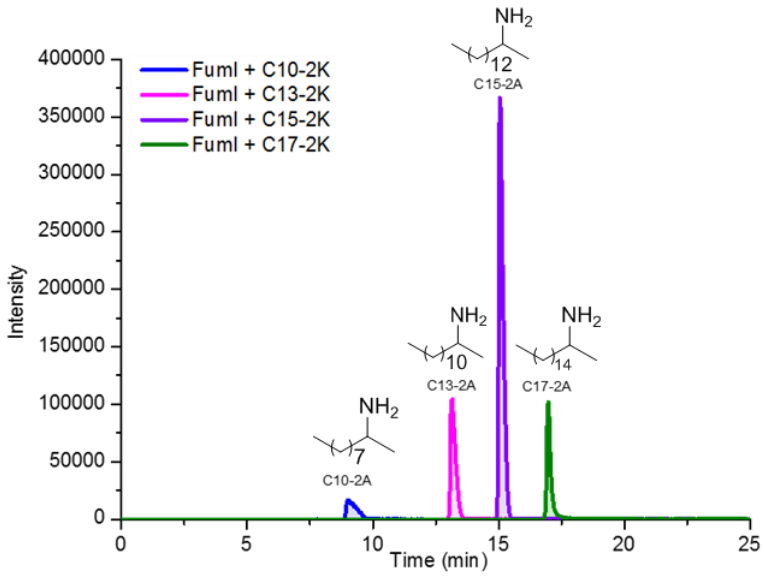


Figure 4.11 Extracted ion chromatogram of the LC ESI-MS FumI reactions with 100 µM of C10-2K, C13-2K, C15-2K and C17-2K and 10 mM L-Ala. The appearance of a peak corresponding to the expected *m/z* value can be observed for each substrate confirming the formation of the amine products.

The ability of FumI to turnover substrates with keto-group substitutions along the alkyl chain (such as 3-decanone (C10-3K) and 5-decanone (C10-5K) (Figure 4.12) was also tested in order to explore the enzyme specificity towards the position of the keto group.

Compound	Expected *m/z*	Observed *m/z*
$C_{10}H_{24}N$ (C10-2A)	158.1909	158.1882
$C_{13}H_{30}N$ (C13-2A)	200.2378	200.2370
$C_{15}H_{34}N$ (C15-2A)	228.2691	228.2704
$C_{17}H_{38}N$ (C17-2A)	256.3004	256.3038

Table 4.3 Summary of the expected and observed *m/z* values for the FumI tested reactions.

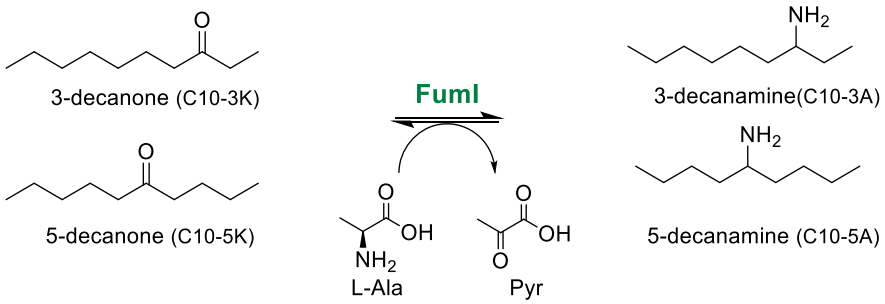


Figure 4.12 Tested reaction for the substrate specificity towards the position of the keto group.

Enzyme assays were performed as described above and product formation was analysed by LC-MS. Surprisingly FumI was capable of utilizing C10-3K, as a peak of 158.1875 Da matching the expected *m/z* value for C10-3A was observed in the corresponding reaction (Figure 4.13).

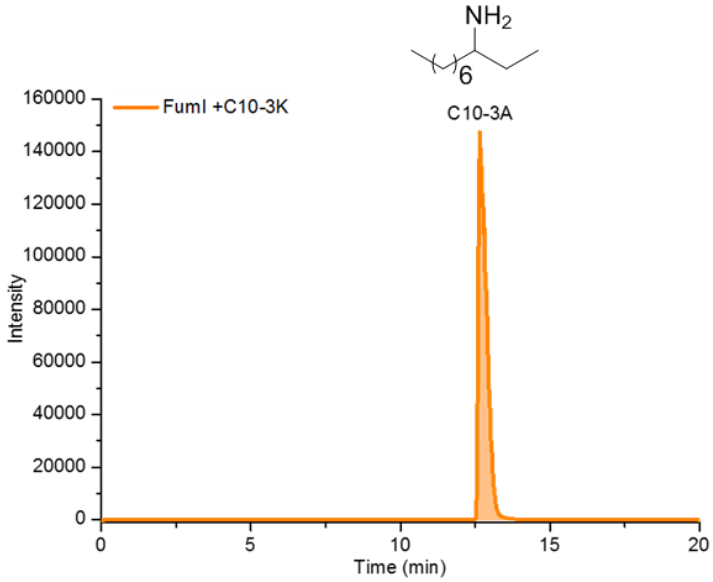


Figure 4.13 Extracted ion chromatogram of the LC ESI-MS from FumI reaction with 100 µM C10-3K and 10 mM L-Ala. An *m/z* value of 158.1875 Da matches the expected *m/z* of 158.1909 Da suggests the formation of the C10-3A product.

However, no peak corresponding to the expected *m/z* value was observed for C10-5K, suggesting that the enzyme has relatively narrow regioselectivity.

4.5.2 MALDI-MS analysis

The ESI LC-MS method described above relies on a C18 column for the LC phase, therefore it was not ideal for testing products shorter than C_{10} (without significant adjustment of the method). Therefore, a Matrix-assisted laser desorption/ionization (MALDI)-MS approach was applied. By using α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix, it was possible to detect product formation for the tested reactions in positive mode. This method has been successfully applied within our group for the detection of sphingolipid products such as 3-ketodihydrosphingosine (KDS) (developed by Dr Van Kelly, University of Edinburgh, data not published) and therefore tested for FumI reactions. A positive control with C10-2A was carried out during this experiment.

This MALDI method was used to explore the chain length specificity of FumI in more detail. In particular, the ability of the enzyme to accept substrates shorter than C_{10} was tested. Hence, FumI was tested with the following amino acceptors: acetone (C3-2K), butanone (C4-2K), hexanone (C6-2K) and octanone (C8-2K) to yield isopropylamine (C3-2A), 2-butylamine (C4-2A), 2-aminohexane (26-2A) and 2-aminooctanone (C8-2A) respectively (Figure 4.14, Figure 4.15).

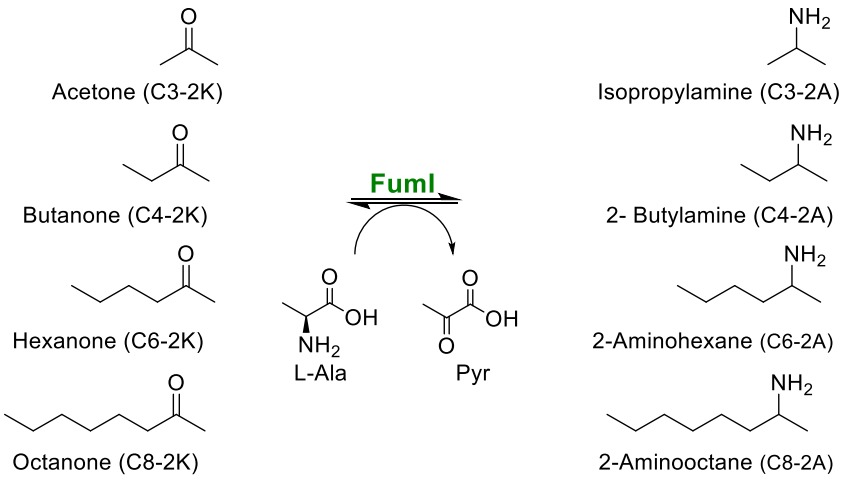


Figure 4.14 FumI reactions for the synthesis of alkyl chain substrates shorter than C_{10}. The amino acceptor tested are: acetone (C3-2K), butanone (C4-2K), hexanone (C6-2K) and octanone (C8-2K). The amino acceptors were dissolved in 100% DMSO for a final concentration of 10% DMSO and used at a final concentration of 100 µM. L-Ala was used at a final concentration of 10 mM.

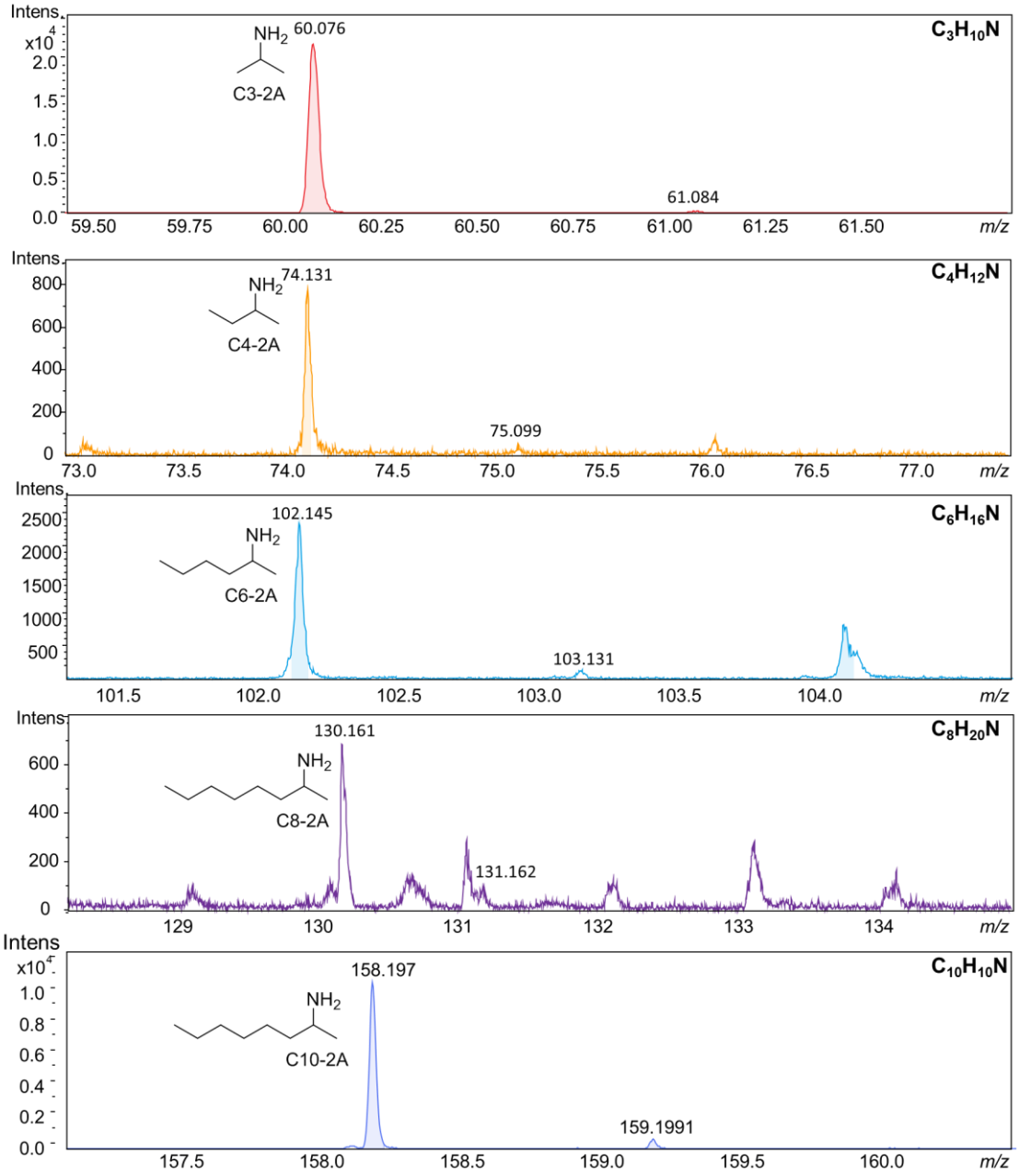


Figure 4.15 MALDI-MS spectrum of the FumI reactions carried out with C3-2K, C4-2K, C6-2K C8-2K and C10-2K and L-Ala. The presence of a peak corresponding to the expected *m/z* value of the amine product can be observed for all the tested reactions.

Surprisingly, the formation of the corresponding amine was observed for substrates down to C_3 in length, suggesting that the enzyme is quite promiscuous and could be potentially

used also for the synthesis of shorter amine products (Figure 4.15). As the natural substrate is 20 carbon long (FB_1) such a chain length tolerance for the FumI enzyme was totally unexpected.

A summary of the expected and observed *m/z* values is shown in Table 4.4.

Compound	Expected *m/z*	Observed *m/z*
$C_3H_{10}N$ (C3-2A)	60.081	60.076
$C_4H_{12}N$ (C4-2A)	74.097	74.131
$C_6H_{16}N$ (C6-2A)	102.128	102.145
$C_8H_{20}N$ (C8-2A)	130.159	130.161
$C_{10}H_{24}N$ (C10-2A)	158.190	158.197

Table 4.4 Summary of the expected and observed *m/z* values for the FumI tested reactions.

The unexpected promiscuity displayed by FumI in accepting substrates ranging from C_3 to C_{17} represents a very interesting property for biocatalysis as it could be used for the production of a range of amines.

4.6 Studying FumI enantioselectivity

The enantioselectivity of FumI AT was determined by using Marfey's reagent, 1-fluoro-2, 4-dinitrophenyl-5-L-alanine amide (FDAA), a pre-column derivatizing reagent used for the separation of enantiomeric isomers of amine compounds (Figure 4.16).[177–179] The use of Marfey's reagent for derivatization of amino acids was firstly reported by Marfey, back in 1984.[180] FDAA reacts by nucleophilic substitution of the aromatic fluorine with the free amino group on the amino acid or amino target molecule. As Marfey's reagent has a chiral centre in the L-Ala group (*S*) when a racemic mixture of an amino compound is treated with this compound the production of analogous diastereomers is observed. The produced diastereomers can be easily separated using reversed-phase high-performance liquid chromatography (HPLC).[145] As the (*R*)- and (*S*)-standards were commercially available for

only the 2-aminooctanone (C8-2A), this was chosen to investigate FumI enantioselectivity (Figure 4.16).

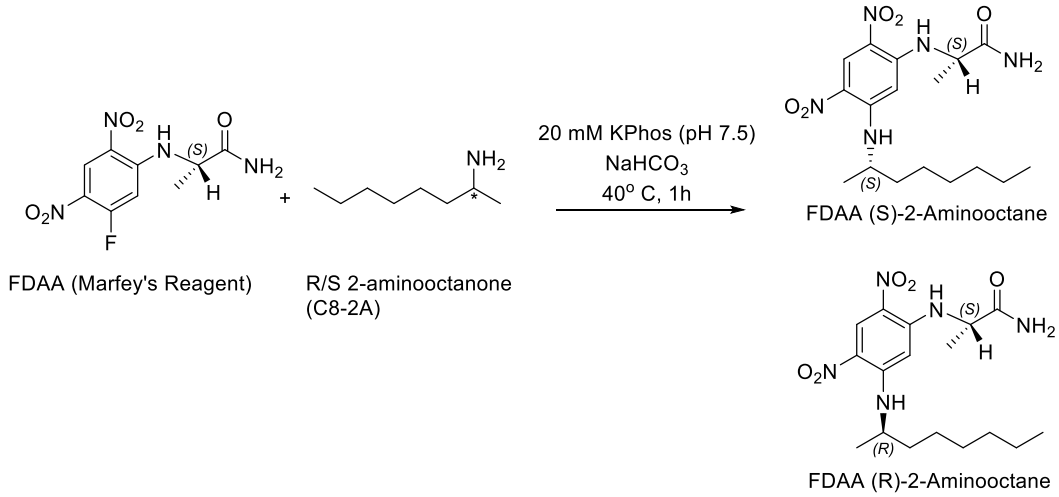


Figure 4.16 Reaction carried out with FDAA and both enantiomers of aminoctanone to check FumI enantioselectivity.

By reviewing Marfey's reagent literature it was discovered that each amine required optimization in order to obtain complete derivatization and enantiomeric separation. Different reaction times (up to 24 h) and different bases (the use of disodium tetraborate decahydrate in place of sodium carbonate) can be used to drive the reaction towards completion.[176– 178,181,182] Different columns and HPLC mobile phases have been reported to obtain enantiomeric separation.

1 mM (*R*) and (*S*)-C8-2A were derivatized using $NaHCO_3$ as a base and the reaction was carried out for 1 h at 40 °C. The reaction was then cooled to room temperature and quenched by the addition of HCl and analysed by reversed phase HPLC using a C18 column. In order to obtain separation of the two diasteroisomers (*S,S* and *S,R*) a HPLC method was developed with a gradient up to 70% ACN and a method length of 50 min (section 5.10). Using this method the retention times for the (*R*) and (*S*)-C8-2A were 42.75 and 43.5 min respectively (Figure 4.17, red dash line).

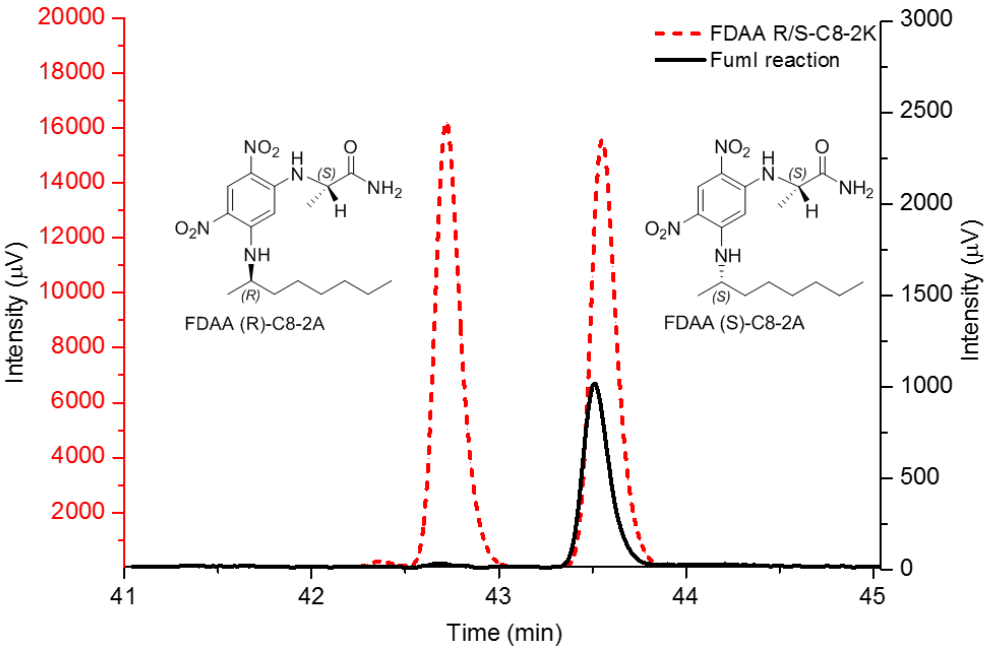


Figure 4.17 C18 HPLC analysis on a calibrated C18 column at λ = 340 nm. The (*R*)- and (*S*)-C8-2A standards (1 mM, dash red line) are well resolved and the AT reaction carried out for 18 h at 35 °C (10 mM L-Ala, 100 µM C8-2A) shows the formation of only the (*S*)-C8-2A enantiomer (black line). The retention times are 42.75 min for (*R*)-C8-2A and 43.5 min for (*S*)-C8-2A respectively.

A reaction was then carried out with FumI and 100 µM C8-2K and 1 mM L-Ala at 35 °C for 18 h in 20 mM KPhos, 150 mM NaCl, 50 µM PLP at pH 7.5,shaking at 250 rpm. The biocatalytic reaction was subjected to the same protocol used for the (*R*) and (*S*)-C8-2A derivatization and analysed by C18 HPLC. Based on the presence of the peak at a retention time corresponding to the FDAA (*S*)-C8-2A enantiomer it can be concluded that the chirality of the final product is (*S*) (Figure 4.17, black line). Thus, it can be concluded that FumI displays (*S*)-enantioselectivity.

In order to achieve complete derivatization and use this assay for quantitative analysis, reaction time and reagents used must be optimised. Moreover, because of the unavailability of commercial standards for all the amines produced using FumI as a catalysts, this assay is not suitable for quantitative analysis.

4.7 Amino donor promiscuity

Having established that FumI has a reasonably wide substrate scope in terms of amino acceptor, FumI was also analysed for its amino donor promiscuity. In particular, attention

was focused on non-natural 'smart' amine donors such as *ortho*-xylylene diamine (OXD),[102] 5-nitro-2, 3-dihydro-1H-inden-2-amine (2,3-DHIA), 4-(2-aminoethyl) benzonitrile (4-AEB) and 2-(4-nitrophenyl) ethan-1-amine (4-NPEA),[103] introduced in section 3.3.4.2 and 3.3.4.3 respectively.

Changes in the UV-Vis profile of FumI were observed upon OXD addition, suggesting the capability of the enzyme of using it as an amino donor (Figure 4.18). A first shift towards the external aldimine (325 nm) could be observed, followed, overtime, by the appearance of an extra peak around ~ 550 nm corresponding to an isoindole moiety, resulting from the production of the isoindole product. However, a very weak purple colour was observed suggesting a low substrate conversion. Moreover the time required to observe the formation of the purple colour (>8 h) suggested a slow reaction rate.

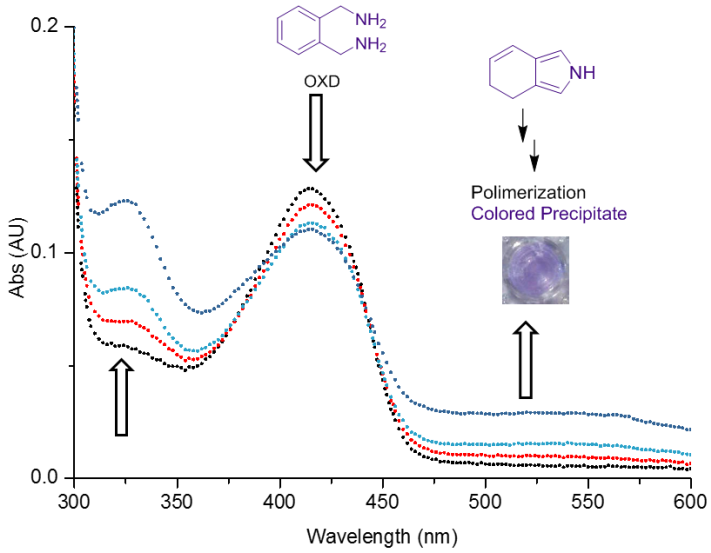


Figure 4.18 FumI UV-Vis studies with OXD as amino donor at 1 mM. Over time the internal aldimine peak at 415 nm disappears and the external aldimine formation at ~ 325 nm is observed. Furthermore, an additional peak at ~ 550 nm corresponding to the isoindole formation is observed suggesting the occurrance of the AT reaction. The formation of a weak purple precipitate can be observed after ~ 8 h of reaction.

The 'smart' amino donors developed by Prof. Helen Hailes (UCL), introduced in section 3.3.4.3, were tested with a wide range of amino acceptors, from C_3 to C_{17} ketones to benzaldehyde (BZALD) and butyraldehyde (NBA) (Figure 4.19).

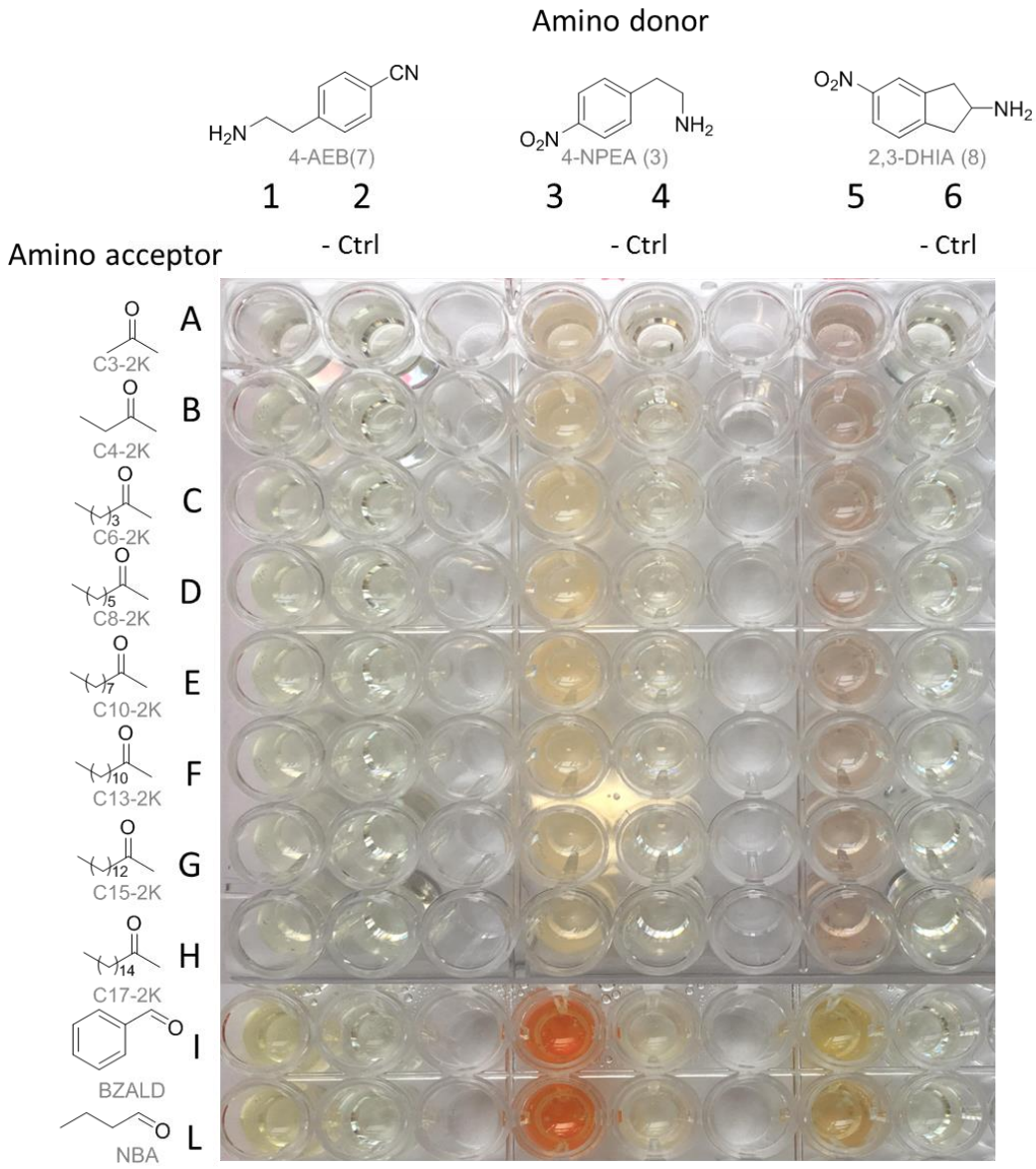


Figure 4.19 Screening of the amine donors 7 (1, 2), 3 (3, 4) and 8 (5, 6) with different amino acceptors such as C3-2K (lane A), C4-2K (lane B), C6-2K (lace C), C8-2K (lane D), C10-2K (lane E), C13-2K (lane F), C15-2K (lane G), C17-2K (lane H), BZALD (lane I) and NBA (lane J) with FumI (0.5 mg mL^{-1}). Amine donors were used at final concentrations of 12.5 Mm, amino acceptors at 100 µM for the keto compounds from C3-2K to C17-2K for solubility reasons, and 5 mM for BZALD and NBA AS suggested in the literature.[153] For each reaction a control with no enzyme was performed (lane 2, 4, 6). Reactions were incubated for 18 h at 37 °C. No significant difference is observed between the reaction and the controls for most of the amino acceptors. A very intense colour is formed when BZALD and NBA are used as amino acceptor.

The formation of a vividly coloured precipitate is observed when the amino donors 4-NPEA (3), 4-AEB (7) and 2,3-DHIA(8) are used with BZALD and NBA as amino acceptors. No significant differences are observed between the reaction (lane 1) and the negative control

(lane 2) with 4-AEB (7) as amino donor and keto compounds ranging from C_3 to C_{17} as amino acceptors. The formation of a pale coloured precipitate is observed when the amino donors 4-NPEA (3) and 2,3-DHIA(8) are incubated with a range of alkyl chain keto acceptors (lane 3 and lane 5). These results are quite surprising as the ketones tested (from C3-2K to C17-2K) were converted by the enzyme when L-Ala is used as amino donor (Figure 4.11, Figure 4.15). This could be due to the lower concentration of the keto acceptor used in the assay (1 mM), compared to the concentration of the aldehydes (5 mM), due to the low solubility of the keto compounds.

The formation of coloured precipitates, observed with BZALD and NBA, suggests an unexpected substrate promiscuity. FumI's capability of accepting aromatic substrates such as BZALD could be further explored as the formation of AT product benzylamine (BZA) can be easily monitored spectrophotometrically.[103]

4.8 FumI and aldehydes

FumI's ability to use aldehyde substrates, as observed in the amino donor assay, was further explored. In particular, FumI was tested for the capability of converting decanal (C10-Al), dodecanal (C12-Al) and hexadecanal (C16-Al) into the corresponding primary amines decylamine (C10-1A), dodecylamine (C12-1A), and hexadecylamine (C16-1A) respectively (Figure 4.20).

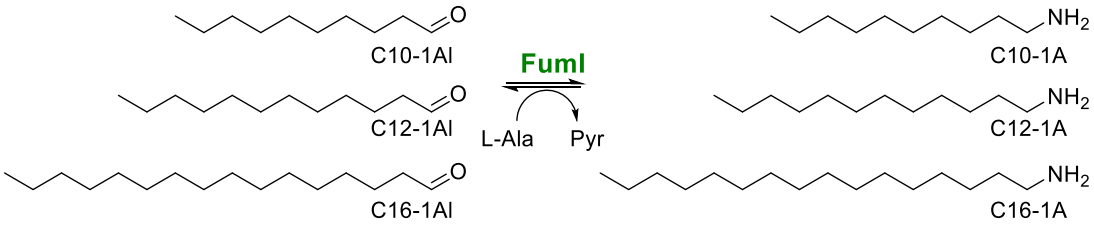


Figure 4.20 Reactions tested with C10-Al, C12-Al and C16-Al as amino acceptor and L-Ala as amino donor to yield C10-1A, C12-1A and C16-1A and pyruvate.

The reactions were carried out following the protocol established for the ketone and product formation was analysed by LC ESI-MS. The presence of a peak matching the expected *m/z* values was observed for all the tested reactions suggesting that the enzyme is able to catalyse the formation of C10-1A, C12-1A and C16-1A. The extracted ion

chromatogram of the reactions are shown in Figure 4.21 and the expected and observed *m/z* values are summarised in Table 4.5.

Compound	Expected *m/z*	Observed *m/z*
$C_{10}H_{24}N$ (C10-1A)	158.1909	158.1875
$C_{12}H_{28}N$ (C12-1A)	186.2222	186.2159
$C_{16}H_{36}N$ (C16-1A)	242.2848	242.2776

Table 4.5 Summary of the expected and observed *m/z* values for the FumI tested reactions with C10-Al, C12-Al and C16-Al.

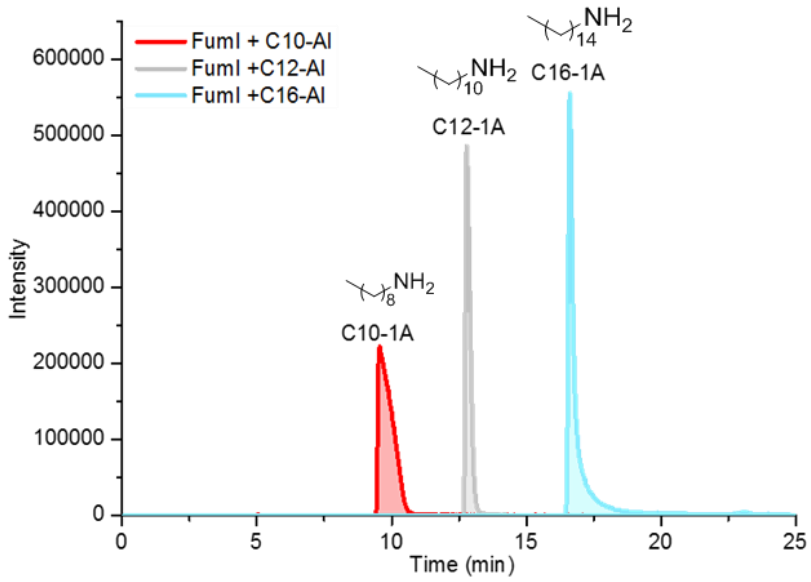


Figure 4.21 Extracted ion chromatogram of the LC ESI-MS FumI reactions with 100 µM of C10-Al, C12-Al, C16-Al and 10 mM L-Ala. The appearance of a peak corresponding to the expected *m/z* value can be observed.

These results reinforce the application of FumI as a biocatalyst for the synthesis of fatty amines as it is able to catalyse the formation of primary amines at the first and second position.

4.8.1 Development of a FumI and CAR cascade

Carboxylic acid reductase (CAR) enzymes catalyse the reduction of carboxylic acids to the corresponding aldehyde *via* nicotinamide adenine dinucleotide phosphate (NADPH) and adenosine 5' phosphate (ATP) consumption (Figure 4.22).[184,185] The hydrolysis of ATP makes the reduction of acids into aldehydes, catalysed by CAR, thermodynamically favourable,

constituting a very attractive green chemical route for the production of aldehyde from cheap carboxylic acids under mild reaction conditions, without the requirement of harsh reducing agents.

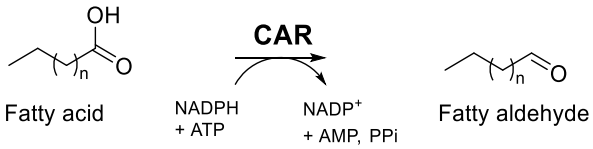


Figure 4.22 Generic reaction scheme for a reaction catalysed by a CAR: a fatty acid is converted into the corresponding aldehyde *via* consumption of NADPH and ATP.

CARs are multidomain enzymes around 130 kDa, consisting of an N-terminal adenylation domain (A-domain) fused *via* a pepidyl carrier protein (PCP) to a C-terminal thioester reductase domain (TR-domain).[186,187]

An elegant *de novo* route in yeasts has been reported for the biosynthesis of vanillin, widely used in fragrances and body care products, starting from glucose using a CAR enzyme from *Nocardia iowensis* (NiCAR) (Figure 4.23 A).[188] CARs have also been recently applied for the production of enantiopure chiral amine building blocks in a cascade reaction by coupling them with an AT and an IRED (Figure 4.23 B).[189]

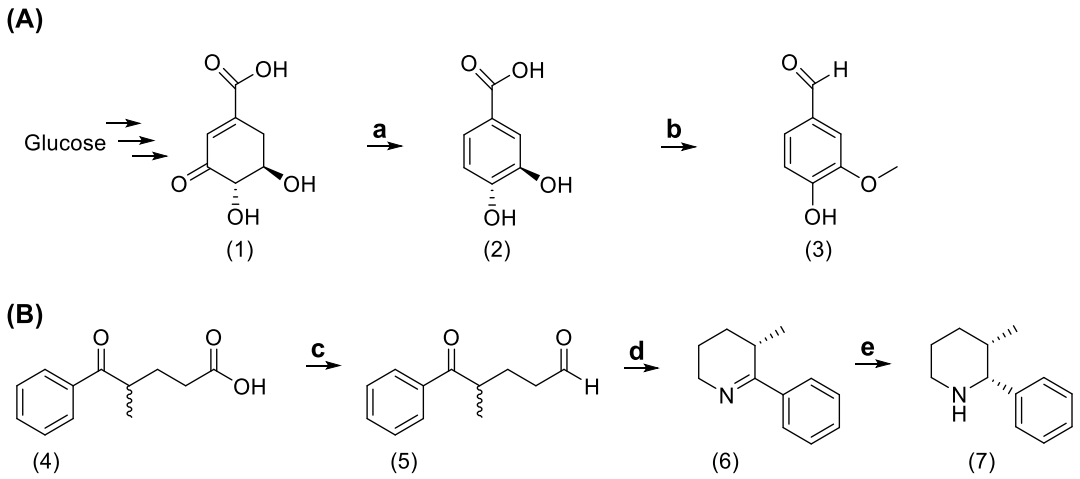


Figure 4.23 Selected examples of cascades developed using CAR enzymes. **(A)** *De novo* biosynthetic pathway from glucose to produce vanillin. The enzyme 3-dehydroshikimate dehydratase (a) catalyses the conversion of 3-dehydroshikimic acid (1) into 3,4-dihydroxybenzoic acid (2) . A CAR enzyme from *Nocardia iowensis* (NiCAR) and an o-methyltransferase (b) reduce the acid to the aldehyde and introduce a methyl group respectively to produce vanillin (3). **(B)** One-pot cascade where the ketoacid (4) is reduced to the aldehyde (5) by *Mycobacterium marinum* CAR (MmCAR) (c). The ATA-113 (d) (Codexis) catalyses the transamination reaction of compound (5), the product undergoes intramolecular cyclization to imine (6). An (*S*)-IRED (e)produces the final (7) compound. Adapted from Winkler *et al.*[185]

As FumI is able to accept long chain aldehydes as substrates the possibility of combining FumI with a CAR in a cascade process was explored, enabling the conversion of simple, inexpensive fatty acids into long chain fatty amines (Figure 4.24).

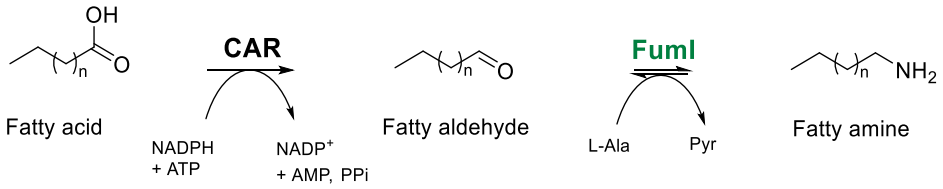


Figure 4.24 Generic reaction scheme of the CAR-FumI cascade. The fatty acid is first converted into aldehyde by CAR and then further converted into the final desired fatty amine by FumI.

In collaboration with Prof. Nick Turner (University of Manchester) and Dr. Sasha Derrington (University of Manchester), a library of CAR enzymes were screened against a range of saturated fatty acids (C_6-C_{12}) by monitoring NADPH depletion versus time at 340 nm. A CAR from *Mycobacterium chlorophenolicum* (McCAR), was identified as the most active for the reduction of long chain ($\geq C_{10}$) fatty acids. The McCAR was previously identified by Finnigan W. *et al.* but not characterised in the study.[190]

CARs require NADPH and ATP for catalysis, which would cause an increase in the production costs for a commercial biocatalytic process. To overcome that, whole cell approaches or cofactor recycling are usually applied. Recycling systems are available for both NADPH and ATP by utilizing the glucose dehydrogenase (GDH) and the adenylate kinase (AdK)/ adenosine 5'-monophosphate (AMP) phosphotransferase (PAP) enzymes, allowing a more sustainable process (Figure 4.25).[49]

In order to check the compatibility of the FumI/McCAR reactions several conditions were tested including the presence or the absence of the recycling system that could potentially affect the reaction efficiency. McCAR and FumI were coupled and the reactions were tested in a cascade process. 1 mM decanoic (C10-FA), lauric (C12-FA), myristic (C14-FA), palmitic (C16-FA), stearic (C18-FA) and arachidic acid (C20-FA) were incubated with 10 mM L-Ala, 2 mM NADPH, 2 mM ATP, 20 mM $MgCl_2$, FumI (0.24 mg mL^{-1}) and CAR (0.25 mg mL^{-1}). Reactions were incubated at 30 °C for 18 h (maximum operational temperature for McCAR) shaking at 200 rpm. Moreover, the CAR-FumI-recycling system (AdK/PAP/GDH) was checked with the reaction with C10-FA as a model substrate. This reaction was carried out

in one pot (all enzymes in one reaction tube) and in two steps (first CAR for 8 h and then FumI was added) and with and without the recycling system.

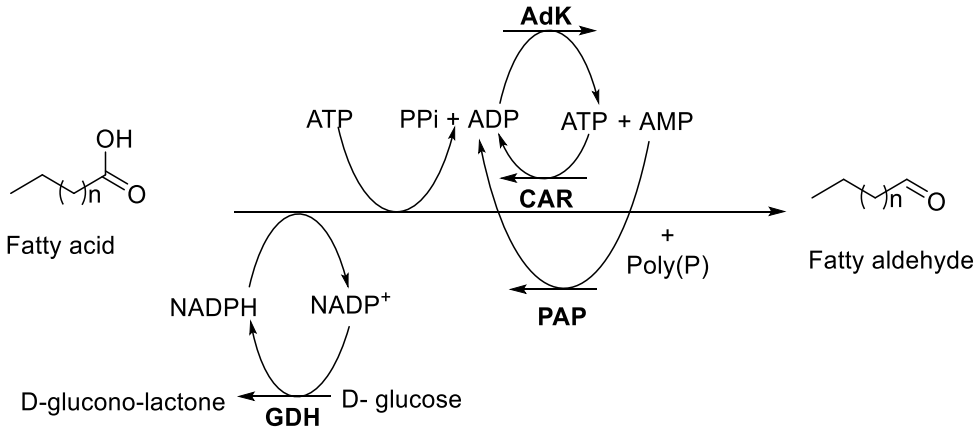


Figure 4.25 Regeneration system used to recycle NADPH and ATP. The $NADP^+$ produced by the reaction is converted into NADPH by GDH and conversion of D-glucose into D-gluconolactone. The adenosine 5'-diphosphate (ADP) generated by the reaction is converted into ATP and AMP by AdK. In turn, the AMP is phosphorylated by PAP to ADP. Poly (P) is the phosphoryl donor.

For all the tested conditions a peak corresponding to the expected *m/z* values for the production of the amines (from C10-1A to C20-1A) was observed (Table 4.6). The observed product formation for all the tested conditions suggests that the enzymes are compatible to work together and that the presence of the recycling system does not affect the enzyme activity. Thus, for potential industrial applications, the recycling system could be used, avoiding the requirement of continuous addition of expensive cofactors and making it more attractive for industrial biocatalysis.

These exciting preliminary data suggest that the FumI/CAR system could be used in a cascade, enabling the conversion of an inexpensive substrate such as a fatty acid into a high value fatty amine. Substrate conversion and reaction rate will need to be determined in order to identify the best reaction conditions and scale up the reaction to an industrially relevant level. The possibility of coupling the enzymes using a whole cell approach could be explored in order to reduce the reaction costs and the requirement of cofactor regeneration systems.

Compound	Expected *m/z*	Observed *m/z*
$C_{10}H_{24}N$ (C10-1A)	158.1909	158.1930
$C_{12}H_{28}N$ (C12-1A)	186.2222	186.2240
$C_{14}H_{32}N$ (C14-1A)	214.2535	214.2550
$C_{16}H_{36}N$ (C16-1A)	242.2848	242.2855
$C_{18}H_{40}N$ (C18-1A)	270.3161	270.3111
$C_{20}H_{44}N$ (C20-1A)	298.3474	298.3409

Table 4.6 Summary of the expected and observed *m/z* values for the FumI/McCAR tested reactions with C10-FA, C12-FA, C14-FA, C16-FA, C18-FA and C20-FA to yield C10-1A, C12-1A, C14-1A, C16-1A, C18-1A and C20-1A respectively.

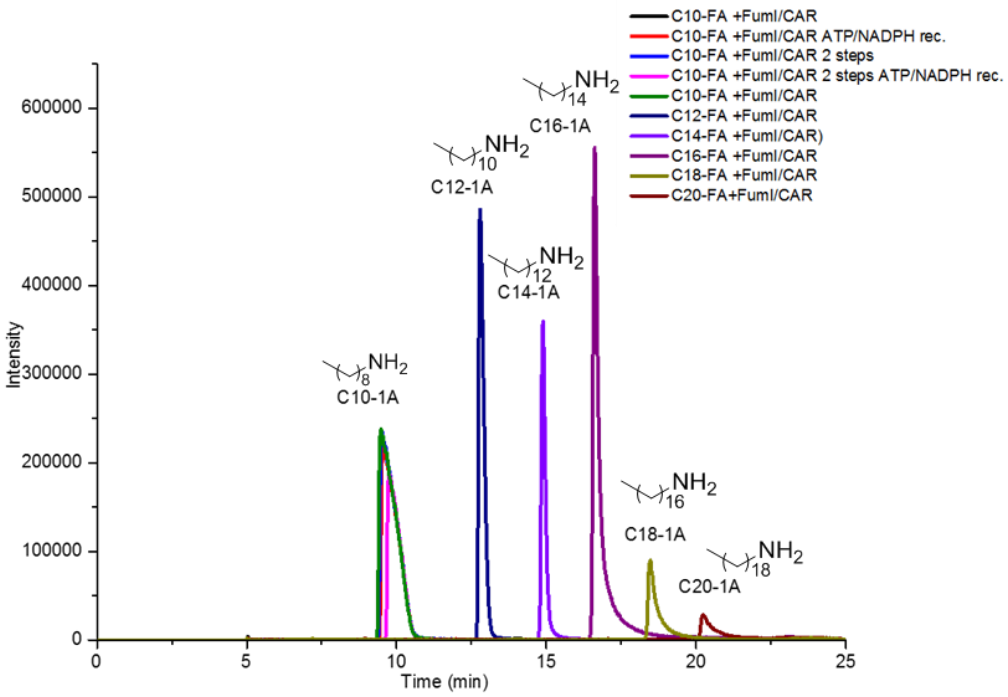


Figure 4.26 Extracted ion chromatogram of the LC ESI-MS coupled reactions with FumI and McCAR exploring different conditions. The appearance of a peaks with the expected *m/z* value corresponding to the production of C10-1A, C12-1A, C14-1A, C16-1A, C18-1A and C20-1A can be observed.

4.9 High resolution structural studies of FumI

To aid rational modification and the understanding of the substrate scope of FumI, the determination of the 3D structure of the enzyme is crucial. In collaboration with Dr. Jon Marles-Wright (Newcastle University), crystal screening with purified FumI (8.1 mg mL^{-1}) in 20 mM KPhos, 500 mM NaCl, 50 µM PLP, pH 7.5 were performed using a wide range of

commercial screens (MIDAS, PEG/Ion, STRUCTURE, JCSG$^+$) in 96 well plates using a Gryphon robot. Crystals were observed under only a few conditions (A4 JCSG$^+$, B4 PEG/ION, C10 MIDAS). The most suitable crystals for X-ray diffraction resulted from C10 MIDAS (30% w/v polyacrylate 2100, sodium salt, 0.2 M ammonium sulfate, 0.1 M HEPES pH 6). These crystals were fished, frozen in cryoprotectant and shot at the Diamond Light synchrotron, Harwell for X-ray diffraction analysis (Figure 4.27).

Several diffraction-quality datasets were collected (Diamond Light) with the highest resolution of 1.60 Å undergoing full refinement to generate structural information (Appendix 8.9). The solving of the FumI structure was carried out by Dr Jon Marles-Wright (Newcastle University) by molecular replacement using the structure of the glutamate 1-semialdehyde AT from *Bacillus subtilis* (PDB code: 3BS8), displaying 29% sequence identity. The final refined model contains one homodimer of FumI with the PLP cofactor present as an internal aldimine with residue Lys256 in each chain. The refined protein structure is deposited under the PDB code 6HBS.

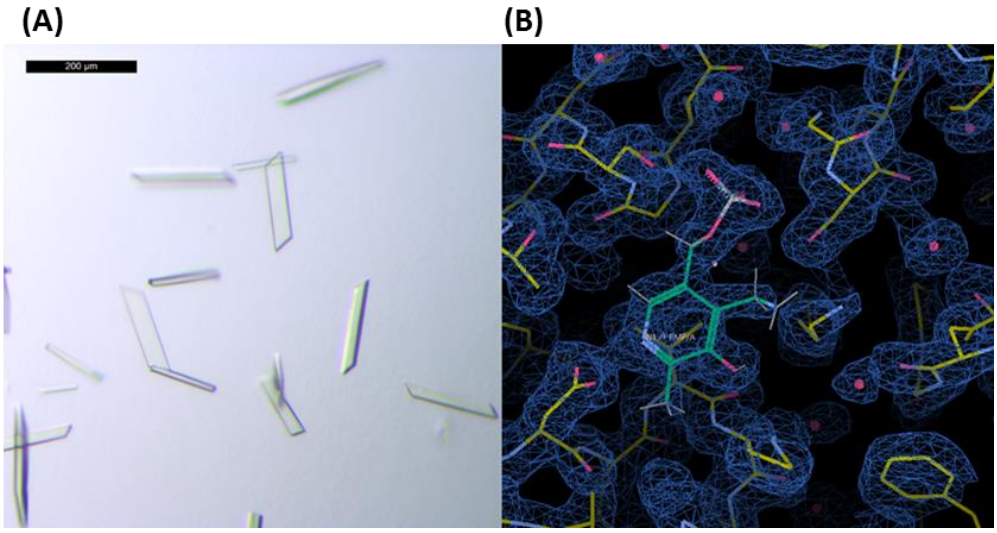


Figure 4.27 (A) Crystal morphology of yellow crystals produced by adding 1 µL of 8.7 mg mL^{-1} FumI in 20 mM KPhos, 500 mM NaCl, 50 µM PLP, pH 7.5, to 1 µL of 30% w/v polyacrylate 2100, sodium salt, 0.2 M ammonium sulfate, 0.1 M HEPES pH 6 (**B)** Electron density map of the active site containing the PLP cofactor contoured at 1 σ.

The overall fold of FumI is typical of the class III AT family (PLP fold type I). FumI is a homodimer (Figure 4.28 A) where each monomer consists of a small and large subunit with the PLP lying at the dimer interface. The monomeric FumI structure can be subdivided into two distinct domains: a small discontinuous domain comprising the residues 1-59 and 308-

417; and a large domain formed by residues 60 to 307 (Figure 4.28 B). The N-terminal portion of the small domain consists of an α-helix followed by a three-stranded antiparallel β-sheet, whilst the C- terminal part consists of an α-helix followed by two antiparallel β-strands, an extended loop interrupted with a short α-helix, followed by a longer α-helix. The large domain consists of the typical central seven- β-strands, with five parallel strands and two in an anti-parallel orientation. The β-strands are connected by α-helices, with the cofactor binding site at the dimerization interface.

The active sites of FumI are located at the dimeric subunit interface, containing amino acids from both subunits. Each active site contains a PLP-binding pocket and a substrate binding region. The cofactor is bound by a covalent Schiff base linkage to the ε-amino group of the catalytic Lys256. The presence of residues involved in PLP stabilization is highlighted in Figure 4.29. The PLP pyridine ring is sandwiched between Tyr143, opposite edge-on against the PLP ring, and Val230, as common in ATs. The nitrogen of the pyridine ring is coordinated by a hydrogen bond to Asp229, acting to maintain the cofactor in the protonated form.[191] The phosphate group of the cofactor is coordinated by a series of hydrogen bonds with Gly117 and Ser118 from one monomer (monomer A) and Thr283, Gly284 and Ser285 from the other one (monomer B), directly or through water molecules, providing a very stable anchoring point for PLP ('phosphate cup').[118] An Arg residue, Arg231 coordinates the hydroxyl group of PLP, which is unusual for ATs.

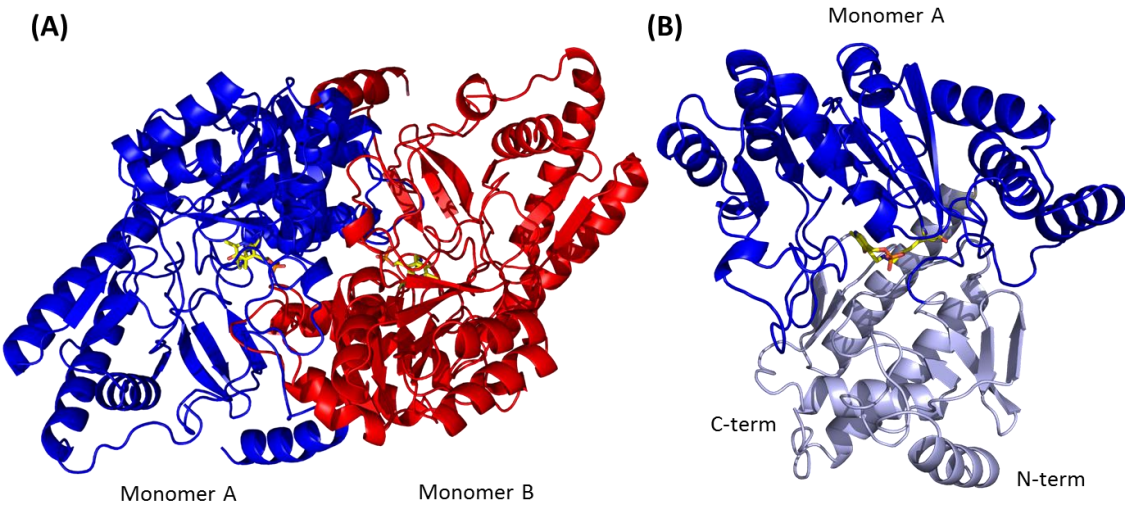


Figure 4.28 **(A)** Crystal structure of the FumI internal aldimine at 1.75 Å resolution Monomer A is shown in red, monomer B in blue. **(B)** Large and small domain of monomer A are highlighted in bright and pale blue. PLP is shown as sticks and highlighted in yellow.

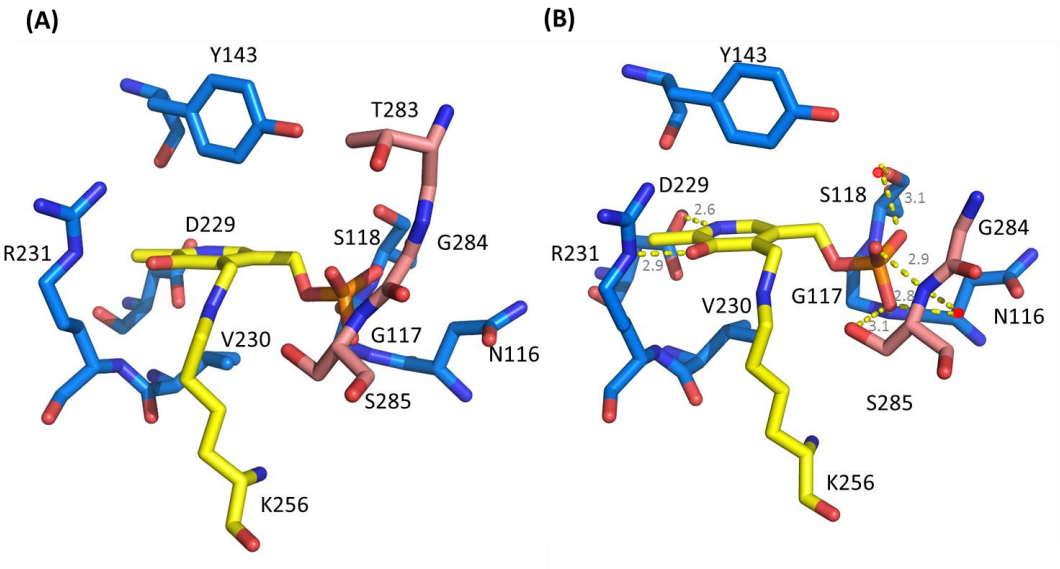


Figure 4.29 (A) Zoom into the active site of the FumI PLP internal aldimine. The PLP internal aldimine and residues involved in cofactor stabilisation are shown as stick representations, with carbon atoms shown in red and blue for the two monomers in the active dimer, oxygens are shown in red, nitrogen in blue and phosphorous in orange. The aromatic ring of PLP is sandwiched between Val230 and Tyr143. The nitrogen of the pyridine ring is fixed by hydrogen bond coordination to Asp229 as is common in AT enzymes. The phosphate group is stabilized by a network of interactions with Ser117, from Monomer A; and Gly284, and Thr283 and Ser285 from monomer B. Unusually, Arg231 interacts with the hydroxyl group of the PLP ring. **(B)** Hydrogen bonds, direct or through water molecules, are shown as dotted lines and distances (Å) in grey.

A long (~ 20 Å) hydrophobic tunnel is observed that leads from the surface down into the PLP-binding site (Figure 4.30).

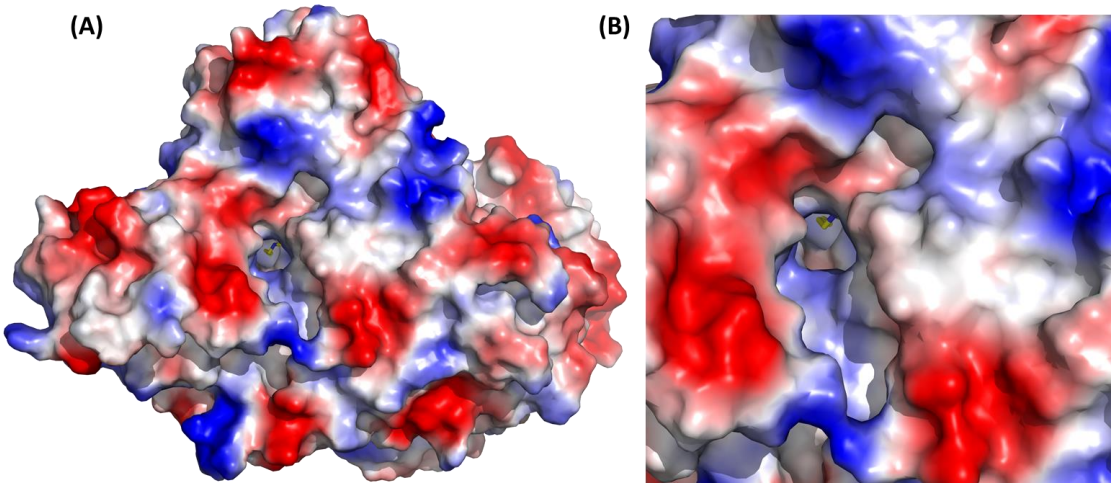


Figure 4.30 (A) FumI surface coloured by electrostatic potential. The presence of a large tunnel to accommodate the HFB_1 substrate can be observed. **(B)** Zoom in of the large hydrophobic 20 Å tunnel.

The hydrophobic nature of this tunnel suggests that it accommodates the long HFB_1 substrate. Moreover, this feature provides insight into the FumI alkyl-chain substrate promiscuity observed in biochemical assays. However, without a co-crystallisation structure of FumI with a bound substrate it is difficult to predict where different substrate functional groups will interact with the FumI structure, which makes the rational design of enzyme modifications somewhat more challenging.

The unusual Arg231 residue interacting with the hydroxyl group of the PLP cofactor is unique amongst the AT class III family. Bioinformatic analysis with the some of the best characterised members of this family such as Cv-AT and Vf-AT and the just characterised D-PhgAT (displaying 22%, 24% and 26% sequence identity respectively) highlight the unusual nature of this residue (Figure 4.31). From the sequence alignment it can be noticed that a DEVI motif is highly conserved in a loop region of the large domain of class III ATs. In FumI the DEVI motif is replaced by DDVR. The usually conserved aliphatic isoleucine residue, a key residue in the large domain, is substituted to the basic Arg231 suggesting that the arginine residue may play a role in substrate recognition and catalysis.

Site directed mutagenesis (SDM) experiments were performed whereby Arg231 was mutated to an alanine. Unfortunately, the R231A mutant was completely insoluble and no protein was recovered after protein purification. As such, no characterization could be performed. This suggests that the interaction of Arg231 with the PLP is crucial for the correct protein folding.

The obtainment of an external aldimine structure could further highlight the role of the Arg residue, known to play a crucial role in AT substrate recognition.[111,113] Thus, substrate soaking experiments have been performed with the natural substrate HFB_1 (kindly supplied from Dr. Moll (BIOMIN, Austria)), to attempt to trap the external aldimine structure. Unfortunately, very low occupancy of the substrate was observed, even after prolonged soaking of HFB_1 (up to 1 h). This is possibly due to the highly hydrophobic nature of the substrate that cannot penetrate the crystal. Alternative approaches, such as substrate docking or molecular dynamics could be performed in the future in order to identify further residues involved in substrate recognition and eventually identify hot-spots for mutations.

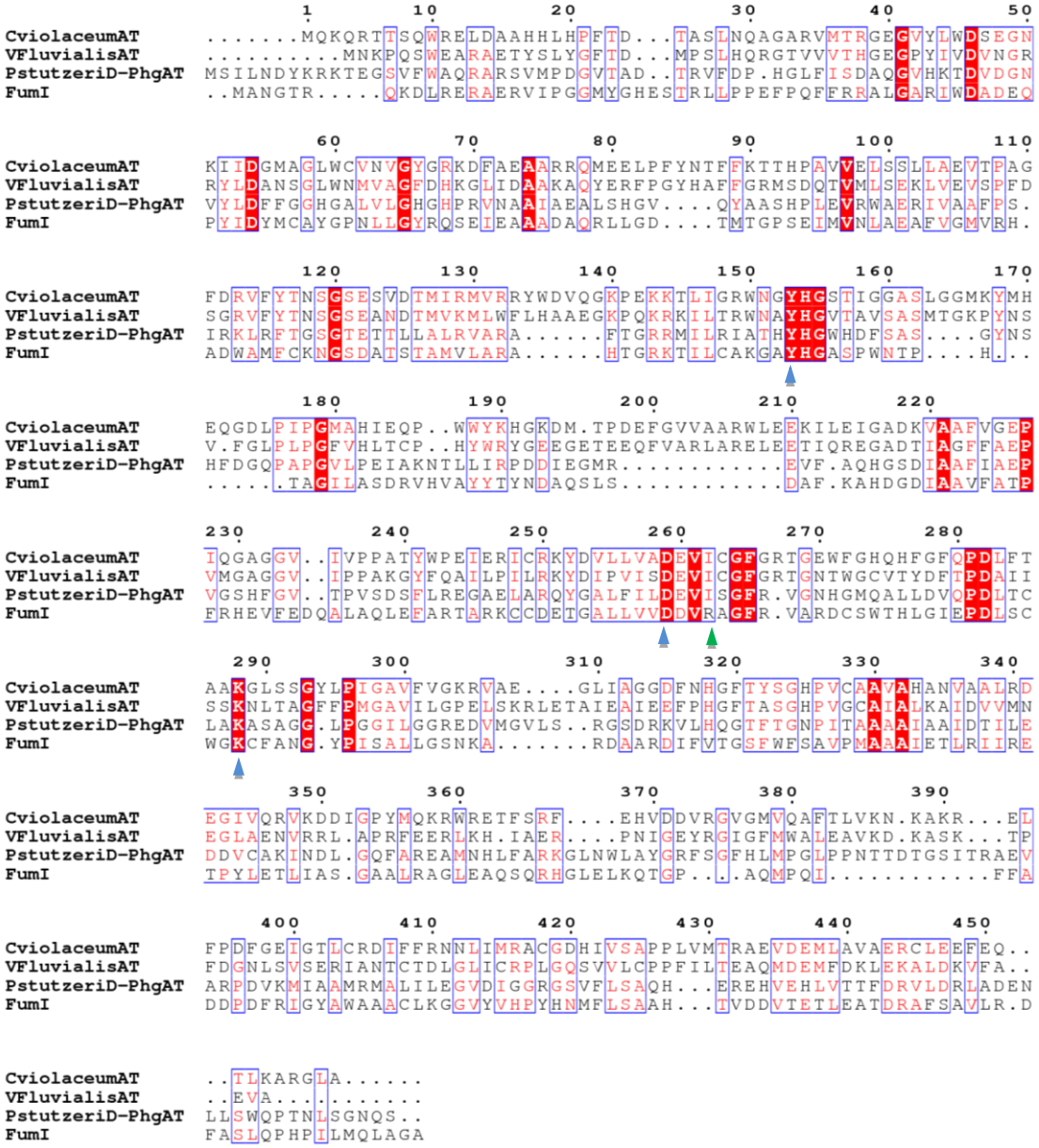


Figure 4.31 ESPript[162] alignment of *Sphingopyxis* FumI with Vf-AT (F2XBU9), Cv-AT (Q7NWG4) and *P. stutzeri* D-PhgAT. The conserved residues involved in cofactor stabilisation (Lys, Asp, Tyr) involved in substrate orientation are marked with a blue triangle and are conserved amongst the AT family. The unique FumI Arg231 residue is shown with a green triangle and is not conserved amongst the AT family. The residues numbering refers to the Cv-AT.

4.10 Conclusions and future work

Recombinant *Sphingopyxis* FumI AT has been successfully expressed, purified and characterised. The FumI substrate scope has been widely explored suggesting an enzyme promiscuity that could be exploited for industrial biocatalysis. FumI was able to transaminate long alkyl chain keto substrates (from C_3 to C_{20}) and long chain aldehydes (from C_{10} to C_{20}). Conversion of substrates with bulkier substituents such as bromo-, fluoro- as well as unsaturated keto compounds will be interesting to be explored to further characterise the enzyme scope.

A coupled spectrophotometric assay, as well as an LC-MS method, have been developed to determine kinetic parameters and observe the product formation respectively.

The enzyme enantioselectivity has been determined by Marfey's derivatization, highlighting the (*S*)-selectivity of FumI.

The possibility of using FumI in a cascade process, by coupling it with a CAR enzyme has been explored and promising preliminary results suggest that the conversion from fatty acid to fatty amine using this coupled system is achievable. Further work will be necessary in order to optimise the reaction conditions and obtain substrates conversion from fatty acid to fatty amines. Whole cells biotransformation could be explored to avoid the requirement of cofactor regeneration systems.

The 3D structure of the FumI: PLP internal aldimine has been solved at 1.6 Å resolution by molecular replacement with a *B. subtilis* AT that displayed 29% sequence identity. The structure highlighted the presence of a wide hydrophobic tunnel that provides some insight into the ability of the enzyme to accept long chain substrates. An unusual Arg231 residue has been identified as potentially important for substrate recognition and catalysis and mutagenesis (R231A) led to the production of insoluble protein. Obtaining the structure of a PLP:substrate or product external aldimine, by crystals soaks or modelling/docking, will be crucial to identify residues for future directed evolution studies which may expand the substrate scope.

Exploring the possibility of immobilising FumI, alone or in the cascade process, using commercial resins such as EziG will have to be explored to reduce the cost of the process and the requirement for using purified enzyme.

Natural product biosynthetic pathways are a source of enzymes that have become attractive candidates for biocatalysis, by using them for the generation of 'unnatural' natural products. In this study, the FB_1 degradation pathway from *Sphingopyxis sp.* MTA144, has provided an attractive biocatalyst, FumI, previously unexplored. FumI represent a very useful addition to the biocatalytic tool box as it can be applied for the biosynthesis of long chain amines from both aldehyde and ketones, providing a much more eco-friendly alternative to fatty amine synthesis.

5 Materials and methods

5.1 Materials and reagents

All reagents, chemicals and media were purchased from Sigma-Aldrich, Thermo Fisher Scientific or Bio-Rad unless otherwise stated. All primers were purchased from Sigma Genosys and competent cells, pET plasmids and restriction enzymes from New England Biolabs (NEB), Novagen and Life Technologies. Chromatography columns were purchased from GE Healthcare. The *dpga* gene was purchased from GenScript.

5.1.1 Culture media

Growth media and Luria Bertani (LB) agar plates were prepared by dissolving the components in deionised water and sterilising by autoclaving for 20 min at 120 °C. 40% Glucose and 1 M $MgCl_2$ stock solutions were prepared separately and sterilised by filtration through a 0.22 µm filter. They were then stored at 4 °C and used within one week and added to the Modified ZB media when cooled down.

- LB: tryptone (10 g L^{-1}), yeast extract (5 g L^{-1}) and NaCl (10 g L^{-1})
- 2YT broth: tryptone (16 g L^{-1}), yeast extract (10 g L^{-1}) and sodium chloride (5 g L^{-1})
- Super optimal broth with catabolite repression (SOC): tryptone (20 g L^{-1}), yeast extract (5 g L^{-1}), sodium chloride (10 mM), potassium chloride (2.5 mM), magnesium chloride (10 mM), magnesium sulphate (10 mM), glucose (2% w/v)
- M9 10X: Na_2HPO_4 (75.2 g L^{-1}), KH_2PO_4 (30 g L^{-1}), NaCl (5 g L^{-1}), NH_4Cl (10 g L^{-1})
- Modified M9ZB medium, 1 X M9 salts solution, tryptone (10 g L^{-1}), yeast extract (5 g L^{-1})

5.1.2 Antibiotic

When growing cell lines harbouring plasmid an appropriate antibiotic was added to cool growth media after sterilisation. Antibiotic solutions were prepared in deionised water, filtered through a 0.22 µm filter and stored at -20 °C.

- Ampicillin (Amp_{100}): 100 mg mL^{-1} stock solution prepared and added to media to give a final concentration of 100 µg mL^{-1} (1:1000)
- Kanamycin (Kan_{30}): 30 mg mL^{-1} stock solution prepared and added to media to give a final concentration of 30 µg mL^{-1} (1:1000)
- Gentamicin (Gen_{20}): 20 mg mL^{-1} stocks were prepared and added to the media to give a final concentration of 20 µg mL^{-1} (1:1000)

5.1.3 Induction

Isopropyl β-D-1-thiogalactopyranoside (IPTG): 0.1 M stock solution was freshly prepared on the same day as used and stored at 4 °C. It was added to a final concentration ranging from 0.05 -1 mM encompassing test expression conditions and large scale inductions.

5.1.4 Cell lines

The different cell lines used in this study for cloning processes and protein expression:

Species	Cell line	Application
E. coli	DH5α™ (Invitrogen)	Transformation – subcloning plasmid amplification etc.
E. coli	C2987 (NEB)	Transformation - PCR product-vector ligations.
E. coli	BL21 (DE3) (Novagen)	Transformation- protein expression.
E. coli	ArcticExpress (Agilent)	Transformation- FumI expression

Table 5.1 *E. coli* strains used in this study.

5.1.5 LB agar

Agar powder was dissolved in deionised water and melted by autoclaving. Appropriate antibiotic was added once cooled to below 50 °C. All plates were stored at 4 ˚C and used within 1 week of preparation.

- LB agar: tryptone (10 g L^{-1}), yeast extract (5 g L^{-1}), sodium chloride (5 g L^{-1}), agar (15 g L^{-1})

5.1.6 Buffers

All buffers were prepared by dissolving components in deionised water and adjusting to correct pH using either hydrochloric acid (HCl) or sodium hydroxide (NaOH). These were filtered through a 0.22 μm filter and degassed by sonication before chromatography.

Buffer	Components
A	0.1 M CAPS, 150 mM NaCl, 20 mM Imidazole, 50 μM PLP pH 9.5
B	0.1 M CAPS, 150 mM NaCl, 500 mM Imidazole, 50 μM PLP, pH 9.5
C	0.1 M CAPS, 150 mM NaCl, 50 μM PLP, pH 9.5
D	0.1 M CAPS, 150 mM NaCl, 75 mM Imidazole, 50 μM PLP, pH 9.5
E	20 mM KPhos, 500 mM NaCl 40 mM Imidazole, 50 μM PLP, pH 7.5
F	20 mM KPhos, 500 mM NaCl 500 mM Imidazole, 50 μM PLP, pH 7.5
G	20 mM KPhos, 500 mM NaCl, 50 μM PLP, pH 7.5
H	20 mM KPhos, 150 mM NaCl, 50 μM PLP, pH 7.5
I	50 mM Tris, 150 mM NaCl, 50 μM PLP, pH 7.5
L	100 mM Tris, pH 7.5, 50 μM PLP

Table 5.2 List of buffers used in this study.

5.2 Molecular Biology

5.2.1 DNA and plasmids

Full-length, codon optimized *dpgA* (*Pseudomonas stutzeri* ST-201, UNIPROT code: Q6VY99) was purchased from GenScript. After cutting with BspHI/BamHI, *dpgA* was sub-cloned into NcoI/BamHI digested pET-15b plasmid to give recombinant D-PhgAT with a non-cleavable N-terminal His_6-tag (pET15b-dpgA) (Appendix 8.1, Appendix 8.2).

The full-length *fumI* gene (*Sphingopyxis sp.* MTA144, UNIPROT code: D2D3B2) was kindly received from Dr Wulf-Dieter Moll from BIOMIN Research Center in Tulln, Austria through a material transfer agreement (MTA). The received clone contains the fumI gene cloned into a pET-30a vector with a non-cleavable C-terminal His_6-tag (pET-30a-AT144HIS) as described in Hartinger *et al.* (Appendix 8.3, Appendix 8.4). [143]

5.2.2 Site directed mutagenesis

5.2.2.1 Primers

The D-PhgAT Q300N, Q300A, R406A and the fumI R231A were constructed according to the overlapping primer site-directed mutagenesis (SDM) method (Table 5.3).[160]

Primer	Sequence
Dpga Q300N Fw	GTGCTGCAC**AAC**GGTACCTTCACCGG
Dpga Q300N Rev	AAGGTACC**GTT**GTGCAGCACTTTACGAT
Dpga Q300A Fw	GTGCTGCAC**GCG**GGTACCTTCACCGG
Dpga Q300A Rev	AAGGTACCCTC**CGC**GTGCAGCACTTTACGAT
Dpga R406A Fw	CATCGGTGGC**GCT**GGCAGCGTTTTTCTGAGCGCG
Dpga R406A Rev	AAAACGCTGCC**AGC**GCCACCGATGTCCACACCTT
FumI R231A Fw	ACGATGTG**GCG**GCAGGTTTCCGGGTGG
FumI R231A Rev	AACCTGC**CGC**CACATCGTCAACGACCAG

Table 5.3 List of oligonucleotides used in this study. The changed sequence is highlighted in bold.

5.2.2.2 Polymerase Chain Reaction (PCR)

The reactions were set up as shown in Table 5.4.

Component	Volume (µL)
5X Phusion Buffer	10
10 mM DNTPs	1
10 uM Rev Primer	2.5
10 uM Fw Primer	2.5
DNA (100 ng)	1-2
DMSO	1.5
H_20	29.5-30.5
DNA Polymerase (Phusion polymerase, NEB)	1

Table 5.4 Volumes and concentrations of the reagents used in the PCR reaction experiments.

The PCR was carried out under the following conditions:

PCR step	Temp (°C)	Time (s)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	T_m -5ºC	30	30

Extension	72	210-240	30
Final extension	72	600	1
Hold	4	∞	--

Table 5.5 Reaction conditions used during the SDM PCR experiments.

5 µL of the reaction of the SDM PCR reaction was loaded into the agarose gel (1 µL of 5x loading buffer) and checked for correct amplification. 5 µL of 10x CutSmart buffer (NEB) and 2 µL of DpnI enzyme (20 U) were added to the reaction and incubated at 37 ˚C overnight (ON). This reaction mixture was used to transform *E. coli* C2987 cells.

5.2.3 Agarose gel electrophoresis

1% w/v agarose gels were prepared as described below according to the number of samples and the size of equipment being used. 0.6 g or 1.5 g agarose powder was added to 60 mL or 150 mL of Tris, acetic acid and EDTA (TAE buffer, Biorad) and heated in a microwave oven until dissolved. This was cooled down and 6 µL or 15 µL of Gelred (Biotum) were added (1:10000 dilution). The gel solution was poured and allowed to set and the appropriate volume of 1x TAE buffer was poured over the gel to cover it. Samples were loaded after addition of 5x loading buffer (NEB) and gels were run at a constant voltage of 100V for around 50 min. Hyperladder I (Bioline) was used as a marker and the gels visualised using UV trans-illumination.

5.2.4 Plasmid DNA purification

Plasmid DNA was purified from cells by picking transformed colonies (DH5α) into 5 mL of LB media containing appropriate antibiotic for selection (Kan_{30} or Amp_{100}). These were grown ON at 37 °C with shaking (220 rpm) with cell pellets harvested by centrifugation at 3,500 x g for 10 min at 4 °C. DNA was isolated using the Qiaprep Spin Miniprep Kit (QIAGEN) using the associated protocols and eluted into 30 µL of deionized water.

5.2.5 Sequencing reaction

The sequence of the recombinant protein was checked by sequencing reaction. 20 µL of plasmid DNA were sent to GATC Biotech Ltd. T7 Fw and Pet-RP primers from GATC were used.

Primer	Sequence
T7	TAATACGACTCACTATAGGG
Pet-RP	CTAGTTATTGCTCAGCGG

Table 5.6 Sequence of the primers used for sequencing reactions.

5.3 Protein Expression

5.3.1 Transformation of BL21 (DE3)

Plasmid DNA (~ 100 ng) was added to an aliquot of competent *E. coli* cells (10-25 µL) and incubated on ice for 20-30 min. The cells were subsequently heat shocked at 42 °C for 30 sec and cooled on ice for 2 min. SOC (100 µL) was added to the cells and these were recovered with shaking at 37 °C for 1 h. The recovered cells were plated on LB agar + appropriate antibiotic and grown at 37 °C ON. Colonies were stored at 4 °C and used within 2 weeks.

5.3.2 Transformation of ArcticExpress

E. coli ArcticExpress (DE3) competent cells were defrosted and kept on ice at all times while aliquoting. Once thawed, the cells were gently mixed and transferred (100 µL) into a prechilled 14- mL BD Falcon polypropylene round-bottom tube. A 1:10 dilution of XL10 Gold β-mercaptoethanol (Agilent) was added (2 µL) to the cells. The content of the tube was mixed gently and incubated on ice for 10 min, swirling gently every 2 min. Plasmid DNA (~ 100 ng) was added to the aliquot and incubated on ice for 30 min. After incubation, cells were heat-shocked in a 42 °C water bath for 20 sec and cooled on ice for 2 min. Preheated (37 °C) SOC medium (900 µL) was added to the cells and subsequently incubated at 37 °C for 1 h with shaking at 230 rpm. Finally, cells were spread onto an LB agar plate containing Kan_{30} and left to grow ON at 37 °C. Colonies were stored at 4 °C and used within 2 weeks.

5.3.3 Small scale test expression in BL21 (DE3)

Test expressions were carried out on small scale to determine optimal conditions for D-PhgAT prior to large scale expression. A single colony of *E. coli* cells harbouring the pet15b-dpgA plasmid was inoculated in LB broth with Amp_{100} and grown ON at 37 °C. In the morning, the optical density at 600nm (OD_{600}) of the cells was taken and this culture added to fresh LB + antibiotic for a final OD600 of 0.1. The new culture was grown to an OD600 of approximately 0.6-0.9 at 37 °C. 0.5 M solution of IPTG was added to the cells for a final concentration of 0.1 mM, 0.5 mM and 1 mM. The cultures were incubated further at 16 °C and 30 °C for 3 h, 5 h and ON. Time points were collected at those times by centrifugation at 3,500 x g for 10min. Protein expression was analysed after cell lysis and separation of pellet and supernatant by denaturing 12% SDS-PAGE gel run at 200 V.

5.3.4 Large scale expression in BL21 (DE3)

The pet15b-dpgA construct was used to transform *E. coli* BL21 (DE3) competent cells and selection was carried out on agar plates containing Amp_{100}. A single colony was used to inoculate 250 mL of LB/Amp_{100} broth and the ON culture was grown at 37 °C with shaking at 250 rpm. The ON culture was used to inoculate 1 L of fresh LB/Amp_{100} broth and grown to an OD_{600} of 0.6–0.9. Protein expression was induced by addition of IPTG to a final concentration of 0.1 mM and growth was continued for 16 h h at 20 °C. Cells were harvested by centrifugation (Thermo Scientific Multifuge X3R) at 4,000 x g for 30 min at 4 °C.

5.3.5 Large scale expression in ArcticExpress

Protein expression was carried out as previously described by Hartinger *et al*.[143] One colony of ArcticExpress (DE3) harbouring pET-30a-AT144HIS derived plasmid was used to inoculate LB broth (250 mL) containing Gen_{20} and the appropriate antibiotic for the pET vector (Kan_{30}) and grown ON at 37 °C with shaking at 250 rpm. The next morning, the inoculant was diluted 50-fold into M9ZB modified medium (10 x 1000 mL) containing M9 salts, 1mM $MgCl_2$ and 0.4% glucose without antibiotics and incubated for 3 h at 37 °C with shaking at 250 rpm (Excella E25 Incubator, New Brunswick Scientific). Cells were allowed to grow to an

OD_{600} of 0.8 before the temperature was lowered to 11 °C. After the temperature was equilibrated, protein expression was induced by the addition of 1 mM IPTG. Protein expression was carried out for 24 h at 11 °C with shaking at 250 rpm.

5.3.6 Cells harvesting

After the appropriate time of induction, bacterial cells were harvested by centrifugation (Thermo Scientific Multifuge X3R) at 4,000 x g for 30 min at 4 °C. The supernatant was discarded and the cell pellet was resuspended in a minimal amount of PBS buffer and further centrifuged for 30 min at 3,500 x g at 4 °C. Finally, the supernatant was discarded and the bacterial cell pellet containing the protein of interest was stored at -20 °C for further protein purification.

5.3.7 Fermentation of ArcticExpress

Fermentation of FumI was carried out by Dr. John White at the University of Edinburgh on a 12 L scale in a BioFlo 4500 Fermentor. The starting colture was carried out as previously described and diluted by 50-fold in the 12 L fermentor. Cells were allowed to grow to an OD_{600} of 1.0 at 37 °C before lowering the temperature down to 11 °C. Protein expression was induced by the addition of 1 mM IPTG for 24 h at 11 °C, harvested and stored at -20 °C.

5.4 Protein Purification

5.4.1 Cell Lysis by sonication

All purification steps were carried out at 4 °C. Cells were resuspended in buffer A/E with addition of DNase (0.2 mg per 10 mL buffer) and lysed by sonication for 15 cycles (30 sec on, 30 sec off). The lysed cell suspension was clarified by centrifugation (Thermo Scientific Multifuge X3R) at 14,000 x g for 40 min at 4 °C. The cell-free extract was loaded onto a 1 mL HisTrap nickel affinity column (GE Healthcare). Samples (50 µL) from all stages were retained for SDS-PAGE analysis.

5.4.2 Cell lysis by high-pressure homogenization

Cell pellets were resuspended in buffer E and particulates were removed *via* filtration through 1 mm sieve. A cell disruptor (Constant Systems) was equilibrated with buffer E and then the resuspended pellet was passed through at a pressure of 20.1 kpsi. The lysed cell suspension was cleared by centrifugation (Thermo Scientific Multifuge X3R) at 14,000 x g for 40 min at 4 °C.

5.4.3 Nickel NTA purification

The cell-free extract was loaded onto a 1 mL HisTrap nickel affinity column (GE Healthcare). The column was washed with binding buffer (buffer A/E) for 20 column volumes (CVs) , then the protein eluted with an imidazole gradient (20 to 500 mM)(buffer B/F) over 30 CVs. Protein containing fractions were analysed by 12% SDS-PAGE and concentrated to 1 mL using a Vivaspin 20 MWCO 30,000 (Sartorius).

5.4.4 Size exclusion chromatography

The concentrated D-PhgAT/FumI (1 mL) was loaded onto a pre-equilibrated (Buffer C/G) HiPrep™ 16/600 Superdex™ S-200 size exclusion column (120 mL). Recombinant protein was eluted at a flow rate of 1 mL min^{-1} monitoring at 280 nm and 420 nm. The purity of the recombinant proteins was analysed by 12% SDS-PAGE. The fractions containing clean protein of interest were combined and frozen at -80 °C with 20% glycerol.

5.5 Protein characterisation

5.5.1 SDS-PAGE protein analysis

An average gel (16 mL) consisted of a 12% running gel (6.9 mL H_20, 4.8 mL of 40% acrylamide, 4 mL of 1.5 M Tris pH 8.8, 150 µL of 10% w/v SDS, 300 µL of 50 mg mL^{-1} ammonium peroxodisulfate and 20 µL TEMED) and a 4% stacking gel (2.9 mL H_20, 0.75 mL of 40% acrylamide, 1.25 mL of 0.5 M Tris pH 6.8, 50 µL of 10% w/v SDS, 100 µL of 50 mg mL^{-1} ammonium peroxodisulfate and 5 µL TEMED). Polyacrylamide gel electrophoresis (PAGE)

of purification stage samples was performed using the BioRad Mini Protean II system and 12% acrylamide gels. 50 µL of each samples were boiled for 5-10 min in 2x SDS loading buffer (2.5 mL Tris, 0.5 M pH 6.8, 2 mL glycerol, 4Ml of 10% w/v SDS, 1 mL β-mercaptoethanol and 0.5 Ml of 0.1% w/v bromophenol blue) with 10 µL loaded onto the gel and run for ~ 50 min at 200V/180mA. Protein was visualised by InstantBlue Coomassie Stain (Expedeon Ltd) and mass estimated using LMW marker (GE Healthcare).

5.5.2 Determination of enzyme concentration

5.5.2.1 Bradford assay

The Bradford protein assay is a method used to quantify the protein concentration in a sample. It is based on the principle that the Coomassie Brilliant Blue G-250 dye (Figure 5.1) changes colour from brown to blue under acidic conditions, measuring in this way the presence of basic amino acids (lysine, arginine and histidine).[172]

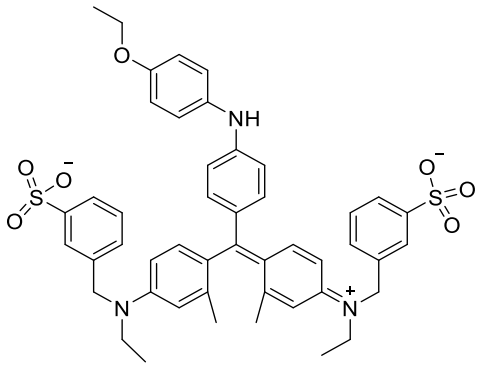


Figure 5.1 Coomassie Brilliant Blue G-250 dye structure

As first,50 µL of four serum albumin bovine (BSA) standard solutions (0.25 mg mL^{-1}, 0.5 mg mL^{-1}, 1 mg mL^{-1}, 1.4 mg mL^{-1}) were mixed with 1500 mL of Bradford reagent (Sigma). The absorbance was read at 595 nm and used to build a calibration curve. 50 µL of the purified protein at different dilutions (1:2, 1:5) were then mixed with 1.5 mL of Bradford reagent and the absorbance at 595 nm was recorded.

5.5.2.2 Beer-Lambert Law

Relevant size exclusion fractions were pooled together and scanned between 200-800 nm wavelengths using a Varian Cary-50 UV-Vis spectrophotometer. Extinction coefficients were

calculated using ProtParam tool (https://web.expasy.org/protparam/) and subsequent protein concentrations determined using the scanned A_{280} absorbance and the Beer-Lambert equation $A = \varepsilon cl$.[192]

5.5.3 UV/Vis spectroscopy

All UV-Vis spectra were recorded on a single-beam Hewlett-Packard 8452A diode array spectrophotometer and analysed using UV-Vis ChemStation software (Agilent). To convert the *apo*-D-PhgAT/FumI to *holo*-D-PhgAT/FumI the enzyme was dialyzed for 2 h at 4 °C against buffer C/H with 50 μM PLP. Excess PLP was removed by passing the protein through a PD-10 (Sephadex G- 25M) desalting column (GE Healthcare) before any spectrophotometric measurements were taken. The concentration of recombinant protein was 20 μM and the spectrophotometer was blanked with buffer C/H without 50 μM PLP.

5.5.4 Measuring enzyme dissociation constants for amino acids

To convert the *apo*- FumI to *holo*- FumI the enzyme was dialyzed for 2 h at 4 °C against buffer H with 50 μM PLP and excess of PLP was removed by buffer exchange as previously described. 10 X stock solutions were prepared for each tested amino donor at each concentration (0-200 mM). The spectrum was recorded immediately prior to the addition and immediately following the addition of the amino donor and left at 35 °C for 18 h. The morning after the spectrum was recorded and the Δ_{325nm} used to calculate the dissociation constant.

Data was analysed by non-linear regression using Origin. Changes in absorbance at 325 nm were plotted against L-alanine (L-Ala) concentrations, and data points were fitted to a hyperbolic saturation curve using Sigma Plot software:

$$\Delta A_{obs} = \frac{\Delta A_{max}\,[Ala]}{K_d + [Ala]}$$

The ΔA_{obs} is the observed change in absorbance at 325 nm and the ΔA_{max} is the maximal absorbance change, [Ala] is the L-Ala concentration, and K_d is the dissociation constant.

5.6 Amino benzenes as 'smart' amino donors

5.6.1 D-PhgAT

The amines 5-nitro-2, 3-dihydro-1H-inden-2-amine (2,3-DHIA), 4-(2-Aminoethyl) benzonitrile
(4-AEB) and 2-(4-nitrophenyl) ethan-1-amine(4-NPEA) were dissolved in buffer D and tested with different amino acceptors (benzoylformate (BZF), α-ketoglutarate (AKG), benzaldehyde (BZALD) and N-butyraldehyde (NBA)) at 0.5 mg mL^{-1} D-PhgAT. Amine donor were used at a final concentrations of 25 mM while amino acceptors at 10 mM, as suggested in the literature[116]. For each reaction a control with no enzyme was performed. Reactions were incubated for 18 h at 37 °C.

5.6.2 FumI

The amine donors 2,3-DHIA, 4-AEB and 4-NPEA were tested with different amino acceptors (from C_3 to C_{20} and BZALD and NBA) at 0.5 mg mL^{-1} FumI. Amine donor were used at a final concentrations of 12.5 mM while amino acceptors at 5mM/1mM final concentration according to their solubility. The keto compounds were dissolved in 100% DMSO and prepared as 10x stock solutions to achieve a final concentration of 10% DMSO. For each reaction a control with no enzyme was performed. Reactions were incubated for 18 h at 37 °C.[153]

5.7 *Ortho*-xylylene diamine (OXD) assay

5.7.1 UV/Vis studies

A solution of 10 mM of *ortho*-xylylene diamine (OXD) was made up in the appropriate buffer (C/H) and added to the enzyme (20 µM), appropriately loaded with PLP and desalted, for a final concentration of 1 mM. The spectrum was recorded from 200 to 800 nm as described above. Overtime the formation of a purple precipitate could be observed.

5.7.2 Solid phase assay

The solid phase assay was performed as described by Anthony P. Green *et al.*[152,102] A glycerol stock of *E. coli* BL21 (DE3) harbouring pET-15b-dpgA was diluted 1/10,000 in sterile H_2O and plated onto a Hybond-N membrane on the surface of LB-agar supplemented with Amp_{100}. Plates were incubated at 30 °C ON. As a control, untransformed *E. coli* BL21 (DE3) cells were grown in parallel. The recombinant protein expression was induced by the addition of 2 mM IPTG to the surface of the LB-agar plates, replacing the membrane and incubating for further 6 h at 30 °C. Blotting paper was then soaked in a solution of OXD (5 mM) in KPhos buffer (100 mM, pH 7.5) and the excess liquid was drained. The membranes were placed on top of the blotting paper and incubated at 30 °C for ON. After 1 h the formation of a purple precipitate could be observed.

5.8 Spectrophotometric assays

5.8.1 Measuring enzyme kinetics with the coupled assay AKGDH/D-PhgAT

The D-PhgAT activity was monitored by coupling the reaction with the α-ketoglutarate dehydrogenase (AKGDH), being AKG the product of the first half reaction. The final volume of the reactions was 250 µL and contained 0.1 M CAPS pH 9.5, 150 mM NaCl, 1mM $MgCl_2$, 1 mM CaCl2, 0.05 mM EDTA, 50 uM PLP, 1 mM CoASH, 3 mM NAD^+ , 0.2 U KGDH, 0 to 50 mM L-Glu, 0 to 20 mM BZF. These reagents were pre-incubated at 35 °C and the reaction was blanked. The reaction was initiated by adding 1 µM D-PhgAT. The increase in absorbance at 340 nm resulting from the enzymatic conversion of NAD^+ to NADH was monitored over 1h on a BioTek Synergy HT plate reader with Costar 96-well plate. The data from the first 15 min were analysed using the Michaelis–Menten model and a nonlinear regression fit on GraphPad gave values of K_M and k_{cat} (ε_{NADH} = 6220 M^{-1} cm^{-1}).

5.8.2 Measuring enzyme kinetics with the coupled assay FumI/PDH

The FumI activity was monitored by coupling the reaction with the pyruvate dehydrogenase (PDH), being pyruvate the product of the first half reaction. The final volume of the reactions was 200 µl and contained 50 mM Tris, 700 mM $MgCl_2$, 35 mM $CaCl_2$, 50 µM, 0.32

mM thiamine pyrophosphate (TPP), 200 µM CoASH, 2 mM NAD 0.2 U PDH, 0 to 100 µM keto substrate, 10 mM L-Ala. These reagents were pre-incubated at 35 °C and the reaction was blanked. The reaction was initiated by adding 1 µM FumI. The increase in absorbance at 340 nm resulting from the enzymatic conversion of NAD^+ to NADH after pyruvate production by the FumI was monitored over 1 h on a BioTek Synergy HT plate reader with a Costar 96-well plate. The data from the first 15 min were analysed using the Michaelis–Menten model and a nonlinear regression fit on GraphPad gave values of K_M and k_{cat} (ε_{NADH} = 6220 M^{-1} cm^{-1}).

5.9 Chiral HPLC

For kinetic parameters calculations, reactions containing 0.51 mg mL^{-1} D-PhgAT, , 0-200 mM L-Glu, 0-20 mM BZF/HBZF/IPA/PPA/HPA in 0.1 M CAPS pH 9.5, 150 mM NaCl, 50 µM PLP buffer were incubated at 37 °C before being terminated at 15 min by diluting 40-fold in the chiral mobile phase: 0.025% triethylammonium acetate (TEAA):MeOH (50:50). Reactions were then analysed by chiral HPLC using a Chirobiotic T column (Astec, chiral phase Teicoplanin, 5 µm, 250 mm × 4.6 mm) under the following isocratic conditions: mobile phase: 0.025% TEAA:MeOH (50:50, v/v), flow rate: 1 mL min^{-1} , detection λ: 205 nm, temperature: RT, run time: 30 min. The column was calibrated with single enantiomers of the amino acids. The conversion was used to measure enzyme activity (Appendix 8.6). The same procedure was applied for the non-natural products such as D-phenylalanine, D-tyrosine and D-tryptophan. Lab solution sofwtare (Shimadzu) was used to measure the area under product peaks.

5.9.1 Determination of the % enantiomeric excess (*ee*)

Enantiomeric excess (*ee*) was calculated for each biotransformation.

The% *ee* was calculated as follows:

$$e.e. = \frac{|(R) - (S)|}{(R) + (S)} \; x\, 100\%$$

Where (*R*)- and (*S)*- are the AUC for each enantiomer.

5.9.2 Amino donor screening D-PhgAT

A D-PhgAT aliquot was thawed and incubated on ice with 50 µM PLP for ~ 2h in order to be fully loaded with fresh PLP. Reactions containing 0.51 mg mL^{-1} D-PhgAT, 10 mM L/D amino donor, 10 mM BZF in buffer C were incubated at 37 °C before being terminated at 18 h by diluting 40-fold in the chiral mobile phase. Reactions were then analysed by cHPLC using a Chirobiotic T column as stated before. Percentage of conversions were normalized by L-Glu conversion which resulted to be the best amino donor.

5.9.3 Large scale biotransformation conditions D-PhgAT

Large scale biotransformation (1g) was performed in 100 mL conical flask with a final volume of 50 mL. For each component of the reaction the pH was adjusted to 9.5 before being added to the reaction mixture. A fresh purified D-PhgAT was added at a final concentration of 0.2 mg mL^{-1}, 0.4 mg mL^{-1}or 1 mg mL^{-1} (4 µM, 8 µM, 20 µM) in buffer C containing 1 g of BZF (~ 133 mM) and 500 mM L- Glu. The biotransformation was carried out at 37 °C at 120 rpm. A negative control with all the reagents but the enzyme was performed at the same time. Time points were collected at time 0, 15 min, 30 min, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, 8 h, 9 h, 24 h, 36 h. Each time point was collected in triplicated and processed as described above. At each time point reactions were quenched by diluting 40-fold in the chiral mobile phase and analysed by cHPLC.

5.9.4 Small scale biotransformation D-PhgAT

Small scale biotransformation was performed in 50 mL conical Falcon tube with a final volume of 5 mL. For each component of the reaction the pH was adjusted to 9.5 before being added to the reaction mixture. A fresh purified D-PhgAT was added at a final concentration of mg mL^{-} in buffer C containing 100 mg of BZF, HBF, IPA, PPA, HPPA and 500 mM L- Glu. The biotransformation was carried out at 37 °C at 120 rpm. A negative control with all the reagents but the enzyme was performed at the same time. Time points were collected at time 0, 15 min, 30 min, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, 8 h, 9 h, 24 h. Each time point was collected in triplicated and processed as described above. At each time point

reactions were terminated by diluting 40-fold in the chiral mobile phase and analysed by cHPLC.

5.9.5 FumI reactions to be analysed by MS

A FumI aliquot was thawed and incubated on ice with 50 µM PLP for ~ 2h in order to be fully loaded with fresh PLP. Reactions containing 0.4 mg mL^{-1} FumI, 10 mM L-Ala, 100 µM keto acceptor (dissolved in DMSO for a final 10% concentration) in buffer H were incubated at 35 °C at 250 rpm using a Thermo Shaker (PCTM) for 18 h. Reactions were terminated by diluting 1:1 with ACN, 0.1% v/v formic acid. Samples were clarified by centrifugation at 16,000 x g for 10 min and analysed by mass spectrometry.

5.9.6 FumI-CAR cascade

The FumI-CAR cascade was set up with C10-FA, C12-FA, C-14-FA, C16-FA, C18-FA and C20-FA in buffer L as follows:

Reagent	Final concentration (mM)
1M $MgCl_2$	20
100 mM ATP	2
100 mM NADPH	2
50 mM FA	1
500 mM Alanine	10
CAR	0.25 mg mL^{-1}
FumI	0.24 mg mL^{-1}

Table 5.7 FumI-CAR couple reaction conditions without the recycling system.

Each fatty acid was dissolved in 100% DMSO. When using the NADH/ATP recycling system the same conditions were applied with the addition of 10 mM Glucose, 4 mg mL^{-1} Poly (P), 0.3 mg mL^{-1} GDH, AdK and PAP.

5.10 Marfey's (FDAA) reagent derivatization

50 µL of 1 mM (*R*)-/(*S*)-C8-2A were mixed with 100 µL of FDAA (dissolved in Acetone) and 20 µL of 1 M $NaHCO_3$. The sample was heated at 40 °C for 1 h in a Thermo Shaker (PCMT). After the sample cooled down to room temperature (RT), 20 µL of 2M HCl were added.

The sample (10 µL) was injected onto Luna 5u C18 RP-HPLC column (100 Å, 250 x 4.60 mm, Phenomenex). The LC gradient ran from 10% ACN and 80% water with 0.1% TFA to 70% ACN 0.1% TFA over 40 min, flow rate 1.0 mL min^{-1}, temperature: RT, detection at 340 nm.
A reaction containing 0.4 mg mL^{-1} FumI, 10 mM L-Ala, 100 µM C8-2K (dissolved in DMSO for a final 10% concentration) in buffer H was incubated at 35 °C at 250 rpm using a Thermo Shaker for 18 h. The reaction was derivatized with Marfey's reagent following the protocol described above.

5.11 Mass Spectrometry

5.11.1 Protein LC-MS

Bradford assay was used to determine protein concentration for LC-MS. Protein was diluted to a concentration of 20 µM and centrifuged for 10 min at 17,000 x g. 5 µL of the supernatant was injected into LC-MS on a Synapt G2-Si Q-TOF (Waters) instrument with Phenomenex C4 3.6 µm LC column coupled to an electrospray ionisation (ESI) source. The LC gradient ran from 5% ACN and 95% water with 0.1% FA to 95% ACN over 15 min. The MS source was set at 120 °C, backing pressure 2 mbar and sampling cone voltage of 54 V. Protein spectra are presented after subtraction using MassLynx V4.1 software. Peak *m/z* annotations were extracted from smoothed and centroided data which was also used in the component algorithm to extract the protein average masses.

5.11.2 FumI reactions LC-MS

FumI reactions (50 µL) with different keto substrates (C_3 to C_{20}) and 10 mM alanine were mixed with ACN, 0.1% v/v formic acid (50 µL) and clarified by centrifugation at 16,000 x g for 10 min. 5 µL of supernatant was subjected to LC-MS on a Synapt G2-Si Q-TOF (Waters) instrument with Phenomenex Jupiter C18 5 µm 300 Å LC column coupled to an ESI source. The LC gradient ran from 5% ACN and 95% water with 0.1% FA to 95% ACN over 30 min. The MS source was set at 120 °C, backing pressure 2 mbar and sampling cone voltage of 54 V. Extracted ion chromatograms (EICs) and masses were determined on MassLynx V4.1 software. A lockmass correction was applied once at the beginning of each analysis against leucine enkephalin peptide.

5.11.3 FumI reactions MALDI

For MALDI sample preparation a volume of 1 µL of FumI reaction sample solution was spotted on a ground steel MALDI target plate, followed by 1 µL of α-cyano-4-hydroxycinnamic acid (CHCA) matrix solution (1% in 50% ACN/0.1% TFA (w/v/v)). The sample and the matrix solutions were mixed together on the plate and allowed to air-dry. Each sample was spotted on to the MALDI target plate in triplicate. For analysis using the ultraflex the data was acquired using FlexControl software version 3.3 (Bruker Daltonics). Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in reflectron mode, equipped with a Nd:YAG smart beam laser. Spectra were acquired over a range of 50-500 *m/z*. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929). Monoisotopic masses were obtained using a SNAP averaging algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a S/N threshold of 2. . Fragmentation was performed in LIFT mode without the introduction of a collision gas. The default calibration was used for MS/MS spectra, which were baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 *m/z*, cycles 4); monoisotopic peak detection used a SNAP averaging algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N of 6. Bruker flexAnalysis software (version 3.3) was used to perform spectral processing and peak list generation.

Before starting the run the laser intensity was first optimised so that maximum signal intensity was obtained, without compromising the baseline of the spectrum. The average laser intensity used was around 43%, however this changed depending on the sample type, with extracts of low concentration typically requiring a higher laser intensity than concentrated extracts. The optimum number of laser shots was investigated at the beginning of each sample run with an average of 800 shots used. It was found that when analysing majority of the samples, the spectra obtained did not increase above 800 laser shots. In addition, before each run, the plate was aligned so that the laser started in the centre of each of each spot with the spot movement for each sample set to random.

5.12 N_{15} D-Phg and N_{15} D-Hpg biosynthesis

5.12.1 C18 HPLC method

Reactions were then analyzed by HPLC using a Luna 5µ C18 (2) RP-HPLC column (100Å, 250×4.60mm, Phenomenex) and the following conditions: 5% to 95% ACN (0.1% TFA, v/v)/water (0.1% TFA, v/v) in 20 min, flow rate: 1mL min^{-1}, detection at λ= 274 nm, temperature: RT. Samples were diluted by 40-fold into 5% ACN (0.1% TFA, v/v) before being analysed by HPLC. A calibration curve of D-Phg and D-Hpg was used to relate the area under the curve (AUC) with the amount of product synthesised.

5.12.2 Synthesis of N_{15} D-Phg and N_{15} D-Hpg

For the preparation of N_{15} D-Phg the reaction mixture consisted of: 66 mM BZF, 340 mM N_{15}-L-Glu and 1 mg mL^{-1} D-PhgAT in a 5 mL scale. Same procedure was followed for the synthesis of ^{15}N-D-Hpg but using a lower concentration of the amino acceptor HBF (33 mM). Reactions were incubated at 37 °C at 250 rpm. The real time conversion of BZF and HBZF into D-Phg and D-Hpg was monitored by C18 HPLC as described above.

5.12.3 Purification by reversed-phase HPLC

Reaction mixtures were acidified by adding few drops of pure trifluoroacetic acid (TFA). Precipitate generated at this stage was removed by centrifugation. Analytical HPLC anaylsis were conducted as described in 5.12.1. Preparative reverse phase HPLC was conducted on a Waters® 600 (225 µL) system using a 486 tuneable absorbance detector recording at 205 nm equipped with a Phenomenex® Luna C18(2) column (5 µm, 250 × 21.2 mm) using 97% water (0.1% TFA) and 3% ACN (0.1% TFA) as eluent at a flow rate of 18.0 mL min^{-1}. Column was pre-equilibrated (3% ACN, 0.1% TFA) before injecting reaction mixtures (2mL in every run) and washed with 97% ACN (0.1% TFA) for approximately 10min after the desired product peaks were collected. Elute fractions were collected and freeze-dried. Due to higher hydrophilicity of D-Hpg than D-Phg, D-Hpg was purified via a modified method with lower ratio of ACN in the mobile phase, using 98% water (0.1% TFA) and 2% ACN (0.1%

TFA). Elute fractions were combined and freeze-dried which afforded the product as yellowish oil.

5.12.4 NMR analysis

Nuclear magnetic resonance (NMR) spectra were recorded at 298 K on Bruker PRO500 or AVA500 spectrometers running at 500 MHz (^{1}H spectra). The solvent used was deuterium oxide (D_2O). Coupling constants J, were calculated using MestreNova (version 9). The following abbreviations (and their combinations) are used to label the multiplicities: s (singlet), d (doublet), m (multiplet).

D-Phg ^{1}H NMR (500 MHz, D_2O): δ_H 7.51-7.47 (5H, m).[166]

D-Hpg ^{1}H NMR (500 MHz, D_2O): δ_H 6.88 (2H, d, J = 8.5 Hz), 7.25 (2H, d, J = 8.5 Hz).[193]

L-Glu ^{1}H NMR (500 MHz, D_2O): δ_H 1.95-2.11 (2H, m), 2.31-2.28 (2H, m), 3.71-3.68 (1H, m).[167]

5.12.5 Mass Spec Analysis

Samples were cleaned prior to analysis using C18 solid-phase tips (Millipore ZipTips). After loading, the tips were washed with 0.1% formic acid and eluted in 50:50(v/v) ACN: 0.1% formic acid). Samples were analysed on the SolariX FT-ICR-MS using direct infusion by nanomate in positive ion mode for N_{15} D-Phg and in negative mode of N_{15} D-Hpg. Spectra were acquired using the solariXcontrol software and processed with DataAnalysis version 4.2 (Bruker Daltonics). The following setting were applied. D-Phg and ^{15}N D-Phg was isolated using the quadrapole set at either 152 (with a 7 Da window) or 153 (with a 5 Da window). For each analysis 100 scans were collected from *m/z* 100-1000 with an accumulation time of 2 sec. This was carried out by Dr Joanna Simpson and Dr Faye Cruikshank (University of Edinburgh).

5.13 X-ray Crystallography

5.13.1 D-PhgAT robot screening and optimization

This work was carried out in collaboration with Dr. Jon Marles-Wright (Newcastle, University). Recombinant purified *P. stutzeri* D-PhgAT at 9.1 mg mL^{-1} was crystallised by

hanging-drop vapour diffusion at 18 °C. Crystallisation experiments were set up in 96 well MRC plates with commercially available screens from Molecular Dimensions Limited with 100 nL drops of protein and 100 nL well solution and equilibrated against 70 µL of well solution. Crystals grew in few conditions (C3/H10 PEG/Ion, B8 JCSG$^+$, H4 MIDAS) that were further optimised. Diffracting quality crystals were obtained by adding 1 µL of 9.1 mg mL^{-1} D-PhgAT in buffer C to 1 µL of 0.1 M Tris (pH 7.5), 0.2 M MgCl2, 10% (v/v) polyethylene glycol 8000. Crystals were cryoprotected with a solution containing 0.1 M Tris (pH 7.5), 0.2 M MgCl2, 10% (w/v) polypropylene glycol 8000, 20% (v/v) polypropylene glycol 200 PEG and then flash-cooled by immersion in liquid nitrogen. Datasets were collected on beamline I03 at the Diamond Light Source (Didcot, UK) at 100 K using a Pilatus 6M detector. Diffraction data were integrated and scaled using XDS[194] and symmetry-related reflections were merged with Aimless.[195] The resolution cut off used for structure determination and refinement was determined based on the CC1/2 criterion proposed by Karplus and Diederichs.[196] The structure of D-PhgAT was determined by molecular replacement using the D-PhgAT incomplete structure with the PDB code 2CY8. A single solution comprising three dimers in the asymmetric unit was found using Phaser.[197] The initial model was rebuilt using Phenix.autobuild[198] followed by cycles of refinement with Phenix.refine[199] and manual rebuilding in Coot.[200] The final model was refined with automatically determined TLS groups and isotropic B-factors. The model was validated using MolProbity.[201] Structural superimpositions were calculated using Coot. Crystallographic figures were generated with PyMOL (Schrodinger LLC). Data collection and refinement statistics are shown in Appendix 8.7. X-ray diffraction images are available online at Zendo (doi:10.5281/zenodo.1059413). The structure is deposited at the PDB with PDB code 6G1F.

5.13.2 FumI robot screening and optimization

This work was carried out in collaboration with Dr. Jon Marles-Wright (Newcastle, University). Crystal screening with recombinant purified *Sphingopyxis* FumI (8.1 mg mL^{-1}) in 20 mM KPhos, 500 mM NaCl,50 µM PLP, pH 7.5 was performed using a wide range of commercial screens (MIDAS, PEG/Ion, STRUCTURE, JCSG$^+$) in 96 well plate using the Gryphon robot. The C10 MIDAS condition (35% w/v polyacrylate 2100 sodium salt, 0.2 M ammonium sulfate, HEPES pH 7.5) was further optimised. The best crystals were produced

by adding 1 µL of 8.1 mg mL^{-1} FumI in buffer G to 1 µL of 100 mM HEPES pH 6, 30% Polyacrilate 2100 sodium salt (v/v), 0.2 M ammonium sulfate. Crystals were cryoprotected with a solution containing 100 mM HEPES pH 6, 30%(w/v) Polyacrilate sodium salt polypropylene glycol 200 PEG and then flash-cooled by immersion in liquid nitrogen. The FumI 3D structure was solved by molecular replacement using the structure of the glutamate 1-semialdehyde AT from *Bacillus subtilis* (PDB code: 3BS8), displaying 29% sequence identity. A single solution comprising one dimer in the asymmetric unit was found using Phaser.[197] Crystallographic figures were generated with PyMOL (Schrodinger LLC). Data collection and refinement statistics are shown in Appendix 8.7. The structure is deposited at the PDB with the PDB code: 6HBS.

5.14 Enzyme immobilization

5.14.1 Optimization

EziG resins were kindly sent by EnginZyme and tested for D-PhgAT immobilization. Three different EziG carriers are commercially available with different properties with regards to surface, pore diameter and bulk density: $EziG^1$ Opal, $EziG^2$Coral and $EziG^3$ Amber.

All resins were tested in order to identify the most suitable carrier for D-PhgAT immobilization. Aiming for 10% loading (9 mg of EziG *per* 1 mg of purified D-PhgAT), as suggested by the manufacturer, the three EziG carriers were tested at different imidazole concentrations, a crucial factor for successful immobilization. 1 mg of D-PhgAT was mixed with 9 mg of carrier and left under orbital shaking for different times (30 and 60 min) at 37 °C. A volume of 200 µL was kept constant in buffer C (0-75 mM imidazole). The amount of immobilized enzyme was evaluated by Bradford assay on the supernatant. This experiment allowed the identification of the best immobilization conditions and the best carrier which was found to be $EziG^2$ with 75 mM imidazole and 30 min. All subsequent studies were conducted under the aforementioned conditions.

5.14.2 Immobilization reactions

1 mg of purified/cell free extract D-PhgAT was mixed with 9 mg of $EziG^2$ in buffer D, left under orbital shaking for 30 min at 37 °C. Reactions were left standing for 10 min to allow

sedimentation. A Bradford assay on the supernatant was carried out to check that the immobilization was successful. The supernatant was then removed and reaction components (10 mM BZF, 300 mM L-Glu) were added to the carrier for a final volume of 200 µL. These reactions were incubated at 37 °C under orbital shaking for 1 h. An aliquot of 20 µL was withdrawn from the reaction, diluted 1:40 in the cHPLC mobile phase and analysed by cHPLC. The amount of D-Phg produced was determined by relating the area under the curve with the amount of D-Phg through a calibration curve. A positive control with the wild enzyme (not immobilized) was always done in parallel in order to compare the enzyme performances.

5.14.3 Optimum of pH and Temperature

Different pH and temperature were tested to identify differences between the immobilized and wild enzyme. The immobilization was carried out under the optimized conditions (pH 9.5, 37 °C). After successfully immobilization the reactions (10 mM BZF, 300 mM L-Glu) were tested at 15 °C, 37 °C, 50 °C and 60 °C in buffer D. Analogously, reactions were carried out at pH 5, 7.5, 9.5 and 11 at 37 °C. The amount of D-Phg produced was evaluated by cHPLC.

5.14.4 Recycling cycles

To test the possibility of using multiple times the immobilized D-PhgAT, recycling tests were performed. Purified/cell free extract D-PhgAT was firstly immobilized according to optimised conditions and 1 h reactions were carried out as previously described. After withdrawing a sample for cHPLC analysis, the supernatant was removed and replaced with 1 mL of buffer D. The samples were orbitally shaken for 30 min at 37 °C. After sedimentation, a Bradford assay was performed to check eventual protein loss and stored at 4 °C. After 24 h the same procedure was repeated. This procedure was repeated for 9 times and the percentage of conversion evaluated by cHPLC.

6 Conclusions and future work

The aim of this thesis was to explore the biocatalytic potential of two very interesting, distantly-related class III aminotransferases (ATs): D-Phenylglycine aminotransferase (D-PhgAT) from *P. stutzeri* and amino-pentol AT (FumI) from *Sphingopyxis sp.* MTA144.

D-PhgAT is a unique member of the AT family as the amino donors in the 'forward' and 'backward' reaction exhibit opposite absolute stereochemical configuration. Thus, D-PhgAT is referred as 'stereo inverting' and is a very appealing target for the production of high value D-amino acids from inexpensive L-amino acids. D-PhgAT was successfully expressed, purified and kinetically characterised. Two different methods, a spectrophotometric method and a high-performance liquid chromatography (HPLC) based method, have been developed to determine the enzyme kinetic parameters and the enantioselectivity. The D-PhgAT amino donor substrate scope has been widely explored identifying L-Glu as the best amino donor. Thus, a large scale biotransformation has been carried out using L-Glu as amino donor and benzoyl formate (BZF) (1 g) as the model substrate and the percentage conversion and *ee* has been determined. A large excess of the amino donor L-Glu (~ 20-fold K_M) is enough to drive the reaction towards product formation and obtain 93% conversion of BZF after 3 h of reaction with an enzyme loading of 1 mg mL^{-1}. Moreover, the bioconversion has been extended to non-natural substrates to yield a wide range of aromatic D-amino acids: D-Phe, D-Tyr and D-Trp exhibiting promising conversions (from 15% to 57%).

The successful determination of the crystal structure of the D-PhgAT at 2.25 Å resolution with its bound PLP cofactor has begun to shed light on the unique stereo-inverting and enantioselective properties of the enzyme. An in depth structural and sequence analysis has highlighted three key active site residues that potentially control the exquisite (*R*)-selectivity of the enzyme; His66, His 213 and Arg34. These residues are shared between the two D-PhgAT homologues from *P. stutzeri* and *P. putida.* Site-directed mutagenesis on these residues will further clarify their role in the substrate recognition.

A strategy to overcome the high costs associated with enzymes consists of immobilizing an enzyme on a solid support to allow the reuse of an enzyme multiple times

without additional isolation and purification. Preliminary studies using a commercial resin, EziG, have been carried out and D-PhgAT was successfully immobilized and recycled up to 9 times. This could be further explored to make the process industrially relevant. All these findings strengthen the D-PhgAT position in the biocatalytic toolbox of the ATs for the production of a wide range of D-aminoacids.

The amino-pentol AT FumI is a very attractive target for biocatalysis as its natural substrate is 20 carbons long, suggesting that it could be potentially applied for the production of long chain amines which are broadly used in detergents, surfactants and cosmetic formulations.

Surprisingly, FumI was able to accept a wide range of ketones and aldehydes of various chain length (from C_3 to C_{20}) yielding the corresponding amines. A coupled assay and an LC ESI-MS method have been developed for the determination of the kinetic parameters and the detection of the product. Derivatisation using the widely used Marfey's reagent has been carried out to check the enzyme enantioselectivity (*S*).

The determination of the FumI 3D structure with its cofactor at 1.6 Å resolution revealed a potential hydrophobic binding site for the long chain substrates. Moreover, an unsual arginine residue, Arg231 directly interacts with the hydroxyl group of the PLP cofactor. Site-directed mutagenesis results were inconclusive as the substitution of the Arg residue to an Ala residue (Q231A) resulted in insoluble protein. The determination of an external aldimine structure with the natural substrate HFB_1 or Ala bound could provide further insight into the residues involved in substrate recognition and highlight hot spots for site directed mutagenesis.

Moreover, the possibility of coupling FumI with a carboxylic acid reductase (CAR) enzyme from *Mycobacterium chlorophenolicum* (McCAR) for the production of fatty amines from fatty acids has been explored. Very promising preliminary results suggest that the two enzymes are compatible and catalyse the production of primary amines from C_{10} to C_{20}. The determination of the % conversions for all the tested substrates will be crucial to optimize the reaction conditions and bring it to industrial level.

Until now, the application of FumI has been limited to detoxification of maize contaminants. In this work FumI is presented in a new light as a promising addition to the biocatalytic toolbox for the production of a wide range of fatty amines.

7 References

(1) Wenda, S.; Illner, S.; Mell, A.; Kragl, U. Industrial Biotechnology—the Future of Green Chemistry? *Green Chem.* **2011**, *13* (11), 3007.

(2) Laird, T. Biocatalysis and Biotransformations. *Org. Process Res. Dev.* **2005**, *9* (5), 521.

(3) Meyer, H.-P.; Eichhorn, E.; Hanlon, S.; Lütz, S.; Schürmann, M.; Wohlgemuth, R.; Coppolecchia, R. The Use of Enzymes in Organic Synthesis and the Life Sciences: Perspectives from the Swiss Industrial Biocatalysis Consortium (SIBC). *Catal. Sci. Technol.* **2013**, *3* (1), 29–40.

(4) Pollard, D. J.; Woodley, J. M. Biocatalysis for Pharmaceutical Intermediates: The Future Is Now. *Trends Biotechnol.* **2007**, *25* (2), 66–73.

(5) Sheldon, R. A. Biocatalysis and Green Chemistry. *Green Biocatal.* **2016**, 1–15.

(6) Sethi, M. K.; Chakraborty, P.; Shukla, R. *Biocatalysis – A Greener Alternative in Synthetic Chemistry*; The Royal Society of Chemistry, 2017.

(7) Sheldon, R. A.; Pereira, P. C. Biocatalysis Engineering: The Big Picture. *Chem. Soc. Rev.* **2017**.

(8) Johannes, T. W.; Simurdiak, M. R.; Zhao, H. Biocatalysis. *Encycl. Chem. Process.* **2006**, 101–110.

(9) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485* (7397), 185–194.

(10) Sedlaczek, L.; Smith, L. L. Biotransformations of Steroids. *Crit. Rev. Biotechnol.* **1988**, *7* (3), 187–236.

(11) Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chemie - Int. Ed.* **2018**, *57* (16), 4143–4148.

(12) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev.*

Genet. **2015**, *16* (7), 379–394.

(13) Bornscheuer, U. T. The Fourth Wave of Biocatalysis Is Approaching. *Philos. Trans. A. Math. Phys. Eng. Sci.* **2018**, *376* (2110), 20170063.

(14) Liese, J. (Alex) T. L. *Biocatalysis for the Pharmaceutical Industry*; Tao, J. (Alex), Lin, G.-Q., Liese, A., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 2009.

(15) Albarrán-Velo, J.; González-Martínez, D.; Gotor-Fernández, V. Stereoselective Biocatalysis: A Mature Technology for the Asymmetric Synthesis of Pharmaceutical Building Blocks. *Biocatal. Biotransformation* **2018**, *36* (2), 102–130.

(16) Gardossi, L.; Molinari, F. Biocatalytic Process. *Catalysis* **2009**, No. 32, 516–538.

(17) Sánchez-Carrón, G.; Campopiano, D. J. *Contemporary Catalysis: Science, Technology, and Applications*; The Royal Society of Chemistry, 2017.

(18) Davis, B. G.; Boyer, V. Biocatalysis and Enzymes in Organic Synthesis. *Nat. Prod. Rep.* **2001**, *18* (6), 618–640.

(19) Choi, J. M.; Han, S. S.; Kim, H. S. Industrial Applications of Enzyme Biocatalysis: Current Status and Future Aspects. *Biotechnol. Adv.* **2015**, *33* (7), 1443–1454.

(20) Yamada, H.; Kobayashi, M. Nitrile Hydratase and Its Application to Industrial Production of Acrylamide. *Biosci. Biotechnol. Biochem.* **1996**, *60* (9), 1391–1400.

(21) Arshad, R.; Farooq, S.; Ali, S. S. 6-Aminopenicillanic Acid Production by Intact Cells of E. Coli Containing Penicillin G Acylase (PGA). *Pakistan J. Biol. Sci.* **2007**, *10* (18), 3190–3194.

(22) Liu, Z. Q.; Zheng, W.; Huang, J. F.; Jin, L. Q.; Jia, D. X.; Zhou, H. Y.; Xu, J. M.; Liao, C. J.; Cheng, X. P.; Mao, B. X.; et al. Improvement and Characterization of a Hyperthermophilic Glucose Isomerase from Thermoanaerobacter Ethanolicus and Its Application in Production of High Fructose Corn Syrup. *J. Ind. Microbiol. Biotechnol.* **2015**, *42* (8), 1091–1103.

(23) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial

Biocatalysis Today and Tomorrow. *Nature* **2001**, *409* (6817), 258–268.

(24) Bommarius, A. S.; Paye, M. F. Stabilizing Biocatalysts. *Chem. Soc. Rev.* **2013**, *42* (15), 6534–6565.

(25) Datta, S.; Christena, L. R.; Rajaram, Y. R. S. Enzyme Immobilization: An Overview on Techniques and Support Materials. *3 Biotech* **2013**, *3* (1), 1–9.

(26) Turner, N. J. Directed Evolution Drives the next Generation of Biocatalysts. *Nat. Chem. Biol.* **2009**, *5* (8), 567–573.

(27) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; et al. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science (80-.).* **2010**, *329* (5989), 305–309.

(28) Desai, A. A. Sitagliptin Manufacture: A Compelling Tale of Green Chemistry, Process Intensification, and Industrial Asymmetric Catalysis. *Angew. Chemie - Int. Ed.* **2011**, *50* (9), 1974–1976.

(29) Reetz, M. T. What Are the Limitations of Enzymes in Synthetic Organic Chemistry? *Chem. Rec.* **2016**, *16* (6), 2449–2459.

(30) Sheldon, R. A.; Brady, D. The Limits to Biocatalysis : Pushing the Envelope. *Chem. Commun.* **2018**, *54*, 6088–6104.

(31) Capelli, R.; Marchetti, F.; Tiana, G.; Colombo, G. SAGE: A Fast Computational Tool for Linear Epitope Grafting onto a Foreign Protein Scaffold. *J. Chem. Inf. Model.* **2017**, *57* (1), 6–10.

(32) Lalonde, J. Highly Engineered Biocatalysts for Efficient Small Molecule Pharmaceutical Synthesis. *Curr. Opin. Biotechnol.* **2016**, *42*, 152–158.

(33) Bencivenni, G. Organocatalytic Strategies for the Synthesis of Axially Chiral Compounds. *Synlett* **2015**, *26* (14), 1915–1922.

(34) Andrushko, V.; Andrushko, N. Principles, Concepts and Strategies of Stereoselective

Synthesis. In *Stereoselective Synthesis of Drugs and Natural Products*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013; pp 1–42.

(35) Nguyen, L. A.; He, H.; Pham-Huy, C. Chiral Drugs: An Overview. *Int. J. Biomed. Sci.* **2006**, *2* (2), 85–100.

(36) Chhabra, N.; Aseri, M.; Padmanabhan, D. A Review of Drug Isomerism and Its Significance. *Int. J. Appl. Basic Med. Res.* **2013**, *3* (1), 16.

(37) Vargesson, N. Thalidomide-Induced Teratogenesis: History and Mechanisms. *Birth Defects Res. Part C - Embryo Today Rev.* **2015**, *105* (2), 140–156.

(38) Calcaterra, A.; D'Acquarica, I. The Market of Chiral Drugs: Chiral Switches versus de Novo Enantiomerically Pure Compounds. *J. Pharm. Biomed. Anal.* **2018**, *147*, 323–340.

(39) Nugent, T. C.; El-Shazly, M. Chiral Amine Synthesis - Recent Developments and Trends for Enamide Reduction, Reductive Amination, and Imine Reduction. *Adv. Synth. Catal.* **2010**, *352* (5), 753–819.

(40) Schrittwieser, J. H.; Velikogne, S.; Kroutil, W. Biocatalytic Imine Reduction and Reductive Amination of Ketones. *Adv. Synth. Catal.* **2015**, *357* (8), 1655–1685.

(41) Ghislieri, D.; Turner, N. J. Biocatalytic Approaches to the Synthesis of Enantiomerically Pure Chiral Amines. *Top. Catal.* **2014**, *57* (5), 284–300.

(42) Wu, H.; West, A. R.; Vickers, M.; Apperley, D. C.; Jones, A. G. Synthesis, Crystallization and Characterization of Diastereomeric Salts Formed by Ephedrine and Malic Acid in Water. *Chem. Eng. Sci.* **2012**, *77*, 47–56.

(43) Xue, Y.-P.; Cao, C.-H.; Zheng, Y.-G. Enzymatic Asymmetric Synthesis of Chiral Amino Acids. *Chem. Soc. Rev.* **2018**, *47*, 1516–1561.

(44) Ismail, H.; Lau, R. M.; Van Rantwijk, F.; Sheldon, R. A. Fully Enzymatic Resolution of Chiral Amines: Acylation and Deacylation in the Presence of Candida Antarctica Lipase B. *Adv. Synth. Catal.* **2008**, *350* (10), 1511–1516.

(45) Gotor-Fernández, V.; Brieva, R.; Gotor, V. Lipases: Useful Biocatalysts for the Preparation of Pharmaceuticals. *J. Mol. Catal. B Enzym.* **2006**, *40* (3–4), 111–120.

(46) Grogan, G. Synthesis of Chiral Amines Using Redox Biocatalysis. *Curr. Opin. Chem. Biol.* **2018**, *43*, 15–22.

(47) Jeon, H.; Yoon, S.; Ahsan, M.; Sung, S.; Kim, G.-H.; Sundaramoorthy, U.; Rhee, S.-K.; Yun, H. The Kinetic Resolution of Racemic Amines Using a Whole-Cell Biocatalyst Co-Expressing Amine Dehydrogenase and NADH Oxidase. *Catalysts* **2017**, *7* (9), 251.

(48) Abrahamson, M. J.; Vázquez-Figueroa, E.; Woodall, N. B.; Moore, J. C.; Bommarius, A. S. Development of an Amine Dehydrogenase for Synthesis of Chiral Amines. *Angew. Chemie - Int. Ed.* **2012**, *51* (16), 3969–3972.

(49) Zhao, H.; Van Der Donk, W. A. Regeneration of Cofactors for Use in Biocatalysis. *Curr. Opin. Biotechnol.* **2003**, *14* (6), 583–589.

(50) Aleku, G. A.; France, S. P.; Man, H.; Mangas-Sanchez, J.; Montgomery, S. L.; Sharma, M.; Leipold, F.; Hussain, S.; Grogan, G.; Turner, N. J. A Reductive Aminase from Aspergillus Oryzae. *Nat. Chem.* **2017**, *9* (10), 961–969.

(51) Cosgrove, S. C.; Brzezniak, A.; France, S. P.; Ramsden, J. I.; Mangas-Sanchez, J.; Montgomery, S. L.; Heath, R. S.; Turner, N. J. Imine Reductases, Reductive Aminases, and Amine Oxidases for the Synthesis of Chiral Amines: Discovery, Characterization, and Synthetic Applications. In *Methods in Enzymology*; Elsevier Inc., 2018; pp 1–19.

(52) Gomm, A.; O'Reilly, E. Transaminases for Chiral Amine Synthesis. *Curr. Opin. Chem. Biol.* **2018**, *43*, 106–112.

(53) Malik, M. S.; Park, E. S.; Shin, J. S. Features and Technical Applications of ω-Transaminases. *Appl. Microbiol. Biotechnol.* **2012**, *94* (5), 1163–1171.

(54) Holden, M. Pyridoxamine Phosphate-Oxidase and Pyridoxal Phosphate- Phosphatase Activities in Escherichia Coli. **1961**, No. 1930, 364–372.

(55) Schneider, G.; Käck, H.; Lindqvist, Y. The Manifold of Vitamin B6 Dependent

Enzymes. *Structure* **2000**, *8* (1), R1–R6.

(56) Heyl, D.; Harris, S. A.; Folkers, K. The Chemistry of Vitamin B6. VI. Pyridoxylamino Acids. *J. Am. Chem. Soc.* **1948**, *70* (10), 3429–3431.

(57) Richard, J. P.; Amyes, T. L.; Crugeiras, J.; Rios, A. The PLP Cofactor : Lessons from Studies on Model Reactions ☆. *BBA - Proteins Proteomics* **2011**, *1814* (11), 1419–1425.

(58) Bugg, T. D. H. *Introduction to Enzyme and Coenzyme Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, 2012.

(59) Percudani, R.; Peracchi, A. A Genomic Overview of Pyridoxal-Phosphate-Dependent Enzymes. *EMBO Rep.* **2003**, *4* (9), 850–854.

(60) Percudani, R.; Peracchi, A. The B6 Database: A Tool for the Description and Classification of Vitamin B6-Dependent Enzymatic Activities and of the Corresponding Protein Families. *BMC Bioinformatics* **2009**, *10*, 273.

(61) Paiardini, A.; Contestabile, R.; Buckle, A. M.; Cellini, B. PLP-Dependent Enzymes. *Biomed Res. Int.* **2014**, *2014*, 1–2.

(62) Mascarenhas, R.; Le, H. V.; Clevenger, K. D.; Lehrer, H. J.; Ringe, D.; Kelleher, N. L.; Silverman, R. B.; Liu, D. Selective Targeting by a Mechanism-Based Inactivator against Pyridoxal 5'-Phosphate-Dependent Enzymes: Mechanisms of Inactivation and Alternative Turnover. *Biochemistry* **2017**, *56* (37), 4951–4961.

(63) Clayton, P. T. B6-Responsive Disorders: A Model of Vitamin Dependency. *J. Inherit. Metab. Dis.* **2006**, *29* (2–3), 317–326.

(64) Harrison, P. J.; Dunn, T. M.; Campopiano, D. J. Sphingolipid Biosynthesis in Man and Microbes. *Nat. Prod. Rep.* **2018**.

(65) Chen, J.; Gong, X.; Li, J.; Li, Y.; Ma, J.; Hou, C.; Zhao, G.; Yuan, W.; Zhao, B. Carbonyl Catalysis Enables a Biomimetic Asymmetric Mannich Reaction. *Science (80-.).* **2018**, *360* (6396), 1438–1442.

(66) John R. A. Pyridoxal Phosphate-Dependent Enzymes. *Biochim. Biophys. Acta* **1995**, *1248*, 81–96.

(67) Metzler, D. E.; Ikawa, M.; Snell, E. E. A General Mechanism for Vitamin B6-Catalyzed Reactions. *J. Am. Chem. Soc.* **1954**, *76* (3), 648–652.

(68) Tidwell, T. T. Hugo (Ugo) Schiff, Schiff Bases, and a Century of β-Lactam Synthesis. *Angew. Chemie - Int. Ed.* **2008**, *47* (6), 1016–1020.

(69) Eliot, A. C.; Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **2004**, *73* (1), 383–415.

(70) Eliot, A. C.; Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **2004**, *73* (1), 383–415.

(71) Raman, M. C. C.; Johnson, K. a.; Yard, B. a.; Lowther, J.; Carter, L. G.; Naismith, J. H.; Campopiano, D. J. The External Aldimine Form of Serine Palmitoyltransferase: Structural, Kinetic and Spectroscopic Analysis of the Wild-Type Enzyme and HSAN1 Mutant Mimics. *J. Biol. Chem.* **2009**, *284* (25), 17328–17339.

(72) Phillips, R. S. Chemistry and Diversity of Pyridoxal-5'-Phosphate Dependent Enzymes. *Biochim. Biophys. Acta - Proteins Proteomics* **2015**, *1854* (9), 1167–1174.

(73) Dunathan, H. C. Conformation and Reaction Specificity in Pyridoxal Phosphate Enzymes. *Proc. Natl. Acad. Sci.* **1966**, *55* (4), 712–716.

(74) Yoshimura, T.; Jhee, K.-H.; Soda, K. Stereospecificity for the Hydrogen Transfer and Molecular Evolution of Pyridoxal Enzymes. *Biosci. Biotechnol. Biochem.* **1996**, *60* (2), 181–187.

(75) Di Salvo, M. L.; Contestabile, R.; Safo, M. K. Vitamin B6salvage Enzymes: Mechanism, Structure and Regulation. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814* (11), 1597–1608.

(76) Toney, M. D. Reaction Specificity in Pyridoxal Phosphate Enzymes. *Arch. Biochem. Biophys.* **2005**, *433* (1), 279–287.

(77) Toney, M. D. Controlling Reaction Specificity in Pyridoxal Phosphate Enzymes. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814* (11), 1407–1418.

(78) Ro, H. S.; Jeon, C. O.; Kim, H. S.; Sung, M. H. Stabilization of Quinonoid Intermediate E-Q by Glu32 of D-Amino Acid Transaminase. *J. Microbiol. Biotechnol.* **2006**, *16* (9), 1434–1440.

(79) Hirayama, A.; Miyanaga, A.; Kudo, F.; Eguchi, T. Mechanism-Based Trapping of the Quinonoid Intermediate by Using the K276R Mutant of PLP-Dependent 3-Aminobenzoate Synthase PctV in the Biosynthesis of Pactamycin. *ChemBioChem* **2015**, *16* (17), 2484–2490.

(80) Alexander, F. W.; Sandmeier, E.; Mehta, P. K.; Christen, P. Evolutionary Relationships Enzymes Regio-Specific a , P and Y Families. *Eur. J. Biochem.* **1994**, *219* (3), 953–960.

(81) Grishin, N. V; Phillips, M. a; Goldsmith, E. J. Modeling of the Spatial Structure of Eukaryotic Ornithine Decarboxylases. *Protein Sci.* **1995**, *4* (7), 1291–1304.

(82) Mehta, P. K.; Hale, T. I.; Christen, P. Aminotransferases: Demonstration of Homology and Division into Evolutionary Subgroups. *Eur. J. Biochem.* **1993**, *214* (2), 549–561.

(83) Slabu, I.; Galman, J. L.; Lloyd, R. C.; Turner, N. J. Discovery, Engineering, and Synthetic Application of Transaminase Biocatalysts. *ACS Catal.* **2017**, *7* (12), 8263–8284.

(84) Soda, K.; Yoshimura, T.; Esaki, N. Stereospecificity for the Hydrogen Transfer of Pyridoxal Enzyme Reactions. *Chem. Rec.* **2001**, *1* (5), 373–384.

(85) Berkovitch, F.; Behshad, E.; Tang, K.-H.; Enns, E. A.; Frey, P. A.; Drennan, C. L. A Locking Mechanism Preventing Radical Damage in the Absence of Substrate, as Revealed by the X-Ray Structure of Lysine 5,6-Aminomutase. *Proc. Natl. Acad. Sci.* **2004**, *101* (45), 15870–15875.

(86) Lepore, B. W.; Ruzicka, F. J.; Frey, P. A.; Ringe, D. The X-Ray Crystal Structure of Lysine-2,3-Aminomutase from Clostridium Subterminale. *Proc. Natl. Acad. Sci.* **2005**, *102* (39), 13819–13824.

(87) Graf von Stosch, A. Aspartate Aminotransferase Complexed with Erythro -β-Hydroxyaspartate: Crystallographic and Spectroscopic Identification of the Carbinolamine Intermediate. *Biochemistry* **1996**, *35* (48), 15260–15268.

(88) Shaw, J. P.; Petsko, G. A.; Ringe, D. Determination of the Structure of Alanine Racemase from Bacillus Stearothermophilus at 1.9-Å Resolution. *Biochemistry* **1997**, *36* (6), 1329–1342.

(89) Sugio, S.; Petsko, G. a.; Manning, J. M.; Soda, K.; Ringe, D. Crystal Structure of a D-Amino Acid Aminotransferase: How the Protein Controls Stereoselectivity. *Biochemistry* **1995**, *34* (Figure 1), 9661–9669.

(90) Barford, D.; Hu, S. H.; Johnson, L. N. Structural Mechanism for Glycogen Phosphorylase Control by Phosphorylation and AMP. *J. Mol. Biol.* **1991**, *218* (1), 233–260.

(91) Łyskowski, A.; Gruber, C.; Steinkellner, G.; Schürmann, M.; Schwab, H.; Gruber, K.; Steiner, K. Crystal Structure of an (R)-Selective ω-Transaminase from Aspergillus Terreus. *PLoS One* **2014**, *9* (1).

(92) Shin, J. S.; Kim, B. G. Comparison of the Omega-Transaminases from Different Microorganisms and Application to Production of Chiral Amines. *Bioscience, biotechnology, and biochemistry*. 2001, pp 1782–1788.

(93) Wybenga, G. G.; Crismaru, C. G.; Janssen, D. B.; Dijkstra, B. W. Structural Determinants of the β-Selectivity of a Bacterial Aminotransferase. *J. Biol. Chem.* **2012**, *287* (34), 28495–28502.

(94) Koszelewski, D.; Tauber, K.; Faber, K.; Kroutil, W. ω-Transaminases for the Synthesis of Non-Racemic α-Chiral Primary Amines. *Trends Biotechnol.* **2010**, *28* (6), 324–332.

(95) Truppo, M. D.; Rozzell, J. D.; Turner, N. J. Efficient Production of Enantiomerically Pure Chiral Amine at Conc 50 g/L Using Transaminase. *Org. Process Res. Dev.* **2010**, *14* (1), 234–237.

(96) Fuchs, C. S.; Simon, R. C.; Riethorst, W.; Zepeck, F.; Kroutil, W. Synthesis of (R)- or

(S)-Valinol Using ω-Transaminases in Aqueous and Organic Media. *Bioorg. Med. Chem.* **2014**, *22* (20), 5558–5562.

(97) Ward, J.; Wohlgemuth, R. High-Yield Biocatalytic Amination Reactions in Organic Synthesis. *Curr. Org. Chem.* **2010**, *14* (17), 1914–1927.

(98) O'Reilly, E.; Turner, N. J. Enzymatic Cascades for the Regio- and Stereoselective Synthesis of Chiral Amines. *Perspect. Sci.* **2015**, *4*, 55–61.

(99) Slabu, I.; Galman, J. L.; Lloyd, R. C.; Turner, N. J. Discovery, Engineering, and Synthetic Application of Transaminase Biocatalysts. *ACS Catal.* **2017**, *7* (12), 8263–8284.

(100) Höhne, M.; Kühl, S.; Robins, K.; Bornscheuer, U. T. Efficient Asymmetric Synthesis of Chiral Amines by Combining Transaminase and Pyruvate Decarboxylase. *ChemBioChem* **2008**, *9* (3), 363–365.

(101) Gomm, A.; Lewis, W.; Green, A. P.; O'Reilly, E. A New Generation of Smart Amine Donors for Transaminase-Mediated Biotransformations. *Chem. - A Eur. J.* **2016**, *22* (36), 12692–12695.

(102) Green, A. P.; Turner, N. J.; O'Reilly, E. Chiral Amine Synthesis Using ω-Transaminases: An Amine Donor That Displaces Equilibria and Enables High-Throughput Screening. *Angew. Chemie - Int. Ed.* **2014**, 10714–10717.

(103) Hailes, H.; Baud, D.; Ladkau, N.; Moody, T.; Ward, J. M. A Rapid, Sensitive Colorimetric Assay for the High-Throughput Screening of Transaminases in Liquid or Solid-Phase. *Chem. Commun.* **2015**, *51*, 17225–17228.

(104) Guan, L.-J.; Ohtsuka, J.; Okai, M.; Miyakawa, T.; Mase, T.; Zhi, Y.; Hou, F.; Ito, N.; Iwasaki, A.; Yasohara, Y.; et al. A New Target Region for Changing the Substrate Specificity of Amine Transaminases. *Sci. Rep.* **2015**, *5*, 10753.

(105) Genz, M.; Vickers, C.; van den Bergh, T.; Joosten, H. J.; Dörr, M.; Höhne, M.; Bornscheuer, U. T. Alteration of the Donor/acceptor Spectrum of the (S)-Amine Transaminase from Vibrio Fluvialis. *Int. J. Mol. Sci.* **2015**, *16* (11), 26953–26963.

(106) Berglund, P.; Humble, M. S.; Branneby, C. C–X Bond Formation: Transaminases as Chiral Catalysts: Mechanism, Engineering, and Applications. In *Comprehensive Chirality*; Elsevier, 2012; Vol. 7, pp 390–401.

(107) Hwang, B. Y.; Cho, B. K.; Yun, H.; Koteshwar, K.; Kim, B. G. Revisit of Aminotransferase in the Genomic Era and Its Application to Biocatalysis. *J. Mol. Catal. B Enzym.* **2005**, *37* (1–6), 47–55.

(108) Steffen-Munsberg, F.; Vickers, C.; Kohls, H.; Land, H.; Mallin, H.; Nobili, A.; Skalden, L.; van den Bergh, T.; Joosten, H.-J.; Berglund, P.; et al. Bioinformatic Analysis of a PLP-Dependent Enzyme Superfamily Suitable for Biocatalytic Applications. *Biotechnol. Adv.* **2015**, *33* (5), 566–604.

(109) Morrison, J. L.; Breitling, R.; Higham, D. J.; Gilbert, D. R. A Lock-and-Key Model for Protein-Protein Interactions. *Bioinformatics* **2006**, *22* (16), 2012–2019.

(110) Hirotsu, K.; Goto, M.; Okamoto, A.; Miyahara, I. Dual Substrate Recognition of Aminotransferases. *Chem. Rec.* **2005**, *5* (3), 160–172.

(111) Malashkevish, V.; Onuffer; Kirsch; Jansonius, J. Alternating Arginine -Modulated Substrate Specificity in an Engineered Tyrosinase Aminotransferase. *Nat. sturctural Biol.* **1995**, *2* (7), 548–553.

(112) Kamitori, S.; Okamoto, A.; Hirotsu, K.; Higuchi, T.; Kuramitsu, S.; Kagamiyama, H.; Matsuura, Y.; Katsube, Y. Three-Dimensional Structures of Aspartate Aminotransferase from *Escherichia Coli* and Its Mutant Enzyme at 2.5 A Resolution. *J. Biochem.* **1990**, *108* (2), 175–184.

(113) Steffen-Munsberg, F.; Vickers, C.; Thontowi, A.; Schätzle, S.; Meinhardt, T.; SvedendahlHumble, M.; Land, H.; Berglund, P.; Bornscheuer, U. T.; Höhne, M. Revealing the Structural Basis of Promiscuous Amine Transaminase Activity. *ChemCatChem* **2013**, *5* (1), 154–157.

(114) Park, E.-S.; Kim, M.; Shin, J.-S. Molecular Determinants for Substrate Selectivity of ω-Transaminases. *Appl. Microbiol. Biotechnol.* **2012**, *93* (6), 2425–2435.

(115) Humble, M. S.; Cassimjee, K. E.; Hãkansson, M.; Kimbung, Y. R.; Walse, B.; Abedi, V.; Federsel, H. J.; Berglund, P.; Logan, D. T. Crystal Structures of the Chromobacterium Violaceumω-Transaminase Reveal Major Structural Rearrangements upon Binding of Coenzyme PLP. *FEBS J.* **2012**, *279* (5), 779–792.

(116) Kaulmann, U.; Smithies, K.; Smith, M. E. B.; Hailes, H. C.; Ward, J. M. Substrate Spectrum of ω-Transaminase from Chromobacterium Violaceum DSM30191 and Its Potential for Biocatalysis. *Enzyme Microb. Technol.* **2007**, *41* (5), 628–637.

(117) Rausch, C.; Lerchner, A.; Schiefner, A.; Skerra, A. Crystal Structure of the ω-Aminotransferase from Paracoccus Denitrificans and Its Phylogenetic Relationship with Other Class III Amino- Transferases That Have Biotechnological Potential. *Proteins Struct. Funct. Bioinforma.* **2013**, *81* (5), 774–787.

(118) Denesyuk, A. I.; Denessiouk, K. a; Korpela, T.; Johnson, M. S. Functional Attributes of the Phosphate Group Binding Cup of Pyridoxal Phosphate-Dependent Enzymes. *J. Mol. Biol.* **2002**, *316* (1), 155–172.

(119) Al Toma, R. S.; Brieke, C.; Cryle, M. J.; Süssmuth, R. D. Structural Aspects of Phenylglycines, Their Biosynthesis and Occurrence in Peptide Natural Products. *Nat. Prod. Rep.* **2015**, *32* (8), 1207–1235.

(120) Mast, Y. J.; Wohlleben, W.; Schinko, E. Identification and Functional Characterization of Phenylglycine Biosynthetic Genes Involved in Pristinamycin Biosynthesis in Streptomyces Pristinaespiralis. *J. Biotechnol.* **2011**, *155* (1), 63–67.

(121) Al Toma, R. S.; Brieke, C.; Cryle, M. J.; Süssmuth, R. D. Structural Aspects of Phenylglycines, Their Biosynthesis and Occurrence in Peptide Natural Products. *Nat. Prod. Rep.* **2015**, *32* (8), 1207–1235.

(122) van Wageningen, A. M. A.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J. M.; Solenberg, P. J. Sequencing and Analysis of Genes Involved in the Biosynthesis of a Vancomycin Group Antibiotic. *Chem. Biol.* **1998**, *5* (3), 155–162.

(123) Hubbard, B. K.; Thomas, M. G.; Walsh, C. T. Biosynthesis of L-P-Hydroxyphenylglycine, a Non-Proteinogenic Amino Acid Constituent of Peptide Antibiotics. *Chem. Biol.* **2000**, *7* (12), 931–942.

(124) Wegman, M. A.; Janssen, M. H. A.; van Rantwijk, F.; Sheldon, R. A. Towards Biocatalytic Synthesis of β-Lactam Antibiotics. *Adv. Synth. Catal.* **2001**, *343* (6–7), 559–576.

(125) Leuchtenberger, W.; Huthmacher, K.; Drauz, K. Biotechnological Production of Amino Acids and Derivatives: Current Status and Prospects. *Appl. Microbiol. Biotechnol.* **2005**, *69* (1), 1–8.

(126) Guo, F.; Berglund, P. Transaminase Biocatalysis: Optimization and Application. *Green Chem.* **2017**, *19* (2), 333–360.

(127) Wiyakrutta, S.; Meevootisom, V. A Stereo-Inverting D-Phenylglycine Aminotransferase from Pseudomonas Stutzeri ST-201: Purification, Characterization and Application for D-Phenylglycine Synthesis. *J. Biotechnol.* **1997**, *55* (3), 193–203.

(128) Rojanarata, T.; Opanasopit, P.; Ngawhirunpat, T.; Saehuan, C.; Wiyakrutta, S.; Meevootisom, V. A Simple, Sensitive and Green Bienzymatic UV-Spectrophotometric Assay of Amoxicillin Formulations. *Enzyme Microb. Technol.* **2010**, *46* (3–4), 292–296.

(129) Khampha, W.; Meevootisom, V.; Wiyakrutta, S. Spectrophotometric Enzymatic Cycling Method Using L-Glutamate Dehydrogenase and D-Phenylglycine Aminotransferase for Determination of L-Glutamate in Foods. *Anal. Chim. Acta* **2004**, *520* (1–2), 133–139.

(130) Müller, U.; van Assema, F.; Gunsior, M.; Orf, S.; Kremer, S.; Schipper, D.; Wagemans, A.; Townsend, C. A.; Sonke, T.; Bovenberg, R.; et al. Metabolic Engineering of the E. Coli L-Phenylalanine Pathway for the Production of D-Phenylglycine (D-Phg). *Metab. Eng.* **2006**, *8* (3), 196–208.

(131) Kongsaeree, P.; Samanchart, C.; Laowanapiban, P.; Wiyakrutta, S.; Meevootisom, V.

Crystallization and Preliminary X-Ray Crystallographic Analysis of D-Phenylglycine Aminotransferase from Pseudomonas Stutzeri ST201. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2003**, *59* (5), 953–954.

(132) Visek, K. Amines, Fatty. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003; Vol. 2, pp 518–537.

(133) Barrault, J.; Pouilloux, Y. Synthesis of Fatty Amines. Selectivity Control in Presence of Multifunctional Catalysts. *Catal. Today* **1997**, *37* (2), 137–153.

(134) Nelson, P. E.; Desjardins, A. E.; Plattner, R. D. Fumonisins, Mycotoxins Produced by Fusarium Species: Biology, Chemistry, and Significance. *Annu. Rev. Phytopathol.* **1993**, *31* (1), 233–252.

(135) Hartinger, D.; Moll, W. Fumonisin Elimination and Prospects for Detoxification by Enzymatic Transformation. *World Mycotoxin J.* **2011**, *4* (3), 271–283.

(136) Zitomer, N. C.; Mitchell, T.; Voss, K. a.; Bondy, G. S.; Pruett, S. T.; Garnier-Amblard, E. C.; Liebeskind, L. S.; Park, H.; Wang, E.; Sulllards, M. C.; et al. Ceramide Synthase Inhibition by Fumonisin B1 Causes Accumulation of 1-Deoxysphinganine. A Novel Category of Bioactive 1-Deoxysphingoid Bases and 1-Deoxydihydroceramides Biosynthesized by Mammalian Cell Lines and Animals. *J. Biol. Chem.* **2009**, *284* (8), 4786–4795.

(137) Marasas, W. F. O.; Riley, R. T.; Hendricks, K. A.; Stevens, V. L.; Sadler, T. W.; Gelineau-van Waes, J.; Missmer, S. A.; Cabrera, J.; Torres, O.; Gelderblom, W. C. A.; et al. Fumonisins Disrupt Sphingolipid Metabolism, Folate Transport, and Neural Tube Development in Embryo Culture and In Vivo: A Potential Risk Factor for Human Neural Tube Defects among Populations Consuming Fumonisin-Contaminated Maize. *J. Nutr.* **2004**, *134* (4), 711–716.

(138) Dragan, Y. P.; Bidlack, W. R.; Cohen, S. M.; Goldsworthy, T. L.; Hard, G. C.; Howard, P. C.; Riley, R. T.; Voss, K. A. Implications of Apoptosis for Toxicity, Carcinogenicity and Risk Assessment: Fumonisin B1 as an Example. *Toxicol. Sci.* **2001**, *61* (1), 6–17.

(139) Gelderblom, W. C. A.; Marasas, W. F. O.; Jaskiewicz, K.; Combrinck, S.; Vanschalkwyk, D. J. Cancer Promoting Potential of Different Strains of Fusarium-Moniliforme in a Short-Term Cancer Initiation Promotion Assay. *Carcinogenesis* **1988**, *9* (8), 1405–1409.

(140) Heinl, S.; Hartinger, D.; Thamhesl, M.; Vekiru, E.; Krska, R.; Schatzmayr, G.; Moll, W. D.; Grabherr, R. Degradation of Fumonisin B1 by the Consecutive Action of Two Bacterial Enzymes. *J. Biotechnol.* **2010**, *145* (2), 120–129.

(141) Heinl, S.; Hartinger, D.; Thamhesl, M.; Schatzmayr, G.; Moll, W. D.; Grabherr, R. An Aminotransferase from Bacterium ATCC 55552 Deaminates Hydrolyzed Fumonisin B1. *Biodegradation* **2011**, *22* (1), 25–30.

(142) Hartinger, D.; Schwartz, H.; Hametner, C.; Schatzmayr, G.; Haltrich, D.; Moll, W. D. Enzyme Characteristics of Aminotransferase FumI of Sphingopyxis Sp. MTA144 for Deamination of Hydrolyzed Fumonisin B 1. *Appl. Microbiol. Biotechnol.* **2011**, *91* (3), 757–768.

(143) Hartinger, D.; Heinl, S.; Schwartz, H. E.; Grabherr, R.; Schatzmayr, G.; Haltrich, D.; Moll, W. Enhancement of Solubility in Escherichia Coli and Purification of an Aminotransferase from Sphingopyxis Sp. MTA144 for Deamination of Hydrolyzed Fumonisin B(1). *Microb. Cell Fact.* **2010**, *9*, 62.

(144) Chantarasiri, A.; Meevootisom, V.; Isarangkul, D.; Wiyakrutta, S. Effective Improvement of D-Phenylglycine Aminotransferase Solubility by Protein Crystal Contact Engineering. *J. Mol. Microbiol. Biotechnol.* **2012**, *22* (3), 147–155.

(145) Jariyachawalid, K.; Laowanapiban, P.; Meevootisom, V.; Wiyakrutta, S. Effective Enhancement of Pseudomonas Stutzeri D-Phenylglycine Aminotransferase Functional Expression in Pichia Pastoris by Co-Expressing Escherichia Coli GroEL-GroES. *Microb. Cell Fact.* **2012**, *11* (1), 47.

(146) Bonissone, S.; Gupta, N.; Romine, M.; Bradshaw, R. A.; Pevzner, P. A. N-Terminal Protein Processing: A Comparative Proteogenomic Analysis. *Mol. Cell. Proteomics* **2013**, *12* (1), 14–28.

(147) Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; et al. Spontaneous α-N-6-Phosphogluconoylation of a "His Tag" inEscherichia coli:The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **1999**, *267* (1), 169–184.

(148) Baxter, S.; Royer, S.; Grogan, G.; Brown, F.; Holt-Tiffin, K. E.; Taylor, I. N.; Fotheringham, I. G.; Campopiano, D. J. An Improved Racemase/acylase Biotransformation for the Preparation of Enantiomerically Pure Amino Acids. *J. Am. Chem. Soc.* **2012**, *134* (47), 19310–19313.

(149) Faber, K. *Biotransformations in Organic Chemistry*; Springer International Publishing: Cham, 2018.

(150) Truppo, M. D.; Rozzell, J. D.; Moore, J. C.; Turner, N. J. Rapid Screening and Scale-up of Transaminase Catalysed Reactions. *Org. Biomol. Chem.* **2009**, *7* (2), 395–398.

(151) Truppo, M. D.; David Rozzell, J.; Turner, N. J. Efficient Production of Enantiomerically Pure Chiral Amines at Concentrations of 50 g/L Using Transaminases. *Org. Process Res. Dev.* **2010**, *14* (1), 234–237.

(152) Green, A. P.; Turner, N. J.; O'Reilly, E. Chiral Amine Synthesis Using ω-Transaminases: An Amine Donor That Displaces Equilibria and Enables High-Throughput Screening. *Angew. Chemie Int. Ed.* **2014**, *53* (40), 10714–10717.

(153) Hailes, H.; Baud, D.; Ladkau, N.; Moody, T.; Ward, J. M. A Rapid, Sensitive Colorimetric Assay for the High-Throughput Screening of Transaminases in Liquid or Solid-Phase. *Chem. Commun.* **2015**, *51*, 17225–17228.

(154) Zhou, Y.; Wu, S.; Li, Z. One-Pot Enantioselective Synthesis of D-Phenylglycines from Racemic Mandelic Acids, Styrenes, or Biobased L-Phenylalanine via Cascade Biocatalysis. *Adv. Synth. Catal.* **2017**, *359* (24), 4305–4316.

(155) Breuer, M.; Ditrich, K.; Habicher, T.; Hauer, B.; Keßeler, M.; Stürmer, R.; Zelinski, T. Industrial Methods for the Production of Optically Active Intermediates. *Angew. Chemie Int. Ed.* **2004**, *43* (7), 788–824.

(156) Tentolouris, N.; Voulgari, C.; Katsilambros, N. A Review of Nateglinide in the Management of Patients with Type 2 Diabetes. *Vasc. Health Risk Manag.* **2007**, *3* (6), 797–807.

(157) Kepert, I.; Fonseca, J.; Müller, C.; Milger, K.; Hochwind, K.; Kostric, M.; Fedoseeva, M.; Ohnmacht, C.; Dehmel, S.; Nathan, P.; et al. D-Tryptophan from Probiotic Bacteria Influences the Gut Microbiome and Allergic Airway Disease. *J. Allergy Clin. Immunol.* **2017**, *139* (5), 1525–1535.

(158) Hansford, K. A.; Reid, R. C.; Clark, C. I.; Tyndall, J. D. A.; Whitehouse, M. W.; Guthrie, T.; McGeary, R. P.; Schafer, K.; Martin, J. L.; Fairlie, D. P. D-Tyrosine as a Chiral Precusor to Potent Inhibitors of Human Nonpancreatic Secretory Phospholipase A2 (IIa) with Antiinflammatory Activity. *ChemBioChem* **2003**, *4* (2–3), 181–185.

(159) Qiu, J.; Su, E.; Wang, W.; Wei, D. High Yield Synthesis of D-Phenylglycine and Its Derivatives by Nitrilase Mediated Dynamic Kinetic Resolution in Aqueous-1-Octanol Biphasic System. *Tetrahedron Lett.* **2014**, *55* (8), 1448–1451.

(160) Liu, H.; Naismith, J. H. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008**, *8*, 91.

(161) van Oosterwijk, N.; Willies, S.; Hekelaar, J.; Terwisscha van Scheltinga, A. C.; Turner, N. J.; Dijkstra, B. W. Structural Basis of the Substrate Range and Enantioselectivity of Two (S)-Selective ω-Transaminases. *Biochemistry* **2016**, *55* (31), 4422–4431.

(162) Robert, X.; Gouet, P. Deciphering Key Features in Protein Structures with the New ENDscript Server. *Nucleic Acids Res.* **2014**, *42* (W1), 320–324.

(163) Jomrit, J.; Summpunn, P.; Meevootisom, V.; Wiyakrutta, S. Sensitive Non-Radioactive Determination of Aminotransferase Stereospecificity for C-4' Hydrogen Transfer on the Coenzyme. *Biochem. Biophys. Res. Commun.* **2011**, *405* (4), 626–631.

(164) Dickschat, J. S. Modern Aspects of Isotopic Labellings in Terpene Biosynthesis. *European J. Org. Chem.* **2017**, *2017* (33), 4872–4882.

(165) Rinkel, J.; Dickschat, J. S. Recent Highlights in Biosynthesis Research Using Stable

Isotopes. *Beilstein J. Org. Chem.* **2015**, *11*, 2493–2508.

(166) Thiverny, M.; Farran, D.; Philouze, C.; Blandin, V.; Chavant, P. Y. Totally Diastereoselective Addition of Aryl Grignard Reagents to the Nitrone-Based Chiral Glycine Equivalent MiPNO. *Tetrahedron: Asymmetry* **2011**, *22* (12), 1274–1281.

(167) Govindaraju, V.; Basus, V. J.; Matson, G. B.; Maudsley, A. A. Measurement of Chemical Shifts and Coupling Constants for Glutamate and Glutamine. *Magn. Reson. Med.* **1998**, *39* (6), 1011–1013.

(168) Ghisalba, O.; Meyer, H.-P.; Wohlgemuth, R. Industrial Biotransformation. *Encycl. Ind. Biotechnol. Bioprocess, Biosep. Cell Technol.* **2009**, *4* (1–2), 1–18.

(169) Straathof, A. J. J.; Panke, S.; Schmid, A. The Production of Fine Chemicals by Biotransformations. *Curr. Opin. Biotechnol.* **2002**, *13* (6), 548–556.

(170) Cassimjee, K. E.; Federsel, H.-J. CHAPTER 13. EziG: A Universal Platform for Enzyme Immobilisation. In *Biocatalysis: An Industrial Perspective*; The Royal Society of Chemistry, 2017; pp 345–362.

(171) Homaei, A. A.; Sariri, R.; Vianello, F.; Stevanato, R. Enzyme Immobilization: An Update. *J. Chem. Biol.* **2013**, *6* (4), 185–205.

(172) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72* (1–2), 248–254.

(173) Zhou, Y.; Wu, S.; Li, Z. One-Pot Enantioselective Synthesis of D -Phenylglycines from Racemic Mandelic Acids , Styrenes , or Biobased L -Phenylalanine via Cascade Biocatalysis. **2017**.

(174) Benedetti, R.; Nazzi, F.; Locci, R.; Firrao, G. Degradation of Fumonisin B1 by a Bacterial Strain Isolated from Soil. *Biodegradation* **2006**, *17* (1), 31–38.

(175) Marasa, W.; Miller, J.; RIley, R.; Visconti, A.; Marasas, W.; Miller, J.; RIley, R.; Visconti, A. Fumonisin B1. *Environ. Heal. Criteria* **2000**, 150.

(176) Walker, J. M. *Mitochondrial Disorders*; Wong, Ph.D., L.-J. C., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; Vol. 837.

(177) Bhushan, R.; Brückner, H. Marfey's Reagent for Chiral Amino Acid Analysis: A Review. *Amino Acids* **2004**, *27* (3–4), 231–247.

(178) Bhushan, R.; Brückner, H. Use of Marfey's Reagent and Analogs for Chiral Amino Acid Analysis: Assessment and Applications to Natural Products and Biological Systems. *J. Chromatogr. B* **2011**, *879* (29), 3148–3161.

(179) B'Hymer, C.; Montes-Bayon, M.; Caruso, J. A. Marfey's Reagent: Past, Present, and Future Uses of 1-Fluoro-2,4-Dinitrophenyl-5-L-Alanine Amide. *J. Sep. Sci.* **2003**, *26* (1–2), 7–19.

(180) Marfey, P. Determination of D-Amino Acids. II. Use of a Bifunctional Reagent, 1,5-Difluoro-2,4-Dinitrobenzene. *Carlsberg Res. Commun.* **1984**, *49* (6), 591–596.

(181) Bhushan, R.; Tanwar, S. RP-LC Resolution of (R,S)-Atenolol via Diastereomerization with Marfey's Reagent and Its Structural Variants Under Conventional and Microwave Heating. *Chromatographia* **2008**, *68* (9–10), 849–853.

(182) Hess, S.; Gustafson, K. R.; Milanowski, D. J.; Alvira, E.; Lipton, M. A.; Pannell, L. K. Chirality Determination of Unusual Amino Acids Using Precolumn Derivatization and Liquid Chromatography-Electrospray Ionization Mass Spectrometry. *J. Chromatogr. A* **2004**, *1035* (2), 211–219.

(183) Del Pozo, A. M.; Merola, M.; Ueno, H.; Manning, J. M.; Tanizawa, K.; Nishimura, K.; Soda, K.; Ringe, D. Stereospecificity of Reactions Catalyzed by Bacterial D-Amino Acid Transaminase. *J. Biol. Chem.* **1989**, *264* (30), 17784–17789.

(184) Akhtar, M. K.; Turner, N. J.; Jones, P. R. Carboxylic Acid Reductase Is a Versatile Enzyme for the Conversion of Fatty Acids into Fuels and Chemical Commodities. *Proc. Natl. Acad. Sci.* **2013**, *110* (1), 87–92.

(185) Winkler, M. Carboxylic Acid Reductase Enzymes (CARs). *Curr. Opin. Chem. Biol.* **2018**, *43* (2), 23–29.

(186) Gahloth, D.; Dunstan, M. S.; Quaglia, D.; Klumbys, E.; Lockhart-Cairns, M. P.; Hill, A. M.; Derrington, S. R.; Scrutton, N. S.; Turner, N. J.; Leys, D. Structures of Carboxylic Acid Reductase Reveal Domain Dynamics Underlying Catalysis. *Nat. Chem. Biol.* **2017**, *13* (9), 975–981.

(187) Stolterfoht, H.; Steinkellner, G.; Schwendenwein, D.; Pavkov-Keller, T.; Gruber, K.; Winkler, M. Identification of Key Residues for Enzymatic Carboxylate Reduction. *Front. Microbiol.* **2018**, *9* (2), 1–15.

(188) Hansen, E. H.; Møller, B. L.; Kock, G. R.; Bünner, C. M.; Kristensen, C.; Jensen, O. R.; Okkels, F. T.; Olsen, C. E.; Motawia, M. S.; Hansen, J. De Novo Biosynthesis of Vanillin in Fission Yeast (Schizosaccharomyces Pombe) and Baker's Yeast (Saccharomyces Cerevisiae). *Appl. Environ. Microbiol.* **2009**, *75* (9), 2765–2774.

(189) France, S. P.; Hussain, S.; Hill, A. M.; Hepworth, L. J.; Howard, R. M.; Mulholland, K. R.; Flitsch, S. L.; Turner, N. J. One-Pot Cascade Synthesis of Mono- and Disubstituted Piperidines and Pyrrolidines Using Carboxylic Acid Reductase (CAR), ω-Transaminase (ω-TA), and Imine Reductase (IRED) Biocatalysts. *ACS Catal.* **2016**, *6* (6), 3753–3759.

(190) Finnigan, W.; Thomas, A.; Cromar, H.; Gough, B.; Snajdrova, R.; Adams, J. P.; Littlechild, J. A.; Harmer, N. J. Characterization of Carboxylic Acid Reductases as Enzymes in the Toolbox for Synthetic Chemistry. *ChemCatChem* **2017**, *9* (6), 1005–1017.

(191) Griswold, W. R.; Toney, M. D. Role of the Pyridine Nitrogen in Pyridoxal 5'-Phosphate Catalysis: Activity of Three Classes of PLP Enzymes Reconstituted with Deazapyridoxal 5'-Phosphate. *J. Am. Chem. Soc.* **2011**, *133* (37), 14823–14830.

(192) Swinehart, D. F. The Beer-Lambert Law. *J. Chem. Educ.* **1962**, *39* (7), 333.

(193) Liu, C.; Molinski, T. F. Preparation of α-Amino Acids by Oxidative Oxazoline-Oxazinone Rearrangement-Hydrogenation (OOOH). Scope and Limitations. *Chem. - An Asian J.* **2011**, *6* (8), 2022–2027.

(194) Kabsch, W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta*

Crystallogr. Sect. D Biol. Crystallogr. **2010**, *66* (2), 133–144.

(195) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (7), 1204–1214.

(196) Karplus, P. A.; Diederichs, K. Linking Crystallographic Model and Data Quality. *Science (80-.).* **2012**, *336* (6084), 1030–1033.

(197) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (Pt 4), 658–674.

(198) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(199) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (4), 352–367.

(200) Nicholls, R. A. Ligand Fitting with CCP 4. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73* (2), 158–170.

(201) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity : All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.

8 Appendices

GGCAGCAGCGGCAGCATTCTGAACGACTATAAGCGTAAAACCGAAGGTAGCGTGTTCTGGGCGCA
GCGTGCGCGTAGCGTGATGCCGGATGGCGTGACCGCGGACACCCGTGTGTTCGATCCGCACGGTC
TGTTTATTAGCGACGCGCAGGGCGTTCACAAGACCGACGTGGATGGTAACGTTTACCTGGATTTCTT
TGGTGGCCATGGTGCGCTGGTGCTGGGTCATGGTCACCCGCGTGTTAACGCGGCGATTGCGGAGG
CGCTGAGCCACGGTGTGCAATATGCGGCGAGCCACCCGCTGGAAGTGCGTTGGGCGGAACGTATT
GTTGCGGCGTTCCCGAGCATCCGTAAGCTGCGTTTTACCGGTAGCGGTACCGAAACCACCCTGCTG
GCGCTGCGTGTGGCGCGTGCGTTTACCGGTCGTCGTATGATCCTGCGTATTGCGACCCACTATCATG
GTTGGCATGACTTCAGCGCGAGCGGTTATAACAGCCATTTTGATGGTCAACCGGCGCCGGGCGTTC
TGCCGGAGATTGCGAAAAACACCCTGCTGATTCGTCCGGACGATATCGAGGGCATGCGTGAAGTGT
TCGCGCAACACGGTAGCGACATTGCGGCGTTCATCGCGGAACCGGTTGGTAGCCACTTTGGCGTGA
CCCCGGTTAGCGATAGCTTCCTGCGTGAAGGTGCTGAACTGGCGCGTCAGTATGGTGCGCTGTTTA
TTCTGGACGAGGTGATCAGCGGTTTCCGTGTTGGTAACCACGGCATGCAGGCGCTGCTGGACGTTC
AACCGGATCTGACCTGCCTGGCGAAGGCGAGCGCGGGTGGCCTGCCGGGTGGCATCCTGGGTGGC
CGTGAAGACGTGATGGGTGTTCTGAGCCGTGGCAGCGATCGTAAAGTGCTGCACCAGGGTACCTT
CACCGGCAACCCGATTACCGCGGCGGCGGCGATTGCGGCGATTGACACCATCCTGGAGGACGATG
TTTGCGCGAAGATCAACGATCTGGGCCAATTTGCGCGTGAAGCGATGAACCACCTGTTCGCGCGTA
AAGGTCTGAACTGGCTGGCGTATGGTCGTTTCAGCGGTTTTCATCTGATGCCGGGCCTGCCGCCGA
ACACCACCGACACCGGTAGCATTACCCGTGCGGAAGTGGCGCGTCCGGATGTTAAAATGATCGCG
GCGATGCGTATGGCGCTGATTCTGGAAGGTGTGGACATCGGTGGCCGTGGCAGCGTTTTTCTGAGC
GCGCAACACGAGCGTGAACACGTGGAGCACCTGGTTACCACCTTCGACCGTGTGCTGGATCGTCTG
GCGGATGAGAATCTGCTGAGCTGGCAACCGACCAACCTGAGCGGCAATCAAAGCTAAATAA

Appendix 8.1 *DpgA* nucleotide sequence.

MSHHHHHHGSSGSILNDYKRKTEGSVFWAQRARSVMPDGVTADTRVFDPHGLFISDAQGVHKTDVD

GNVYLDFFGGHGALVLGHGHPRVNAAIAEALSHGVQYAASHPLEVRWAERIVAAFPSIRKLRFTGSGTE

TTLLALRVARAFTGRRMILRIATHYHGWHDFSASGYNSHFDGQPAPGVLPEIAKNTLLIRPDDIEGMREV

FAQHGSDIAAFIAEPVGSHFGVTPVSDSFLREGAELARQYGALFILDEVISGFRVGNHGMQALLDVQPDL

TCLAKASAGGLPGGILGGREDVMGVLSRGSDRKVLHQGTFTGNPITAAAAIAAIDTILEDDVCAKINDLG

QFAREAMNHLFARKGLNWLAYGRFSGFHLMPGLPPNTTDTGSITRAEVARPDVKMIAAMRMALILEG

VDIGGRGSVFLSAQHEREHVEHLVTTFDRVLDRLADENLLSWQPTNLSGNQS

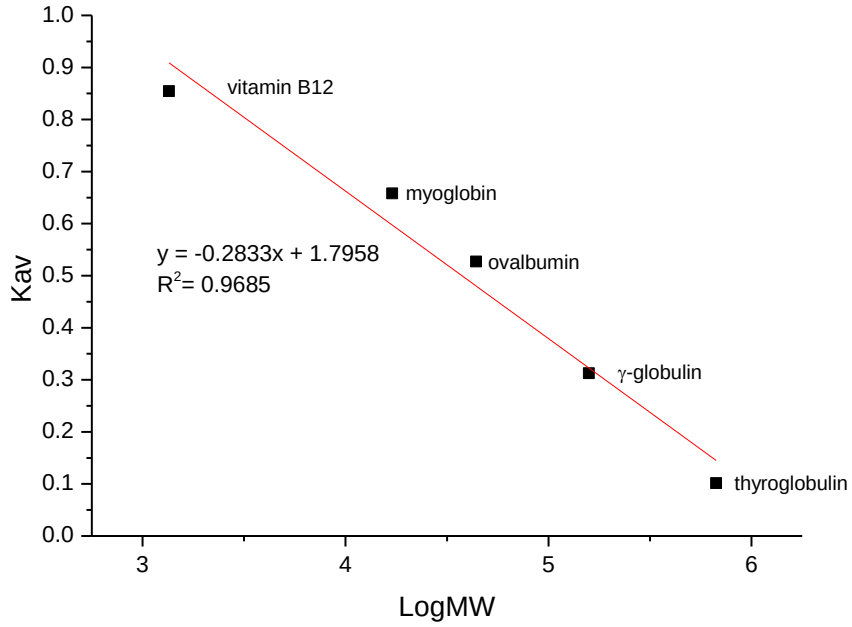
Appendix 8.2 Recombinant DpgA amino acid sequence. The N-terminal His6-tag is highlighted in bold.

ATGGCGAACGGAACAAGGCAGAAAGATCTCAGAGAACGCGCCGAACGGGTCATTCCGGGCGGGA
TGTACGGCCACGAGTCGACACGGTTGCTGCCGCCAGAATTCCCCCAGTTCTTCAGGCGCGCGCTGG
GGGCACGAATTTGGGACGCCGACGAGCAGCCCTATATCGACTATATGTGCGCGTATGGGCCAAATT
TGCTCGGTTACCGGCAATCCGAAATCGAAGCCGCGGCTGATGCGCAGCGACTTCTCGGCGACACCA
TGACCGGTCCTTCGGAGATCATGGTCAACCTCGCCGAAGCCTTTGTGGGCATGGTCCGTCATGCGG
ATTGGGCGATGTTCTGCAAAAATGGCAGCGATGCCACCTCAACGGCGATGGTTCTCGCGCGTGCCC
ATACGGGGCGCAAAACCATATTATGCGCCAAAGGCGCCTATCATGGCGCTTCCCCGTGGAACACTC
CGCATACTGCCGGGATTCTCGCTTCCGATCGCGTGCATGTCGCATATTATACCTATAACGACGCCCA
AAGCTTATCGGACGCGTTCAAGGCGCACGATGGCGATATTGCGGCTGTCTTTGCCACACCTTTCCGA
CACGAAGTATTTGAGGACCAGGCCCTCGCCCAGCTTGAGTTCGCGCGCACCGCTCGAAAATGTTGT
GACGAGACCGGTGCGCTTCTGGTCGTTGACGATGTGCGCGCAGGTTTCCGGGTGGCGCGCGATTG
CAGCTGGACGCATTTGGGTATCGAACCCGATCTCAGTTGCTGGGGAAAATGCTTTGCGAATGGCTA
TCCGATCTCCGCCCTGCTGGGCTCGAACAAGGCGCGCGATGCGGCGCGGGATATATTTGTGACCGG
CTCCTTCTGGTTCTCTGCGGTACCGATGGCGGCCGCGATCGAAACCCTCAGGATCATTCGAGAGAC
GCCTTATCTCGAAACGCTGATCGCCAGCGGCGCCGCCCTGCGGGCAGGCCTGGAGGCACAGTCTCA
GCGCCATGGTCTTGAGTTGAAGCAGACGGGCCCGGCGCAGATGCCGCAAATATTCTTTGCGGACG
ATCCCGATTTTCGGATCGGCTATGCGTGGGCCGCGGCGTGCCTGAAGGGCGGCGTCTATGTTCATC
CCTATCACAATATGTTTCTCTCTGCGGCCCATACAGTTGACGATGTAACGGAGACCCTCGAGGCGAC
GGATCGCGCGTTCAGCGCGGTCCTCAGAGATTTTGCGTCTCTCCAGCCTCATCCCATTTTAATGCAA
CTCGCCGGTGCTTGA

Appendix 8.3 *FumI* nucleotide sequence.

MANGTRQKDLRERAERVIPGGMYGHESTRLLPPEFPQFFRRALGARIWDADEQPYIDYMCAYGPNLLG
YRQSEIEAAADAQRLLGDTMTGPSEIMVNLAEAFVGMVRHADWAMFCKNGSDATSTAMVLARAHT
GRKTILCAKGAYHGASPWNTPHTAGILASDRVHVAYYTYNDAQSLSDAFKAHDGDIAAVFATPFRHEVF
EDQALAQLEFARTARKCCDETGALLVVDDVRAGFRVARDCSWTHLGIEPDLSCWGKCFANGYPISALLG
SNKARDAARDIFVTGSFWFSAVPMAAAIETLRIIRETPYLETLIASGAALRAGLEAQSQRHGLELKQTGPA
QMPQIFFADDPDFRIGYAWAAACLKGGVYVHPYHNMFLSAAHTVDDVTETLEATDRAFSAVLRDFASL
QPHPILMQLAGA**ASSVDKLAAALEHHHHHH**

Appendix 8.4 Recombinant FumI amino acid sequence. The C-terminal His6-tag is highlighted in b



Appendix 8.5 Calibration curve for the HiLoad 16/600 Superdex 200 column (GE Healthcare) used to estimate the molecular weight of the protein based on its retention volume.

The MW of the protein is estimated as follows:

$$MW = e\left(\frac{Kav-1.7958}{-0.2833}\right) \text{ Where } Kav = \frac{Ve-Vo}{Vt-Vo}$$

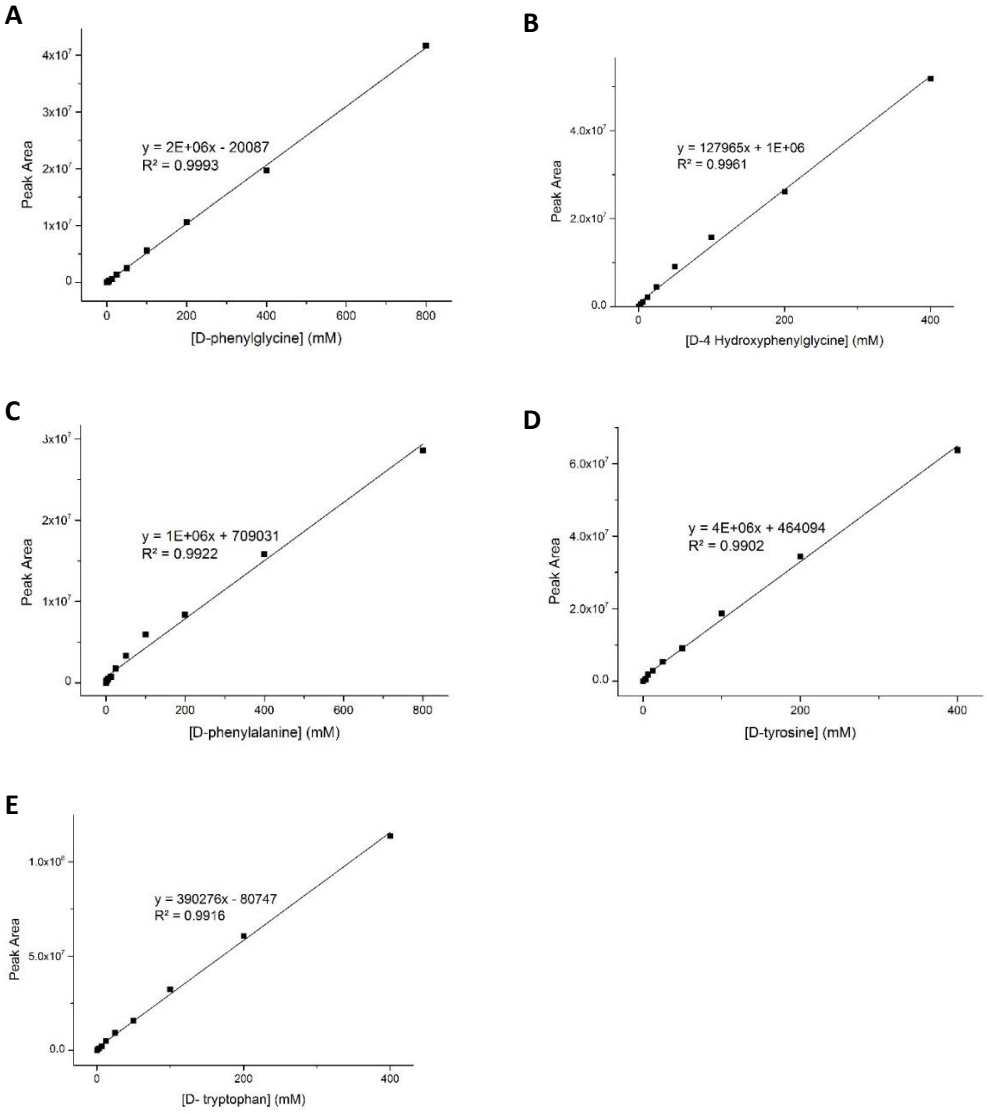
Where:

MW is measured in Da

V_e = elution volume

V_o = Void volume (41.94 mL)

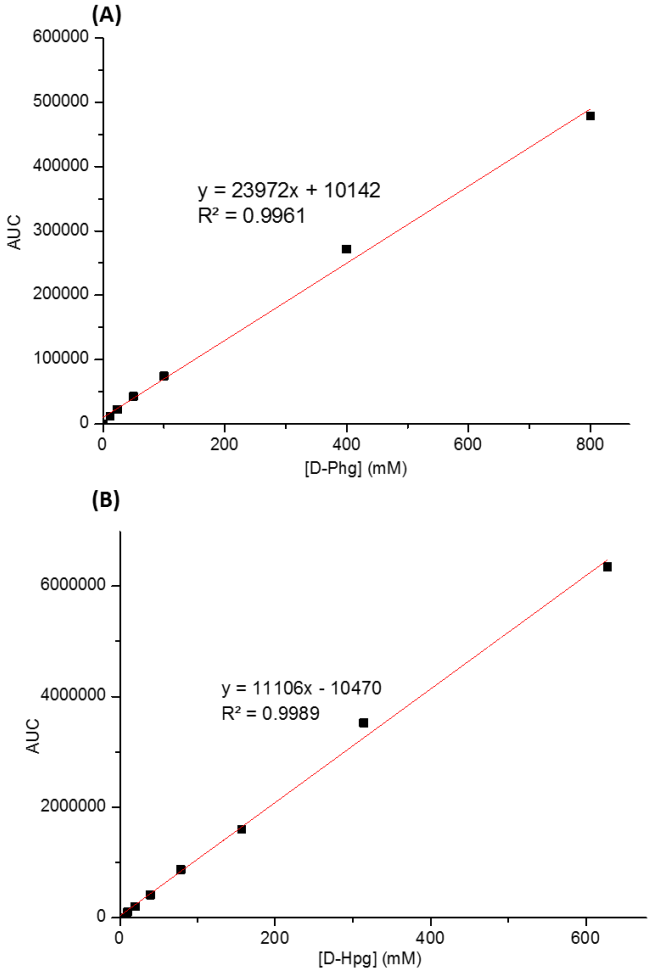
V_t = Total bed volume (120 mL)



Appendix 8.6 Chiral HPLC calibration curves of: **(A)** D-phenylgylcine **(B)** D-4 hydroxyphenylglycine **(C)** D- phenylalanine **(D)** D- tyrosine **(E)** D-tryptophan. Each amino acid concentration (0 – 800 mM) was analysed using the following solvents (0.025% triethylammonium acetate (TEAA):MeOH (50:50). The flow rate was 1 min mL^{-1} with elution monitored at λ = 205 nm. Data was analysed by non-linear regression using Origin.

	D-PhgAT
Data collection	
Wavelength (Å)	0.9795
Resolution range (Å)	49.47 - 2.248 (2.329 - 2.248)
Space group	C 1 2 1
Unit cell (Å)	*a* = 329.28, *b* = 83.90, *c* = 133.42
Unit cell (°)	β = 111.56
Total reflections	60,0728 (56,842)
Unique reflections	160,521 (15,800)
Multiplicity	3.7 (3.6)
Completeness (%)	99.76 (98.19)
Mean I/sigma(I)	12.70 (1.39)
Wilson B-factor (Å^2)	48.16
R_{merge}	0.062 (0.869)
R_{meas}	0.072 (1.018)
R_{pim}	0.037 (0.522)
$CC_{1/2}$	0.998 (0.609)
CC*	1 (0.87)
Diffraction images (DOI)	10.5281/zenodo.1059413
Refinement	
Reflections used in refinement	160,770 (15,797)
Reflections used for R_{free}	8,144 (782)
R_{work}	0.1828 (0.2874)
R_{free}	0.2106 (0.3125)
CC(work)	0.963 (0.794)
CC(free)	0.952 (0.763)
Number of non-hydrogen atoms	20,642
Macromolecules	20,233
solvent	409
Protein residues	2641
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.97
Ramachandran favored (%)	97.13
Ramachandran allowed (%)	2.80
Ramachandran outliers (%)	0.08
Rotamer outliers (%)	2.20
Clashscore	1.92
Average B-factor (Å^2)	60.63
Macromolecules (Å^2)	60.80
Solvent (Å^2)	52.36
Number of TLS groups	32
PDB code	**6G1F**

Appendix 8.7 D-PhgAT X-ray crystallographic data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.



Appendix 8.8 C18 HPLC calibration curves of **(A)** D-phenylgylcine **(B)** D-4 hydroxyphenylglycine. Each amino acid concentration (0 – 800 mM) was analysed using a C18 column. The flow rate was 1min mL^{-1} with elution monitored at λ = 274 nm. Data was analysed by non-linear regression

FumI PLP	
Data collection	
Wavelength (Å)	0.976
Resolution range (Å)	46.79 - 1.65 (1.709 - 1.65)
Space group	P 1 21 1
Unit cell (Å)	63.18 104.419 65.25 90 109.264 90
Unit cell (°)	303657 (28488)
Total reflections	95437 (9514)
Unique reflections	3.2 (3.0)
Multiplicity	99.56 (99.09)
Completeness (%)	12.40 (1.57)
Mean I/sigma(I)	20.90
Wilson B-factor (Å2)	0.05703 (0.6757)
R_{merge}	0.06848 (0.8182)
R_{meas}	0.03749 (0.4555)
R_{pim}	0.999 (0.712)
$CC_{1/2}$	1 (0.912)
Refinement	
Reflections used in refinement	95418 (9512)
Reflections used for R_{free}	4734 (430)
R_{work}	0.1751 (0.3584)
R_{free}	0.2031 (0.3419)
CC(work)	0.970 (0.830)
CC(free)	0.968 (0.866)
Number of non-hydrogen atoms	7041
Macromolecules	6406
solvent	32
Protein residues	603
RMS (bonds) (Å)	826
RMS (angles) (°)	0.013
Ramachandran favored (%)	1.50
Ramachandran allowed (%)	97.08
Ramachandran outliers (%)	2.68
Rotamer outliers (%)	0.24
Clashscore	1.24
Average B-factor ($Å^2$)	1.35
Macromolecules ($Å^2$)	25.98
Solvent ($Å^2$)	25.29
Reflections used in refinement	22.56
Reflections used for R_{free}	33.44
PDB code	**6HBS**

Appendix 8.9 FumI X-ray crystallographic data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.