



University of
Nottingham
UK | CHINA | MALAYSIA

JM Johnson Matthey
Inspiring science, enhancing life

Biocatalytic synthesis of chiral amine building blocks

Christian M. Heckmann

Supervisors:

Prof. Francesca Paradisi

Dr Beatriz Dominguez

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

July 2021

Abstract

Many commercially important molecules, such as agrochemicals and active pharmaceutical ingredients (APIs), contain chiral amines. However, the synthesis of chiral amines by chemical means is often challenging and in particular in the case of aliphatic amines only low enantiomeric excesses (*ees*) are achieved. By using enzymes, high pressure hydrogen, high temperatures, precious metals, and organic solvents can often be avoided.

In the production of chiral amines, lipases are the most important class commercially, catalysing the enantioselective acylation of chiral amines, resulting in a kinetic resolution. However, while their enantio-selectivity is usually excellent when substituents on the α -carbon are sufficiently different in size, it tends to be poor if substituents are similar (e.g. 2-aminobutane). Additionally, while it is possible to recover both enantiomers and racemize and recycle the unwanted enantiomer, a synthesis of just the desired enantiomer from a prochiral precursor, such as a ketone, has advantages. Here, amine transaminases are the most developed class of enzymes, with several examples of enzyme engineering and scale-up in the literature.

While most research with transaminases is focussed on bulky-bulky ketones, in this work the use of transaminases for the synthesis small chiral amines is being explored. Here, a significant limitation of wild-type transaminases proved to be an advantage: the small pocket that typically does not accept substituents significantly larger than a methyl-group allows for excellent enantioselectivity (> 99.5% *ee*) even for very small chiral amines, such as 2-aminobutane, for which a multi-gram scale synthesis in continuous flow is described. However, attempts at engineering a transaminase for the synthesis of 2,2-dimethylhexan-3-amine were less successful, with only traces of activity being observed. With the cyclic prochiral ketone, the enantiomeric outcome of the reaction depended on the reaction conditions (ionic strength and concentration of organic molecules), with *ees* ranging from 70% (*S*) to 19% (*R*).

The discovery of the tetrameric quaternary structure of two (*R*)-selective transaminases (RTA) (from *Aspergillus terreus* and *Thermomyces stellatus*) is also described. Using this information, a rational mutation stabilizing the tetramer was introduced, which resulted in an overall more stable catalyst that could be used at higher substrate concentrations compared to the wild-type.

Finally, a sequential cascade involving transaminases followed by a Buchwald-Hartwig amination (BHA) is described, which allows access to chiral *N*-arylamines without the need for purification of the intermediate. Employing a biphasic water-toluene system and using a 3rd-generation Buchwald precatalyst, the BHA showed excellent formation of the desired amine also in the presence of excess amine donor (in particular with alanine), allowing for the quick generation of diverse libraries of compounds which may be of use during drug discovery.

Acknowledgments

Foremost, I would like to thank Prof. Francesca Paradisi for being incredibly supportive while also giving me complete freedom over my project, allowing me to explore several avenues during my PhD. I would also like to thank my industrial supervisor, Dr Beatriz Dominguez for all the advice during my project, as well as all the people in U260 at Johnson Matthey (JM) (in particular Dr Ahir Pushpanath, Dr Annette Alcasabas, and Dr Serena Bisagni) for giving me advice and for creating a nice working environment during my three-month placement at JM. I would also like to thank JM for their kind supply of transaminases and for trusting me with plasmids for some of their commercial enzymes, significantly speeding up my research. I would also like to thank all the analytical staff at UoN, in particular Ben Pointer-Gleadhill, who frequently came back from the CNL late in the evening to place my samples on the HPLC when I was not allowed to do so myself because of Covid-restrictions. Lastly, I would like to thank all the members (past and present) of the Paradisi group, Dr Rhys Griffith of the Mitchell group, and Dr Freya Taday of the O'Reilly group for advice, help, and a great working environment.

Table of Contents

1	Introduction	1
1.1	Historical developments leading to modern Biocatalysis	1
1.1.1	Early enzymology—demystifying life	1
1.1.2	Enzyme structures and elucidation of mechanisms	3
1.1.3	The DNA revolution.....	8
1.1.4	Directed evolution and the beginnings of modern biocatalysis	13
1.1.5	“Smart” libraries and applications of enzyme engineering	17
1.1.6	Enzyme immobilization and flow chemistry	21
1.2	Amines	27
1.2.1	Chemical synthesis of chiral amines	27
1.2.2	Enzymatic synthesis of amines.....	29
1.3	Aims.....	41
1.4	References	42
2	Materials and Methods.....	59
2.1	Materials	59
2.2	Media	59
2.2.1	LB medium and LB agar.....	59
2.2.2	SOC medium.....	60
2.2.3	ZYP–5052 auto–induction medium ¹	60
2.2.4	TB-lac medium	60
2.3	Preparation of chemically competent cells	60
2.4	Preparation of electrocompetent cells	61
2.5	Transformation	61
2.5.1	Routine heat shock.....	61
2.5.2	Higher efficiency heat shock	61
2.5.3	Electroporation	61
2.6	Molecular biology techniques.....	62
2.6.1	DNA agarose gel electrophoresis	62

2.6.2	Discovery of novel RTA sequences.....	62
2.6.3	Cloning	62
2.6.4	Site-directed mutagenesis	63
2.7	Substrate Docking	63
2.8	Directed Evolution	64
2.8.1	CASTing library generation.....	64
2.8.2	Colony-based screening	64
2.8.3	DWP screening	65
2.9	Enzyme preparation and characterization.....	65
2.9.1	Expression of HEWT	65
2.9.2	Expression of TsRTA, AtRTA, *RTA-43_His, LDH-4, and GDH.....	66
2.9.3	Enzyme purification	66
2.9.4	Enzyme lyophilization	67
2.9.5	Enzyme immobilization	67
2.9.6	Gel filtration	67
2.9.7	SDS-PAGE	68
2.9.8	Stability, activity, and kinetic assays	68
2.9.9	MALDI-TOF MS.....	69
2.10	TsRTA crystallization and data collection.....	69
2.11	Batch biotransformations	72
2.11.1	Substrate scope analysis TsRTA.....	72
2.11.2	Batch Biotransformations chapter 3—IPA method.....	72
2.11.3	Batch Biotransformations chapter 3—Alanine method.....	72
2.11.4	Batch biotransformations chapter 4.....	73
2.12	Flow biotransformations TsRTA.....	73
2.12.1	Operational stability tests	73
2.13	Scaled-up 2-aminobutane synthesis in flow	73
2.14	Synthesis of primary amine standards.....	75
2.14.1	2,2-dimethylhexan-3-oxime	75
2.14.2	2,2-dimethylhexan-3-amine hydrochloride ²⁰	75
2.14.3	2,2-dimethylpentan-3-one	76
2.14.4	2,2-dimethylpentan-3-oxime	76

2.14.5	2,2-dimethylpentan-3-amine hydrochloride ²⁰	77
2.14.6	(S)-tetrahydrothiophene-3-amine ²²	77
2.14.7	(S)-2-phenoxyisopropylamine	77
2.15	Buchwald-Hartwig Amination project.....	78
2.15.1	General method for the sequential cascade to produce N-aryl amines	78
2.15.2	3-benzylaminotoluene hydrochloride	79
2.15.3	Synthesis of all other N-arylamine standards.....	79
2.16	Analytical techniques	83
2.16.1	Fmoc derivatization	83
2.16.2	DNP derivatization	83
2.16.3	Reverse-phase HPLC analysis of conversions	83
2.16.4	Reverse-phase LC-MS analysis of conversions	84
2.16.5	Reverse-phase UPLC analysis of conversions	85
2.16.6	Chiral GC-FID.....	85
2.16.7	Chiral RP-HPLC (primary amines).....	86
2.16.8	Chiral RP-HPLC (N-arylamines)	86
2.17	References	88
3	Transaminases for the synthesis of 2-aminobutane and 2,2-dimethylhexan-3-amine .	91
3.1	2-aminobutane and 2,2-dimethylhexan-3-amine as target molecules	91
3.2	Asymmetric synthesis of 2-aminobutane from butanone	93
3.2.1	Screening of STAs and RTAs against butanone	93
3.2.2	Determination of immobilization and operating conditions in flow	99
3.2.3	Scaled-up synthesis of both enantiomers of 2-aminobutane in flow..	100
3.3	Screening and attempted engineering of transaminases toward the synthesis of 2,2-dimethylhexan-3-amine	106
3.3.1	Screening of STAs and RTAs against 2,2-dimethylhexan-3-one.....	106
3.3.2	Library design	107
3.3.3	Library screening—colony-based screening	111
3.3.4	Screening of additional transaminases	114
3.3.5	Limitations of the screening.....	116
3.4	Conclusions and future work	119
3.5	References	119

4	Further investigations of HEwT wild-type, F18W, and F84W	123
4.1	Comparison of HEwT wild-type, F18W, and F84W variants	123
4.2	Enantioselectivity THF-ketone—Kinetic and thermodynamic effects	125
4.3	Enantioselectivity THF-ketone—Concentration effects.....	127
4.4	Conclusions	130
4.5	References	131
5	An (<i>R</i>)-Selective Transaminase from <i>Thermomyces stellatus</i> : Stabilizing the Tetrameric Form.....	133
5.1	Background	133
5.2	Search for RTA sequences.....	133
5.3	Expression and initial characterization of TsRTA	134
5.4	3D Structural analysis and relation to thermostability	138
5.5	Substrate scope.....	145
5.6	Immobilization of TsRTA	149
5.7	Conclusion.....	152
5.8	References	153
6	Accessing chiral <i>N</i> -aryl amines in a sequential transaminase–Buchwald-Hartwig amination cascade	155
6.1	Initial optimizations using benzylamine as a model substrate	157
6.2	Scale-up and ketone substrate scope	161
6.3	Aryl halide substrate scope.....	163
6.4	Conclusions	165
6.5	References	166
7	Conclusions and Outlook	167

Abbreviations

ACN	Acetonitrile
ADH	Alcohol dehydrogenase
AlaDH	Alanine dehydrogenase
AmDH	Amine dehydrogenase
Amp	Ampicillin
API	Active Pharmaceutical Ingredient
AtRTA	<i>Aspergillus terreus</i> (R)-selective transaminase
BHA	Buchwald-Hartwig amination
BLAST	Basic local alignment search tool
bp	Base pairs
BSA	Bovine serum albumin
CAST	Combinatorial active-site saturation test
carb	Carbenicillin
cfe	Cell free extract
CV	Column volume
CvSTA	<i>Chromobacterium violaceum</i> (S)-selective transaminase
DCM	Dichloromethane
DMF	Dimethylformamide
DMHamine	2,2-dimethylhexan-3-amine
DMHone	2,2-dimethylhexan-3-one
DMPamine	2,2-dimethylpentan-3-amine
DMPone	2,2-dimethylpentan-3-one
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNPH	2,4-Dinitrophenylhydrazine
dNTP	Deoxynucleoside triphosphate
DTT	Dithiothreitol
DWP	Deep-well plate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
<i>ee</i>	Enantiomeric excess
epPCR	error-prone Polymerase Chain Reaction
ESI-MS	Electrospray ionization mass spectrometry
Fmoc	Fluorenylmethyloxycarbonyl
g	Gravitational acceleration, 9.81 m s ⁻²
GC-FID	Gas chromatography–flame ionization detector
GDH	Glucose dehydrogenase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEwT	<i>Halomonas elongata</i> ω-Transaminase
HSA	Human serum albumin
IMAC	Immobilized metal affinity chromatography
IPA	isopropylamine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IREd	Imine reductase
JM	Johnson Matthey
Kan	Kanamycin
LB	Lysogeny Broth

LC-MS	Liquid chromatography – mass spectrometry
LDH	Lactate dehydrogenase
m/z	Mass to charge ratio
MALDI-TOF MS	Matrix-assisted laser desorption/ionization Time of Flight Mass spectrometry
MAO	Monoamine oxidase
MWCO	Molecular weight cut-off
NAD ⁺ /NADH	Nicotinamide adenine dinucleotide (oxidized, reduced)
NMR	Nuclear magnetic resonance
nt	Nucleotides
NZY	N-Z-Amine + Yeast extract
OD _n	Optical Density at n nm
<i>o</i> XDA	<i>o</i> -xylylenediamine
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PLP	Pyridoxal 5'-phosphate
<i>p</i> -NPEA	<i>p</i> -nitrophenethylamine
RedAm	Reductive Aminase
RP-HPLC	Reverse-phase high-performance liquid chromatography
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SMBA/RMBA	(<i>S</i>)/(<i>R</i>)-methylbenzylamine
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
STA/RTA	(<i>S</i>)/(<i>R</i>)-selective transaminase
STY	Space-time yield
Strep	Streptomycin
TA	Transaminase
TAE	Tris-acetic acid-EDTA
TB	Terrific broth
TBAB	Tetrabutylammonium bromide
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
TFB	4,4,4-trifluorobutane
THF	Tetrahydrofuran
THT	Tetrahydrothiophene
TPGS-750-M	DL- α -Tocopherol methoxypolyethylene glycol succinate
Tris	Tris(hydroxymethyl)aminomethane
<i>Ts</i> RTA	<i>Thermomyces stellatus</i> (<i>R</i>)-selective transaminase
U	Unit of enzymatic activity
UV	Ultra-violet
wt	Wild-type
ZYP-AI	N-Z-Amine + Yeast extract + phosphate - Autoinduction
β -ME	β -mercaptoethanol
ϵ	Molar extinction coefficient

Publications

Christian M. Heckmann contributed to the following publications during his PhD. Associated chapters in this thesis in parentheses where applicable.

1. Planchestainer, M., Hegarty, E., Heckmann, C. M., Gourlay, L. J. & Paradisi, F. Widely applicable background depletion step enables transaminase evolution through solid-phase screening. *Chem. Sci.* **10**, 5952–5958 (2019).
2. Heckmann, C. M., Gourlay, L. J., Dominguez, B. & Paradisi, F. An (R)-Selective Transaminase From *Thermomyces stellatus*: Stabilizing the Tetrameric Form. *Front. Bioeng. Biotechnol.* **8**, 707 (2020). (chapter 5)
3. Heckmann, C. M. & Paradisi, F. Looking Back: A Short History of the Discovery of Enzymes and How They Became Powerful Chemical Tools. *ChemCatChem* **12**, 6082–6102 (2020). (chapter 1)
4. Delgado, L., Heckmann, C. M., Di Pisa, F., Gourlay, L. & Paradisi, F. Release of Soybean Isoflavones by Using a β -Glucosidase from *Alicyclobacillus herbarius*. *ChemBioChem* **22**, 1223–1231 (2021).
5. Delgado, L. *et al.* Producing natural vanilla extract from green vanilla beans using a β -glucosidase from *Alicyclobacillus acidiphilus*. *J. Biotechnol.* **329**, 21–28 (2021).
6. Heckmann, C. M., Dominguez, B. & Paradisi, F. Enantio-Complementary Continuous-Flow Synthesis of 2-Aminobutane Using Covalently Immobilized Transaminases. *ACS Sustain. Chem. Eng.* **9**, 4122–4129 (2021). (chapter 3)

1 Introduction

1.1 *Historical developments leading to modern Biocatalysis*

Living systems carry out reactions under mild conditions, enabled by excellent catalysts: enzymes. Humans have, unknowingly at first, used these enzymes to their advantage since prehistory, for example to ferment sugars into alcohol. So how did we get from biocatalysis being used unknowingly to the modern application at the forefront of chemical synthesis?

1.1.1 *Early enzymology—demystifying life*

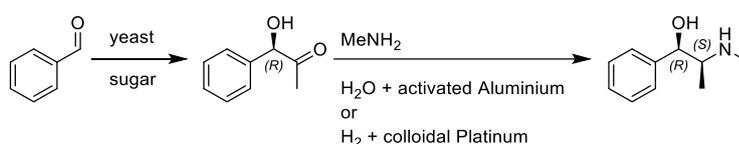
In 1833, diastase (a mixture of amylases) was the first enzyme to be discovered,¹ quickly followed by other hydrolytic enzymes such as pepsin and invertase,² but the term enzyme was only coined in 1877 by Wilhelm Kühne.³ The concept of catalysts (chemicals facilitating a reaction without undergoing any change themselves, was introduced in 1836⁴) by Berzelius who quickly hypothesized that enzymes were such catalysts.⁵ Yeast, which had been observed in ethanolic fermentations, was also viewed as a catalyst, but soon it was discovered that it was a living organism which at the time seemed to contradict that concept.⁶ Evidence from Pasteur that fermentation occurs in the absence of oxygen (Liebig thought that the decay of yeast in its presence was catalysing alcohol from sugar), and failed attempts to isolate an enzyme able to carry out this transformation (including by Pasteur himself), was claimed as evidence by vitalists that a “life-force” was necessary for these more complex transformations and that enzymes only carried out “simple” hydrolysis reactions.⁷ Indeed, this is often framed as a dispute between Pasteur and Liebig, with the former supporting vitalism and the latter supporting a mechanistic view that ascribes no special place to life. However, it appears more accurate to say that Pasteur supported the idea that fermentation was carried out by yeast through chemical means whereas Liebig opposed the idea of any causal link between yeast as a living organism and the catalytic fermentation reaction.^{7,8} Finally, in 1897, Eduard Büchner showed that a dead yeast extract could carry out the same fermentation reaction as living yeast, thus dealing the final blow to vitalism, which had already been on the decline (Nobel Prize in Chemistry 1907).^{2,7}

The fermentation of sugars into ethanol and carbon dioxide was attributed to “zymase.” Further investigations started to reveal reaction intermediates, and dependency on phosphate and “co-zymase” (A. Harden and H. von Euler-Chelpin; Nobel Prize in Chemistry 1929) and started to untangle glycolysis. However, the chemical nature of

enzymes was still being debated. In 1926, James B. Sumner crystalized the first enzyme (urease), and confirmed it was a protein.⁹ John H. Northrop also crystalized several other proteins, amongst them pepsin, trypsin, and chymotrypsin.⁵ They were awarded the Nobel Prize in Chemistry in 1946; in his Nobel lecture,⁹ Sumner remarks that

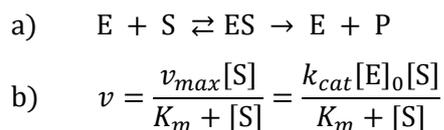
“The organic chemist has never been able to synthesize cane sugar, but by using enzymes, the biological chemist can synthesize not only cane sugar but also gum dextran, gum levan, starch and glycogen.”

Indeed, a whole range of industrial applications of mainly whole organisms but also some enzyme preparations had already been developed. For example, glycerol was produced on a 1000 ton per month scale in Germany during world war I, employing fermentation in yeast with the final acetaldehyde reduction step inhibited by sulfite, resulting in dihydroxyacetone phosphate reduction. By 1949, citric acid was almost exclusively produced using the fungus *Aspergillus niger* (ca. 26,000,000 pounds per year in the US alone), even though it was not understood how the organism produced it.^{10,11} In 1934, a patent was granted for the condensation of acetaldehyde (produced in-situ from glucose) with benzaldehyde catalysed by whole yeast, giving L-phenylacetylcarbinol, which was then further reacted to give L-ephedrine (**Scheme 1-1**).¹² This procedure is still used today.¹³ Enzymes prepared from mould or bacteria became alternatives to those initially obtained from plants or animals (e.g. amylases and proteases). Purified proteases were used to clarify beer since 1911, pectinases (from various fungi or malt) were used to clarify juices and wine.¹¹



Scheme 1-1: Benzoin-type addition of acetaldehyde to benzaldehyde (now known to be catalysed by pyruvate decarboxylase),¹⁴ followed by reductive amination with methylamine.¹²

By 1949 a vast number of enzyme classes had been discovered and characterized extensively. Many pathways and intermediates were fully uncovered, yet little was known with regard to the mechanism by which individual enzymes worked.¹¹ Through the famous lock-and-key model, proposed by Emil Fischer in 1894,¹⁵ as well as the Michaelis-Menten model of enzyme kinetics from 1913 (**Equation 1-1**)¹⁶ it was understood that a substrate has to bind to the enzyme prior to catalysis, yet how this binding proceeds and how catalysis occurs afterwards was unsolved.



Equation 1-1: The Michaelis-Menten model and equation, in modern form: a) Enzyme and substrate combine in a reversible fashion to form the enzyme-substrate complex, which then goes on to release the enzyme and product in an irreversible reaction. b) Under the assumption of a steady state concentration of the enzyme-substrate complex, the Michaelis-Menten equation can be written, describing the consumption of substrate depending on substrate concentration $[\text{S}]$, maximum velocity v_{max} (itself dependent on enzyme concentration $[\text{E}]_0$), and substrate affinity K_m .

1.1.2 Enzyme structures and elucidation of mechanisms

In 1948 Linus Pauling proposed that enzymes had to stabilize the transition state rather than the substrate as proposed by Fischer.¹⁷ The detailed concept of a transition state itself had only been developed less than two decades earlier.¹⁸ This was further refined by Koshland in 1958,^{19,20} proposing the concept of “induced-fit,” explaining the specificity of enzymes on an abstract level. In parallel to this abstract understanding, a more detailed understanding of the structure of proteins was being developed. It had been hypothesised since the beginning of the 20th century that proteins were composed of chains of amino acids connected via amide bonds;²¹ however, the order or even the relative amount of amino acids was not well understood, and indeed the peptide hypothesis itself was frequently questioned.²² In 1951, Sanger determined the amino acid sequence (referred to as primary structure) of insulin, revealing that indeed as expected it was a well-defined sequence of amino acids linked by amide bonds.^{23,24} He received his first Nobel prize in Chemistry in 1958 for this work.

Famously, Linus Pauling proposed how a chain of amino acids might fold into regular geometric features (i.e. α -helices and β -sheets; referred to as secondary structure, **Figure 1-1**) while sick in bed, based on his detailed understanding of the rigidity of the amide bond and “reasonable” interatomic distances. The rigorous understanding of chemical bonds had just been developed for which Pauling received his Nobel Prize in Chemistry in 1954.^{25–28}

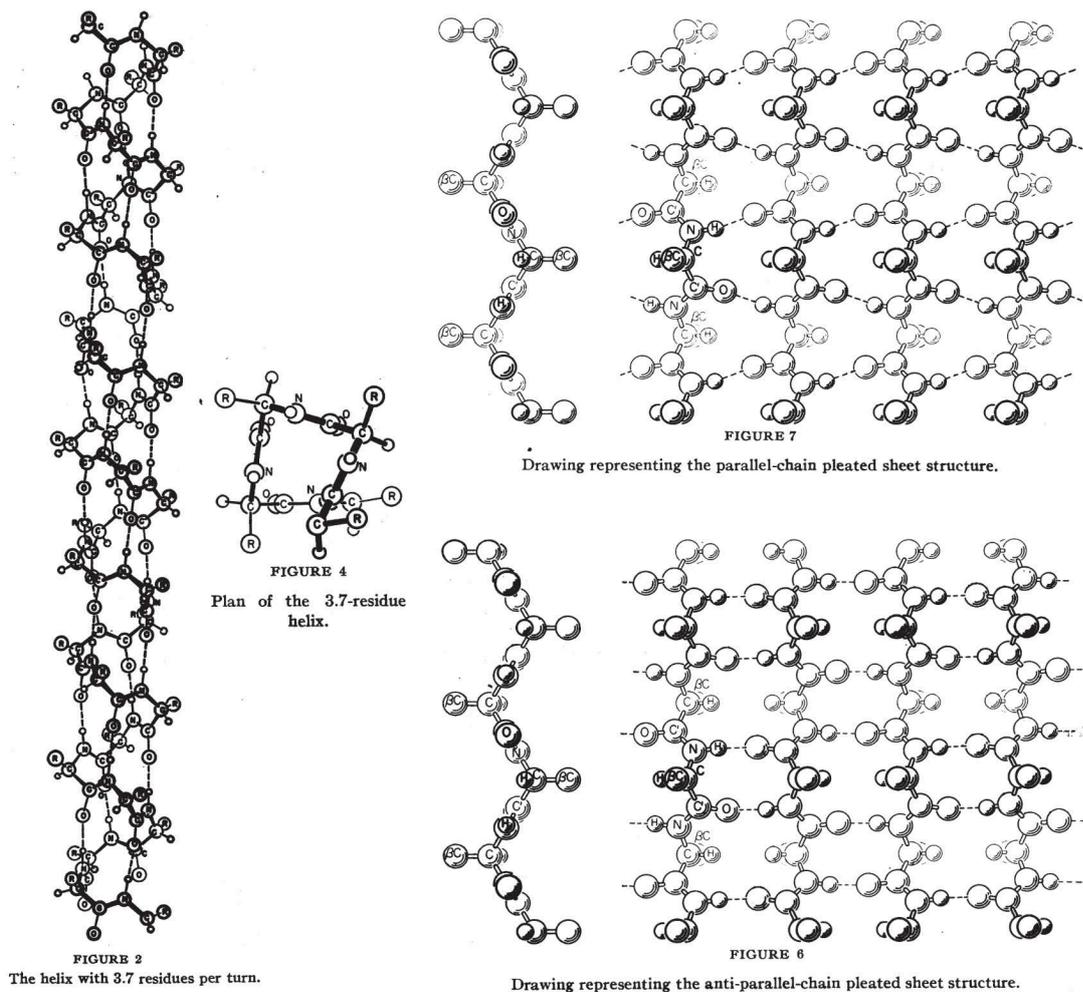


Figure 1-1: The original drawings of an α -helix (left) and parallel and anti-parallel β -sheets (right), published by Pauling in 1951, copyright the original authors.^{26,28}

In the meantime, x-ray crystallography, developed from 1912 onward by Max von Laue (Nobel prize in Physics in 1914) and William and Lawrence Bragg (Nobel prize in Physics in 1915), had become more sophisticated and was being applied to increasingly complex compounds. Evidence for Pauling's secondary structure from x-ray diffraction was reported by Max Perutz in 1951.²⁹ The first structures of proteins were solved in 1958-1960 by John Kendrew and Max Perutz (Nobel prize in Chemistry in 1962).³⁰⁻³⁴ This was initially met with some degree of disappointment as it revealed that proteins were "messy" (**Figure 1-2**) and squashed the hope that solving the structure of one protein would reveal the structure of all proteins (in contrast to DNA where that expectation largely held true).^{35,36} However, as higher resolution structures were obtained the insight that could be gained into the mystery world of enzymes became apparent and many groups set forth to investigate not just proteins but enzymes (**Figure 1-2**).³⁷

Structures of lysozyme were solved in 1965,^{38,39} and included structures of the enzyme with inhibitors bound to it, revealing the location and residues of its active site. Other enzyme structures solved around this time include bovine carboxypeptidase A in 1967,⁴⁰ both with and without substrate bound—revealing conformational changes (in agreement with the induced fit hypothesis) as well as key interactions between substrate and enzyme. The crystal structure of chymotrypsin (also in 1967)^{41–44} paved the way to uncover the classic catalytic triad and oxyanion hole of proteases (as well as esterases and other hydrolytic enzymes; **Figure 1-2**). In 1971, the Protein Data Bank (PDB) was founded with 7 structures,³⁷ reaching 50 structures in 1979, and 100 structures 3 years later. At the end of 2019 it contained almost 160,000 structures.⁴⁵

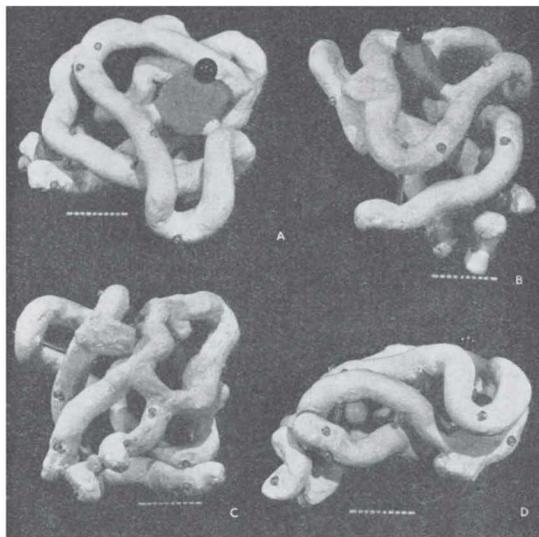
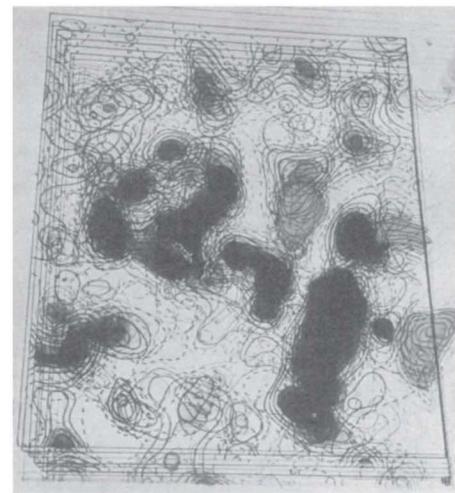
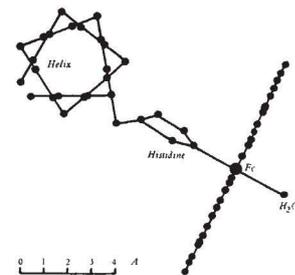


Fig. 2. Photographs of a model of the myoglobin molecule. Polypeptide chains are white; the grey disk is the heme group. The three spheres show positions at which heavy atoms were attached to the molecule (black: Hg of *p*-chloro-mercuri-benzene-sulphonate; dark grey: Hg of mercury diammine; light grey: Au of auri-chloride). The marks on the scale are 1 Å apart



(a)



(b)

Fig. 3. (a) Photograph of a set of sections normal to the plane of the heme groups showing, from left to right, a helix in cross-section, the histidine residue nearly edge-on, the heme group edge-on, and a presumed water molecule. (b) Sketch showing the atomic arrangement in Fig. 3a

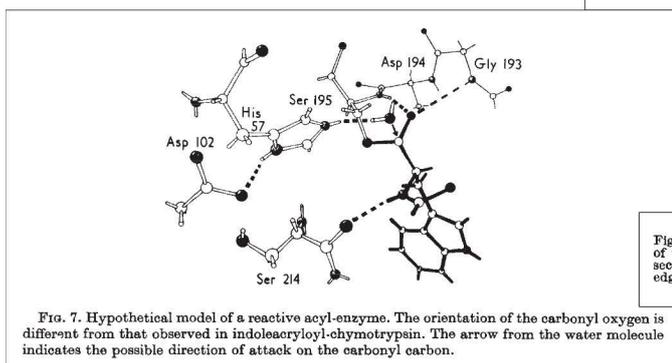


Fig. 7. Hypothetical model of a reactive acyl-enzyme. The orientation of the carbonyl oxygen is different from that observed in indoleacryloyl-chymotrypsin. The arrow from the water molecule indicates the possible direction of attack on the carbonyl carbon.

Figure 1-2: Top left: Clay model of the first x-ray structure of a protein, myoglobin, at 6 Å resolution. Reprinted with permission from Springer Nature (ref.³⁰), copyright 1958. Right: electron density sections of myoglobin at 2 Å resolution and sketch of groups coordinated to iron. Reprinted with permission from Springer Nature (ref.³¹), copyright 1960. Bottom left: Model of the catalytic triad and oxyanion hole of chymotrypsin, as inferred from crystal-structures. Reprinted from ref.,⁴³ copyright 1970, with permission from Elsevier.

In parallel to the increasing understanding of the structure of enzymes, newly developed physical-chemistry techniques were also employed to elucidate mechanisms, such as detailed kinetics, isotopic labelling, isotope effects, and spectroscopic techniques.^{46,47} The first mechanisms to be elucidated in that way were of enzymes employing co-enzymes, as the structures (fragments) of co-enzymes were determined before the structures of whole proteins. Indeed, as early as 1936,⁴⁸ Otto Warburg showed that certain pyridines (analogous to nicotinamide that could be obtained from hydrolysis “co-zymase”) could transfer hydrides reversibly, implying that such a hydride transfer plays a role during glycolysis (he had previously received the Nobel Prize in physiology or medicine in 1931 for his work on the role of iron in respiration).

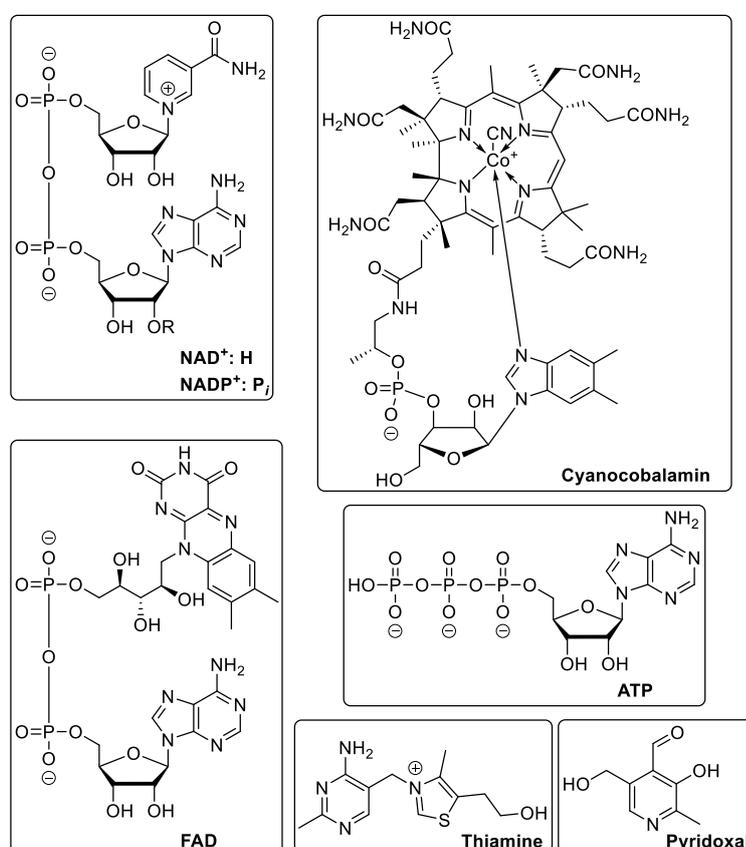


Figure 1-3: Selected structures of common cofactors that were known by 1955: NAD⁺, NADP⁺ and FAD are redox catalysts, ATP transfers energy released during glycolysis, thiamine is the co-factor of pyruvate decarboxylase during fermentation, and pyridoxal is the co-factor of transaminases.

The full structure of thiamine (cocarboxylase) had been proved in 1936 by Williams and Cline.⁴⁹ The structures of pyridoxine, as well as of biologically relevant derivatives pyridoxal and pyridoxamine were established in the early 1940s by Esmond Snell soon after the discovery of transaminases (also see **Section 1.2.2.1**).^{50–54} Full structures of NAD(P)(H) (co-zymase),^{55–57} ATP,^{58,59} and FAD^{60,61} were proved by Alexander Todd in the late 1940s and

50s (Nobel Prize in Chemistry in 1957). The structure of Vitamin B₁₂ (cyanocobalamin) was solved through x-ray crystallography by Dorothy Hodgkin in 1955^{62,63} (**Figure 1-3**; Nobel Prize in Chemistry in 1964).

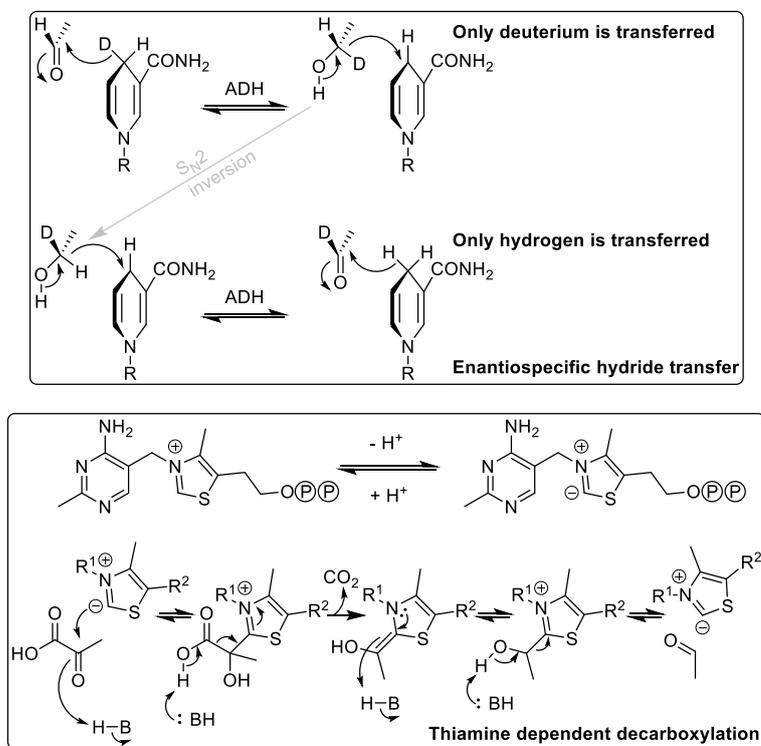
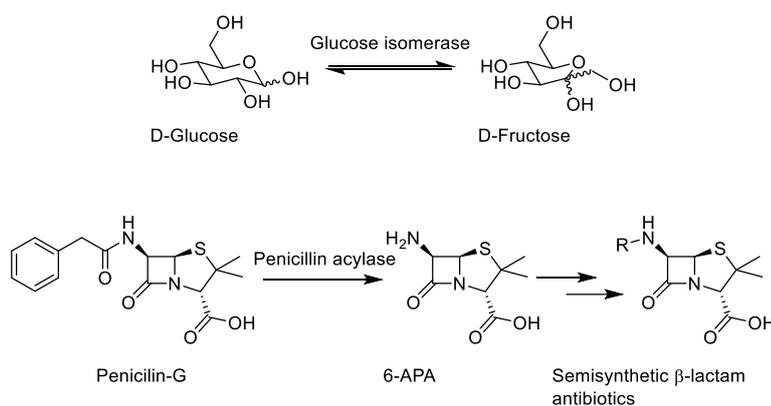


Figure 1-4: Selected mechanisms of co-factors that were being elucidated in the 1950s: enantiospecificity during hydride transfer from NAD(P)H in alcohol dehydrogenases, and thiamine dependent decarboxylation.⁴⁷

The chemistries of those co-factors could be investigated in the absence as well as in the presence of their enzymes, and from this, mechanistic details could be inferred. In addition, structural analogues could be synthesized and their reactivities compared. For example, careful isotope labelling studies in the early 1950s revealed that one hydride of the pyridine ring of NAD(P)H was transferred during reduction/oxidation in a stereospecific manner, giving additional detail to Otto Warburg's mechanism (**Figure 1-4**).^{64,65} In 1957, Breslow showed by NMR that an anion in position 2 of a thiazolium ring could exist, revealing the reactive centre of thiamine (**Figure 1-4**).^{66,67} The observation that pyridoxal, the co-factor of transaminases, as well as structural analogues with electron withdrawing groups on the aromatic ring, can catalyse transamination in the absence of the enzyme allowed Alexander Braunstein and Esmond Snell to postulate independently a likely catalytic cycle in 1954, which later proved to be correct (**Section 1.2.2.1, Figure 1-17**).^{47,51,68-70}

However, these advances in the knowledge of how enzymes work had no immediate impact on industrial biocatalysis, which was largely limited by the low quantities most

enzymes could be obtained in. Major developments at the time include the application of glucose isomerase for the production of high fructose corn syrup and the development of a penicillin acylase process for the production of 6-aminopenicillanic acid (6-APA, at the time obtained from chemical cleavage of penicillin-G), a building block for semi-synthetic antibiotics such as ampicillin and amoxicillin (**Scheme 1-2**).⁷¹ Key for the success of both applications was the discovery of the possibility to immobilize proteins with retention of their function as discovered in the 1950s.^{72–77} This allowed the enzymes to be recycled and used in a continuous fashion, reducing cost by reducing the quantity of enzyme that has to be isolated. The HFCS processes became wide-spread in the 1970s.⁷⁸ However, the production of 6-APA via chemical hydrolysis predominated until the early 1990s, at least partially due to the difficulty of obtaining sufficient quantities of penicillin acylase before then.⁷⁹ A notable exception is Bayer, who used an immobilized penicillin acylase since 1972 as a closely guarded secret, employing *E. coli* strains that achieved a penicillin acylase content of ca. 20%.⁸⁰ Processes to synthesise amino acids using immobilized enzymes (as well as whole cells) were being commercialized in Japan from 1973.⁸¹



Scheme 1-2: Top: Isomerization of D-glucose to D-fructose, catalysed by an immobilized glucose isomerase as used in the production of HFCS. Bottom: Hydrolysis of Penicillin-G to give 6-APA, which can then be acylated to give several semi-synthetic antibiotics.

1.1.3 The DNA revolution

In order for enzymes to enjoy more widespread use, their production had to be ramped up dramatically. In addition, to apply the insights into enzyme mechanisms described above, enzyme active sites had to be tweaked somehow. The key to both these problems was the understanding of how DNA encodes proteins, as well as the development of efficient ways of manipulating DNA.⁸² Of course, in parallel to the research into proteins described above, research into DNA was also ongoing. While DNA was originally viewed as less important than proteins (due to its simple make-up of four building blocks), this view quickly changed with the discovery that it was the carrier of hereditary information by Avery in

1944.⁸³ Of course, the correct structure of DNA was postulated in 1953 by Watson and Crick⁸⁴ (Nobel Prize in Physiology or Medicine in 1962) from x-ray diffraction data by Rosalind Franklin.

This quickly led to a postulation of how genetic information is encoded in DNA: the hypothesis that DNA encodes amino acid sequences³⁴ and that the amino acid sequence alone determines the structure of proteins, as demonstrated by Anfinsen in 1961⁸⁵ (Nobel Prize in Chemistry in 1972). The process of transcription of DNA into mRNA and the translation of mRNA into protein itself was broadly solved within a decade of the discovery of the structure of DNA; transcription as a concept was proposed by François Jacob and Jacques Monod in 1961⁸⁶ (Nobel Prize in Physiology or medicine in 1965). In the same year, mRNA was discovered,^{87,88} the triplet code was established,⁸⁹ and the first codon (UUU) was solved by Nirenberg.⁹⁰ In competition with several other groups,^{91–96} amongst them Gobind Khorana developing a sequence specific chemical synthesis of polynucleotides, the genetic code (**Figure 1-5**) was fully solved by 1966 and its universality established by 1967⁹⁷ (Nirenberg, Khorana and Robert W. Holley (for the isolation of tRNA) received the Nobel Prize in Physiology or Medicine in 1968).

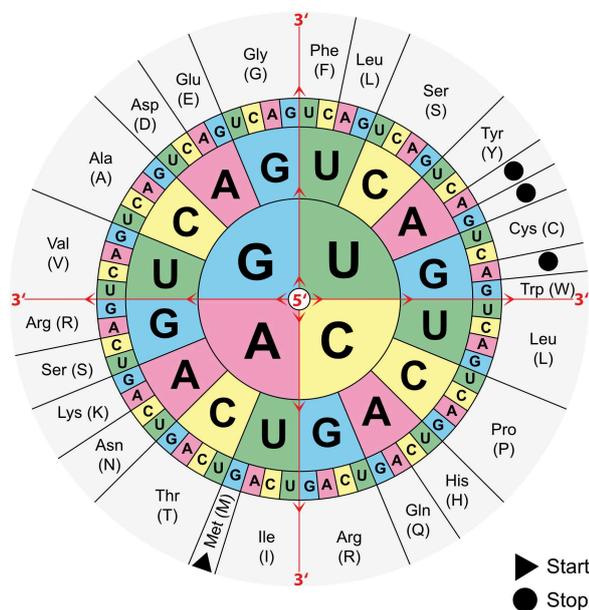


Figure 1-5: The genetic code. https://commons.wikimedia.org/wiki/File:Aminoacids_table.svg, public domain.

Understanding the meaning of the genetic code of course is of limited use, unless one can also read the DNA sequence. Frederick Sanger developed an ingenious hi-jacking of normal DNA replication in 1977 (**Figure 1-6**):⁹⁸ by supplying a small quantity of nucleotides that could not be further extended (because they missed the 3'-OH), strands of DNA

truncated after every A (or C, G, or T; depending on which was supplied as the dideoxynucleotide) were produced, which could be separated by size using gel electrophoresis. Repeating this experiment for all 4 nucleotides, the sequence of bases in the template could be deduced. This earned Frederick Sanger his second Noble Prize in Chemistry in 1980. While initially DNA was visualized using radioactive labels, this was quickly replaced by using fluorescently labelled dideoxynucleotides, allowing for all four bases to be present in the same reaction mixture, increasing throughput. Using capillary electrophoresis, automated sequencing became possible.^{99,100}

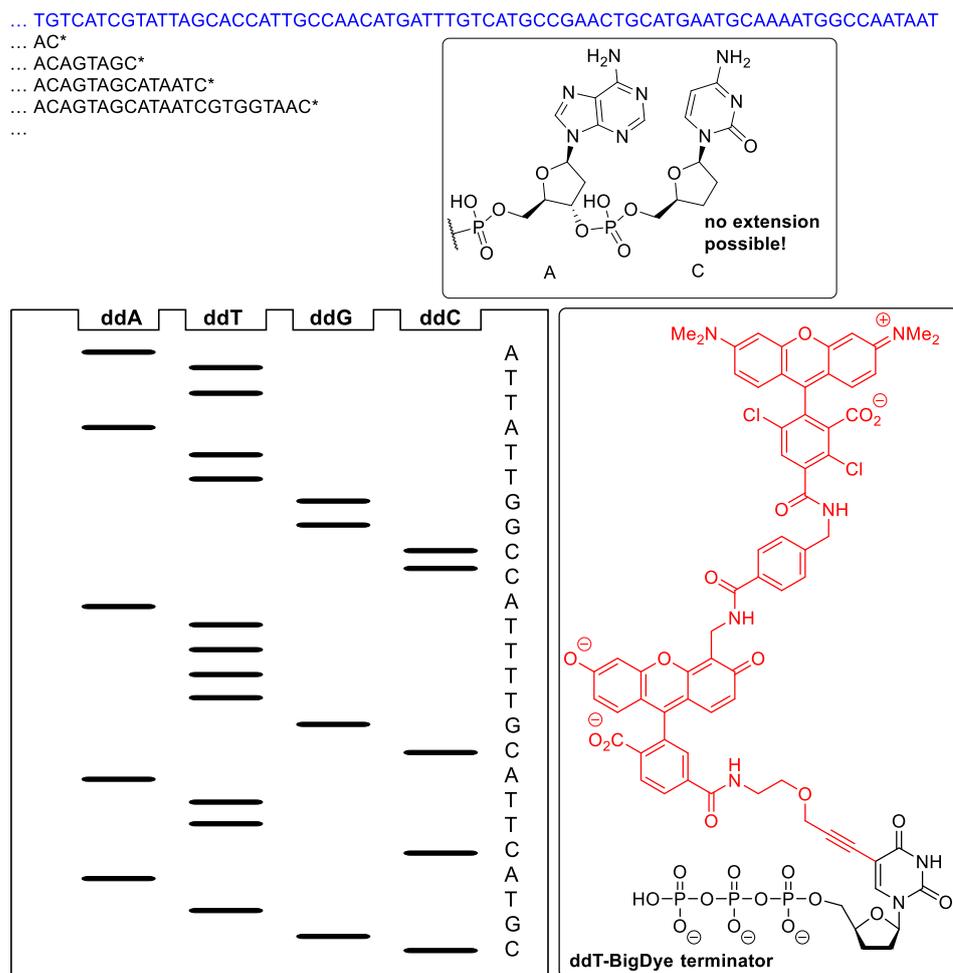


Figure 1-6: Principle of Sanger sequencing: A DNA strand (blue) is copied by DNA polymerase. If a small quantity of dideoxynucleotides (ddNTP) is offered in addition to deoxynucleotides (for example ddCTP), chains will terminate whenever a ddCTP is incorporated instead of a dCTP, as the 3'-OH needed for chain-extension is missing. If this experiment is repeated for all four nucleotides, and the products are separated by size, the sequence of the DNA template can be inferred. Modern Sanger sequencing includes all four ddNTPs in a single sequencing reaction, and distinguishes incorporation of the different bases at the termination site via fluorescent labels, such as the label (red) for the ddT—BigDye terminator shown.^{99,100}

At the same time and leading on from Avery's experiment, the transmission of genetic information in bacteria was being investigated. In 1952, Joshua Lederberg coined the term "plasmid" to describe such transmissible DNA and discovered the nature of its transmission

(Nobel Prize in Physiology or Medicine in 1958).^{101,102} Also in 1952, Salvador Luria¹⁰³ and Giuseppe Bertani¹⁰⁴ observed that bacteriophages from one strain of *E. coli* have a decreased virulence in another, but upon growth in the second strain would show increased virulence for it and a decreased virulence for the original strain, observing the effect of restriction enzymes (so called because they restrict the growth of bacteriophage). This effect was then also observed for several other bacteria. However, it was only in the early 1960s that the nature of these enzymes as site specific endonucleases, and that host bacteria protect their own DNA through modification (methylation), was proposed by Werner Arber.^{105–107} Mathew Meselson isolated the first such restriction endonuclease in 1968 (EcoKI).^{108,109} These early restriction enzymes recognized a specific sequence but did not cut in a specific location, and are now known as type I restriction enzymes. Type II restriction enzymes, cutting DNA in specific locations (**Figure 1-7**), were discovered by Hamilton Smith in 1970 (HindII and HindIII).^{110,111} In conjunction with gel electrophoresis, this allowed the digestion of DNA into fragments of defined size which could then be separated, as shown by Daniel Nathans in 1971.^{112,113} Arber, Smith, and Nathans won the Nobel Prize in Physiology or Medicine in 1978.⁸²

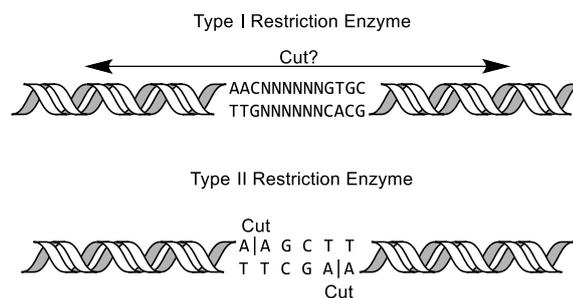


Figure 1-7: Type I restriction enzymes cut at a non-defined remote location from the recognition site (example of EcoKI). Type II restriction enzymes cut at a well-defined position within (or close to) the recognition sequence, often in a staggered way, producing cohesive ends (example of HindIII). Figure and caption reproduced from ref.⁸²

As many type II restriction enzymes produce palindromic single stranded overhangs, it was then realized that DNA from different sources could be stitched together if cut with the same restriction enzyme, through the action of DNA ligase. The first such “recombinant” DNA was reported by Paul Berg in 1972 (Nobel Prize in Chemistry in 1980).^{114–117} It thus became possible to introduce any piece of DNA from any organism into (for example) *E. coli*.¹¹⁸ Thus, plasmids for convenient introduction of such recombinant DNA were being developed in the 1970s.¹¹⁶ One of the most famous of these plasmids is pBR322, developed by Bolivar and Rodriguez (BR) in 1977.^{116,119,120} This plasmid made use of two antibiotic resistance genes and several unique restriction sites within them to allow for the selection of colonies that had a) up-taken the plasmid and b) up-taken a plasmid containing an insert

(**Figure 1-8**). The propagation of recombinant DNA in a new host is referred to as cloning. The pUC series of plasmids (UC for University of California), derived from pBR322, allowed for colourimetric detection of inserts.^{121,122} Finally, the pET series of vectors, also derived from pBR322, was created in the late 1980s and included a T7 promoter, allowing for the selective expression of the DNA insert (ET for Expression by T7 RNA Polymerase).^{123,124} Strains of *E. coli* were generated containing the gene for the T7 polymerase, under the control of a modified *lac* promoter (*lacUV5*), as a lysogen of the DE3 phage.^{82,124–127} Alternative promoter systems were also developed, such as the aforementioned *lac* promoter, as well as the *trc* promoter, *pL* promoter, and *tetA* promoter, and more, each with their own advantages and draw-backs.^{128,129}

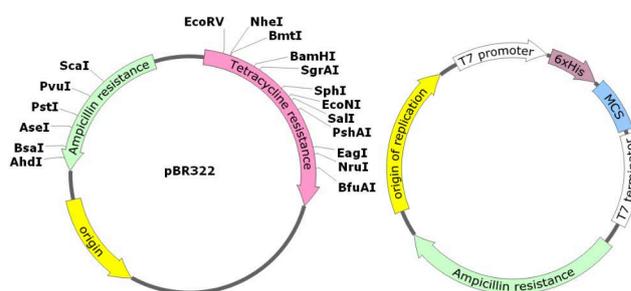
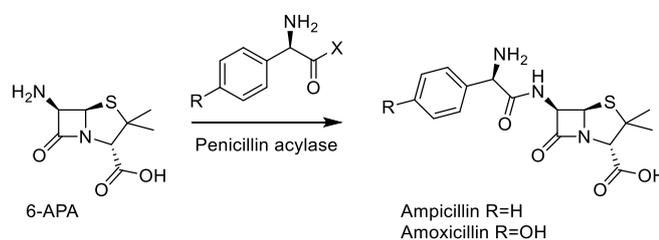


Figure 1-8: Left: Map of pBR322, showing the unique restriction sites inside both antibiotic resistance genes. Right: Generic map of a typical (empty) expression vector, having an origin of replication (for replication in vivo), a selective marker (Ampicillin resistance in this instance), and the T7 promoter and terminator flanking a His-tag (to allow purification of the insert) and the multiple cloning site, which contains a large number of unique restriction sites (not shown) for easy cloning. Figure and caption reproduced from ref.⁸²

This revolutionized enzymology and biocatalysis. Finally, the DNA sequence of an enzyme of interest could be determined and cloned, and the enzyme could be over-expressed in *E. coli* (or another suitable organism) and thus be obtained in sufficient quantities to be studied and used in industrial applications. The first recombinant protein produced was insulin in 1978, and the commercial production of human insulin assumed in 1982.¹³⁰ Prior to that, insulin had to be isolated from pigs or cows and often had limited and inconsistent efficacy, as well as inconsistent supply.¹³¹ This technology also allowed penicillin acylase to be obtained in sufficient quantities and enabled its widespread application toward the synthesis of 6-APA, as mentioned above.^{76,80} Indeed, penicillin acylase was one of the first enzymes expressed recombinantly, in 1979, only one year after insulin.^{75,80} The availability of this enzyme also enabled the development of its application in the reverse direction, catalysing the amide bond formed between 6-APA and the side chains found in semi-synthetic antibiotics such as amoxicillin and ampicillin (**Scheme 1-3**, see **Section 1.1.6** for a discussion of the role of immobilization).^{71,79,132} Around the same time, recombinant

chymosin started replacing natural rennet, obtained from calf stomachs, in cheese-making.^{133–136} This provided a cheaper, more stable supply for cheesemaking as well as more consistent results due to a higher purity. By 2006, up to 80% of all rennet was recombinant chymosin and cheese production in the US had increased over two-fold.¹³⁷



Scheme 1-3: Application of penicillin acylase for the synthesis of amoxicillin and ampicillin from 6-APA. X = NH₂ or OMe.

1.1.4 Directed evolution and the beginnings of modern biocatalysis

As great as natural enzymes are at carrying out their function, they often present drawbacks which make them unsuitable for industrial applications such as their lack of stability, (co-)solvent-tolerance, or a very limited substrate scope. While immobilization can address the stability problems (see **Section 1.1.6**),¹³⁸ it quickly became desirable to be able to change the properties of the enzymes themselves. Of course, to some extent this had already been done routinely, through strain optimization, since the late 1920s. Whole organisms were subjected to mutation-inducing conditions, such as radiation or chemical agents, and the resulting strains were screened for favourable phenotypes.^{139–143} Through this method, strains producing larger quantities of desirable products, either specific enzymes (such as in the case of penicillin acylase at Bayer mentioned above) or chemicals could be obtained, and entirely new pathways could be introduced.^{144–146} However, these approaches were slow, unlikely to directly change the properties of any specific enzyme, and could only really be applied to organisms with sufficiently short replication-cycles. However, with the availability of recombinant DNA, as well as the understanding of enzymes and their mechanisms as outlined above, introduction of specific mutations into a single target enzyme was now within reach, irrespective of the organism it originated from. Indeed, a general method for site-directed mutagenesis was being reported by Michael Smith in 1978 (Nobel Prize in Chemistry in 1993), the same year as the cloning of insulin was achieved. By designing DNA primers, harbouring the desired mutations, complimentary to the target sequence to be mutated, and extending with DNA polymerase using the target sequence as a template, copies containing the mutation could be made (**Figure 1-9**).¹⁴⁷ Of course, efficient syntheses

of specific DNA sequences such as those developed by Khorana,¹⁴⁸ Gillam,¹⁴⁹ and Caruthers¹⁵⁰ were instrumental for this.¹⁵¹

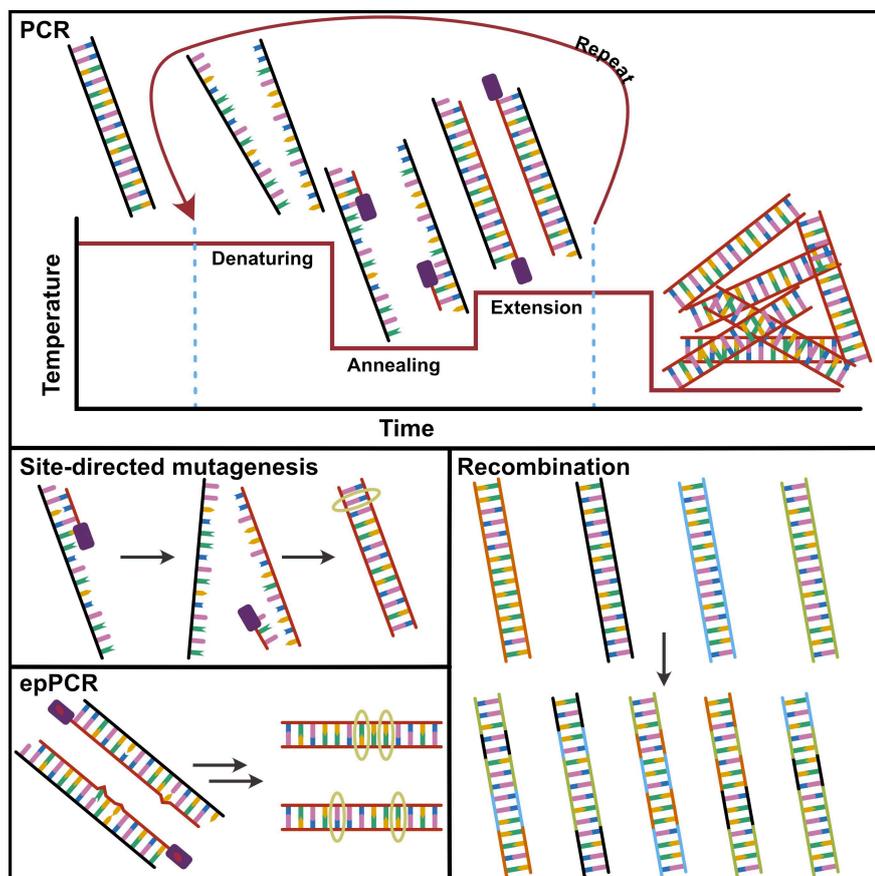


Figure 1-9: The principles of PCR: DNA is denatured at high temperature, primers supplied in the reaction mixture are annealed, and the template is copied. Repeated cycles exponentially amplify the target sequence. Site directed mutagenesis: a mutation is incorporated in the primer; the amplified product now contains the changed base-pair. epPCR: a polymerase that occasionally incorporates incorrect nucleotides is used. The product now contains a set of different sequences that differ from the parent in a few positions. Recombination: several sequences are shuffled to produce a diverse set of new sequences from the parents.

Using this approach, the role of catalytic residues could be directly investigated. For example changing the cysteine in the catalytic triad found in tyrosyl tRNA synthase to a serine (as found in esterases mentioned above) greatly reduces the efficiency of the enzyme.¹⁵² Indeed, the role of several residues in several enzymes could now be quantified, confirming and in some cases revising mechanisms that had been postulated based on crystallography.¹⁵³ However, using this technique to introduce desirable properties into enzymes was quickly met with the realization that the effect of mutations was often unpredictable, and rational engineering of enzymes was often not successful. Thus, a shift was made toward a more random approach such as the use of site-saturation mutagenesis, where targeted residues were changed to all possible amino acids rather than a specific one. Through this, some progress was made such as the introduction of a stabilizing mutation into subtilisin, a

protease with application in laundry detergents,^{138,153,154} or enhanced thermostability of glucose isomerase.¹⁵⁵

With the development of the polymerase chain reaction (PCR, **Figure 1-9**) in the 1980s by Kary Mullis (Nobel Prize in Chemistry in 1993), it became possible to produce large numbers of copies of DNA sequences from a single template.^{156–158} By modulating the fidelity of the polymerase, random mutations could be introduced into the amplified product (error-prone or epPCR, **Figure 1-9**). In the early 1990s, Frances Arnold used this technique to create large libraries of mutants to which she then applied evolutionary pressure. In her own words,¹⁵⁹ she

“rejected microbial growth or survival selections favored by microbiologists and geneticists. Thus we turned to good old-fashioned analytical chemistry to develop reproducible, reliable screens that reported what mattered to us.”

In doing so she managed to produce a variant of subtilisin E that could tolerate high concentrations of DMF, introducing a total of 10 mutations.^{160,161} Thus, the field of directed evolution was born. In 1994, Pim Stemmer introduced the concept of DNA reshuffling (**Figure 1-9**), mimicking DNA recombination which occurs in organisms as a way to increase genetic diversity, and applying it to recombinant DNA *in vitro*. Without being restricted to genes from a single species, very diverse proteins could be mixed together to create new sequences very distant from natural ones.^{162,163} This technique proved very powerful on its own, but especially when combined with epPCR, allowing the combination of mutations from several mutants without the need for assumptions on which mutations would be additive.¹⁶⁴ Frances Arnold received the Nobel Prize in Chemistry in 2018 for this work. In addition to co-solvent tolerance, directed evolution was quickly used to create enzymes with improved thermostability,^{165–168} pH stability,¹⁶⁷ as well as enhanced activity at low temperatures,^{167,169} activity toward unnatural substrates,^{164,170,171} modified enantioselectivity,^{172,173} or combinations of the above. Thus, it quickly established itself as a powerful tool in protein engineering across structurally and functionally diverse classes of enzymes. It was also quickly realized that beneficial mutations were often found in unexpected parts of the enzymes, explaining why early rational attempts struggled at accomplishing these modifications.^{138,159,174–178}

The sudden availability of biocatalysts with properties suitable for industrial applications, as well as the ability to create those properties at will, made them very

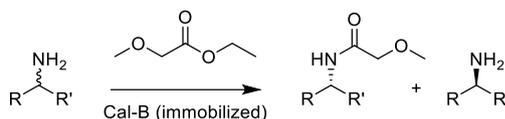
attractive for use in synthetic applications that so far had been considered out of reach.^{179,180} Of course, catalysis itself had become a major field of interest in synthetic chemistry in the second half of the 20th century, as the environmental impact of traditional (stoichiometric) chemistry was gaining attention.^{181–183} This gained more traction with the conceptual development of green chemistry in the 1990, in parallel to the advances made in biocatalysis outlined above.^{184,185} Thus, it is not surprising that biocatalysis formed a key strategy of accomplishing the goals of green chemistry from the start.¹⁸² Indeed, it promises to address many of the “12 principles of green chemistry” (**Figure 1-10**), in particular with regard to hazardous reagents and waste, energy requirements, number of steps, and their inherently renewable and biodegradable nature.^{186,187} Indeed, the number of biocatalytic processes in industry started increasing rapidly and continues to do so to this day: there were around 60 processes in 1990, 134 processes in 2002, and several hundred by 2019.^{188–190}



Figure 1-10: The 12 principles of green chemistry, reproduced with permission from ACS Green Chemistry Institute® (<https://www.acs.org/content/acs/en/greenchemistry/principles/12-principles-of-green-chemistry.html>). Copyright 2020 American Chemical Society.

Perhaps one of the most successful examples developed at the time was the use of lipases, in particularly CalB from *Candida antarctica*, in organic solvents, allowing ester and amide formation without competing hydrolysis, which is frequently employed in (dynamic) kinetic resolutions of chiral alcohols and amines. The latter was developed at BASF¹⁹¹ and is

often referred to as the “BASF process” (**Scheme 1-4**).¹⁹² By 2004, multiple BASF plants produced chiral amines on a >1000 ton scale per year, and this process is still in use today. The reactions can be carried out without solvent, are often nearly quantitative, both amine and amide are readily isolated, and the undesired product can be recycled, making this process highly efficient.^{193,194}



Scheme 1-4: The lipase catalysed BASF process for the kinetic resolution of amines. Enantioselectivity is often essentially perfect and conversions quantitative, the amide and amine can be separated by distillation, the amide is readily hydrolysed (giving access to both enantiomers), and the undesired enantiomer can be racemized and recycled. The process can also be run neat (e.g. in the case of 1-methoxy-2-aminopropane).¹⁹⁴ Other esters than the ethyl ester may be used; however, the methoxy-group is critical for an efficient reaction.

While the BASF process uses a wild-type enzyme, stabilized through immobilization, a strong interest in enzyme engineering developed in industry. On major barrier to directed evolution is screening. In Frances Arnold’s original paper,¹⁶⁰ the enzyme was secreted from colonies of bacteria and digestion of the substrate in plates could be easily observed by a decrease in turbidity, thus allowing the screening of a large number of variants relatively easily. However, in general screening is not straightforward and usually it is the bottleneck. This is tackled on two fronts: development of faster high-throughput screens, as well as reduction of the size of libraries by increasing the proportion of hits. While the former is often highly specific to the (class of) enzymes being evolved, more general concepts exist for the latter. For example, the structure of the enzyme may be used to assess which recombinations are more likely not to disrupt the overall fold of the enzyme, in a process called SCHEMA,¹⁹⁵ which can then be used to reduce the size of combinatorial libraries.

1.1.5 “Smart” libraries and applications of enzyme engineering

Advances in the understanding of protein structures as well dynamics have increasingly allowed target residues to be identified with more reliability than was previously possible, reducing the need for random mutagenesis across the whole gene although it remains a valuable tool.¹⁹⁶ The availability of increasing numbers of structures of diverse enzymes within a given family, and the even larger availability of sequences allow points of natural variation to be identified which may then be targeted. In addition, the amino acids found in nature for a given position can inform which substitutions to include in a given library.^{197,198} Alternatively, random mutagenesis might be used to identify hotspots which are then further investigated by more targeted mutagenesis.¹⁹⁹

Several residues may be targeted together, to increase the chance of detecting synergistic effects of mutations. One such approach is combinatorial active site testing (CASTing),^{200–202} whereby multiple residues lining the active site are saturated at the same time, allowing for synergistic effects between mutations to emerge. This has been particularly successful in changing enantioselectivities and substrate scopes of enzymes.^{192,203} Amine dehydrogenases (AmDH) were created from amino acid dehydrogenases in this way (**Section 1.2.2**). In addition to interactions in the active site (traditional docking), the trajectory of the substrate as it enters the active site may also reveal key residues (e.g. steric clashes or bottlenecks). One recent tool that enable studying this is CaverDock.²⁰⁴

Statistical tools and machine learning are also a powerful way to increase the efficiency of directed evolution,²⁰⁵ such as the use of protein sequence activity relationships (ProSAR).²⁰⁶ In an initial library, mutations are classified as beneficial, neutral, or detrimental and can inform which mutations to incorporate into subsequent libraries, as opposed to taking the best overall variant and generating a new library. This strategy was successfully applied by Codexis in the engineering of a halohydrin dehalogenase (HHDH) for the synthesis of (*R*)-4-cyano-3-hydroxybutyrate, a key intermediate for the synthesis of atorvastatin, a cholesterol-lowering drug.²⁰⁷ Overall, the volumetric productivity was improved 4000-fold over 18 rounds of evolution and 35 mutations were introduced, meeting the process requirements for the enzyme. The authors note that half of the mutations introduced in the final variant were initially not present in the best variant when selected and would have been missed in a hit-based approach. While this approach can reduce screening efforts, it requires a larger sequencing effort. However, this has become increasingly possible as the cost of DNA sequencing has steadily declined.¹³⁸

Codexis and Merck combined several of these approaches to engineer a transaminase for the synthesis of sitagliptin (more on transaminases in **Section 1.2.2.1**). Starting from an enzyme with no activity toward the substrate and minimal activity toward a truncated analogue, 11 rounds of engineering (**Figure 1-11**) led to a catalyst that outcompeted the alternative rhodium catalysed reductive amination process in terms of efficiency, yield, enantioselectivity, and waste formation. Overall, 27 mutations were introduced using a combination of site-saturation mutagenesis, combinatorial libraries (including diversity from homologous sequences), proSAR, and epPCR, screening a total of 36480 variants.²⁰⁸

Highlighting that, even with the use of tools to maximize the efficiency of evolution, a huge screening effort may still be required for significant catalyst improvements.

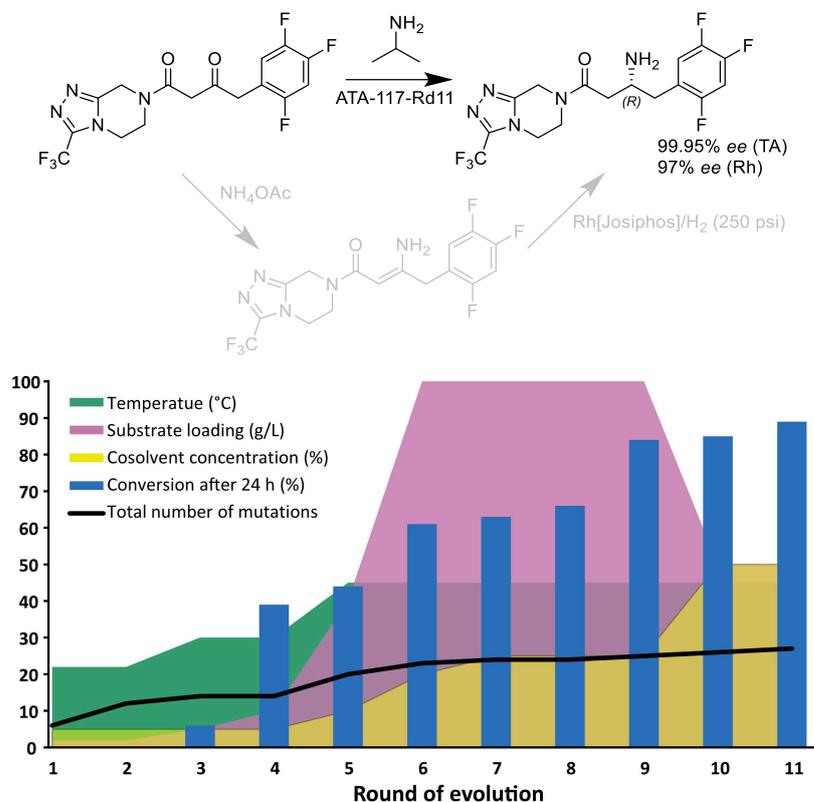


Figure 1-11: Evolution of ATA-117 to produce sitagliptin,²⁰⁸ compared to the chemocatalytic route using a rhodium catalyst. Over 11 rounds of evolution, the conditions of the screening (substrate loading, temperature, cosolvent concentration (Rd 3-6 MeOH, otherwise DMSO)) were gradually increased to the process level. Overlaid is the steady increase in conversion under process conditions, as well as the increase in the total number of mutations (note, several mutations changed throughout the process).

In another more recent example, GSK evolved an imine reductase (IRED) to meet the process requirements for the synthesis of the LSD1 inhibitor GSK2879552, currently in clinical trials.²⁰⁹ Screening of their in-house panel (of at least 85 IREDS)²¹⁰ revealed a suitable candidate for mutagenesis. Given the scarcity of structural data on IREDs and that their highly dynamic mechanism is not fully understood, an initial round of site-saturation mutagenesis was carried out on 256 out of 296 positions. Beneficial mutations from that round were then used to generate combinatorial libraries, which were then analysed using the proprietary CodeEvolver software from Codexis. Statistical analysis was performed to identify pairwise interactions of beneficial mutations which were then included in another combinatorial library in a final third round of evolution, yielding an enzyme with 13 mutations that met or exceeded the process requirements, resulting in improved sustainability metrics over previous route (**Figure 1-12**). The enzyme was then used to synthesize 1.4 kg of GSK2879552 for use in additional rounds of clinical trials.

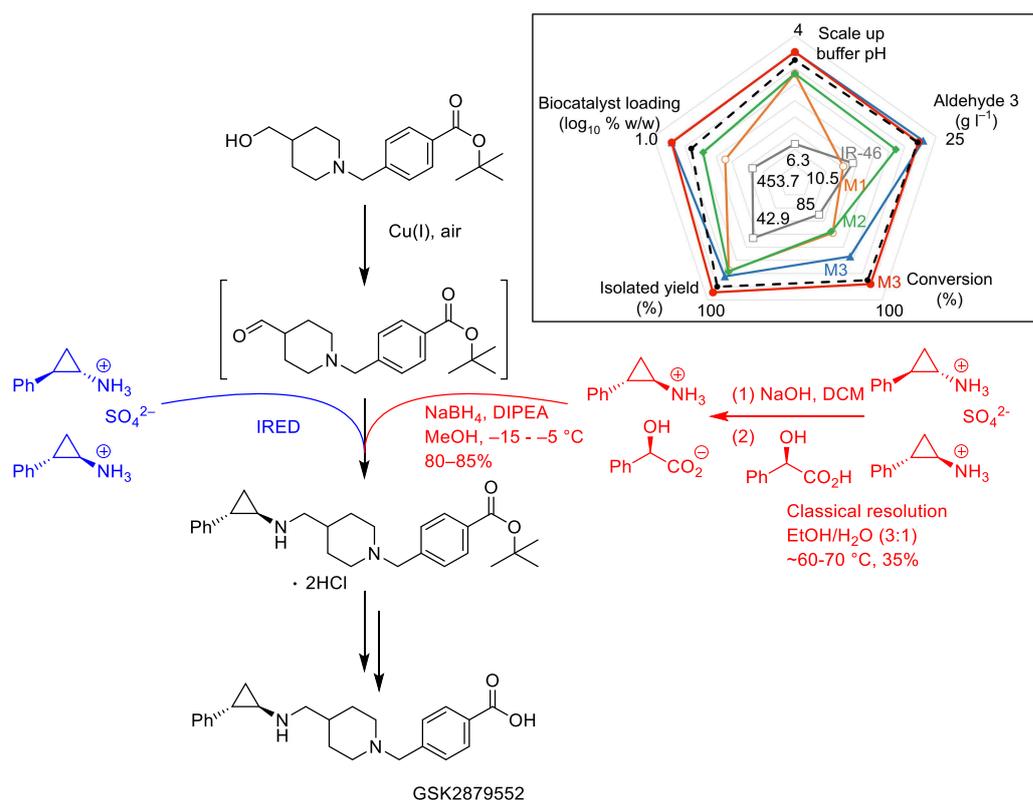


Figure 1-12: Engineering of an IRED for the synthesis of GSK2879552, and the alternative chemical route. Insert: improvement of the catalyst over 3 rounds of evolution; acceptable operating space (black dotted line), wild type IR-46 (grey), M1 (orange), M2 (green), M3 at small scale (blue) and process scale (red). Adapted with permission from Springer Nature (ref.²⁰⁹), copyright 2019.

Increasingly, directed evolution has also been applied to create enzymes carrying out reactions not observed in nature, by exploiting the promiscuous nature of enzymes (**Figure 1-13**). Frances Arnold's group reported²¹¹ the evolution of a cytochrome c (cyt c), a protein without any catalytic role in nature, to form carbon-silicon bonds, a reaction not observed in nature. After screening several P450 enzymes, myoglobins, and cyt c variants, they identified a cyt c from *Rhodothermus marinus* with low levels of catalytic activity for this reaction. Iterative site-saturation mutagenesis of just three key residues—an iron co-ordinating methionine, and two additional residues close to the heme group—resulted in a catalyst with a total turnover number (TTN) of >1500, a >33-fold improvement over the wild-type (wt) and a >375-fold improvement over free heme, outperforming the best chemical catalysts for this reaction. In addition, the turnover frequency (TOF) was increased 7-fold, the reaction proceeded with nearly perfect enantioselectivity, and was chemoselective for carbene insertion into silanes over alcohols and amines (**Figure 1-13**).

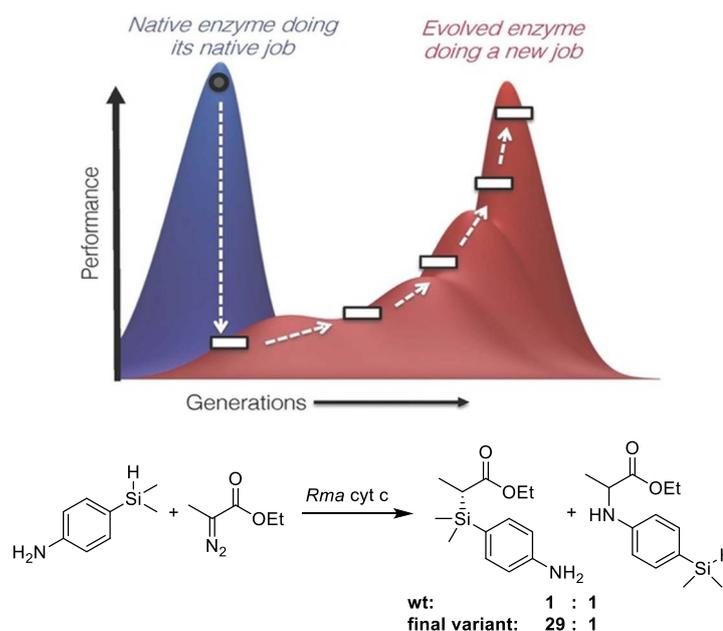


Figure 1-13: Top: Exploiting enzyme promiscuity to evolve new catalytic activity. Enzymes often exhibit promiscuous activity toward non-native substrates or reactions. By applying directed evolution, a “specialist” enzyme might be transformed into another specialist enzyme for the new activity, at the cost of diminishing its original function. Such a transformation proceeds through a “low-fitness valley” where the enzyme is not very good at either the new or the original function. Figure reproduced with permission from ref.¹⁵⁹ Bottom: This concept was applied to evolve a cytochrome c from *Rhodothermus marinus* (without any native catalytic function) to catalyse Si-H carbene insertions.²¹¹ The final variant was 33x more active than the parent and became more specialized for Si-H insertion over N-H insertion chemistry, both promiscuous activities of the wt.

Clearly, in addition to the generation of efficient libraries, the choice of template was key in both examples above. Indeed, modern genomics has created huge databases of genes and it has become incredibly easy to identify new sequences that likely have a given catalytic function. The dramatic reduction in cost of synthetic genes has made it possible to create panels of these sequences, in a way producing a “smart” library of sequences already pre-selected by natural evolution. The likely function of a DNA sequence may be determined from sequence similarity to other proteins of known function (using search algorithms such as BLAST),²¹² as well as the identification of motifs.²¹³ In addition, structure-based searches of crystal structures of unknown function may yield new biocatalyst.²¹⁴

1.1.6 Enzyme immobilization and flow chemistry

Protein immobilization, which, as already outlined above, is a key strategy for enzyme stabilization and reusability, has also become more advanced and a whole plethora of different strategies and supports are available (**Figure 1-14**). Those are needed in part because protein immobilization can be highly unpredictable, and because of application-dependent requirements on the immobilized catalyst (such as particle size, swelling, hydrophilicity/hydrophobicity, etc.). Enzyme immobilization may be mechanical or

physiochemical, the latter can be further divided into covalent or noncovalent/adsorption immobilization. Immobilized proteins may also be used in continuous (i.e. flow chemistry) processes, which is particularly attractive for its scalability, improved efficiency, and increased control.²¹⁵

Mechanical immobilization relies on the entrapment of the enzyme in a matrix that restricts its movement (**Figure 1-14A**). This has the advantage that the enzyme itself is not being modified, while allowing for its environment to be fine-tuned. However, mass transfer to and from the enzyme is often impaired.²¹⁶ Furthermore, leaching of the enzyme can occur.²¹⁷ Adsorption onto a solid support is another simple immobilization strategy. However, it too often suffers from leaching of the enzyme from the support. In its simplest form (hydrophobic, hydrophilic or ionic interactions between support and enzyme. **Figure 1-14B**), no control over the orientation of the enzyme is achieved and the entrance to the active site may become blocked.²¹⁸ One of the most widely used biocatalysts, CalB, is immobilized in this way (Novozym 435), through hydrophobic interactions between the enzyme and an acrylic resin support. While leaching is an issue in an aqueous environment, this is suppressed in the organic solvents in which it is usually used.²¹⁹

More specific adsorption is possible with the use of tags. Attached at either the N- or C-terminal of the protein, they can help orient the enzyme in a favourable position. Examples include the use of a polyhistidine-tag (His-tag, **Figure 1-14C**), originally developed for efficient protein purification in 1988,²²⁰ which coordinates to transition metal cations, streptavidin with its remarkably high affinity for biotin, and sugar-lectin interactions. While a His-tag is encoded genetically, the other two examples require biotinylation or glycosylation, respectively. This also allows for enzyme purification and immobilization to be combined into a single step. While these interactions are stronger than the simple adsorption described above, low levels of leaching can still occur and pose problems for applications in flow, where any enzyme leaching from the column will be lost (as opposed to batch processes where temporarily detached enzyme remains in the vicinity of the support and can, in principle, reattach). One example are EziG beads, made of controlled porosity glass which has been modified to coordinate to metal ions.²²¹ They appear to be predominantly used in organic solvents,^{221,222} which appear to suppress leaching although use in an aqueous environment without leaching has also been reported (using Fe³⁺ as the cation).²²³

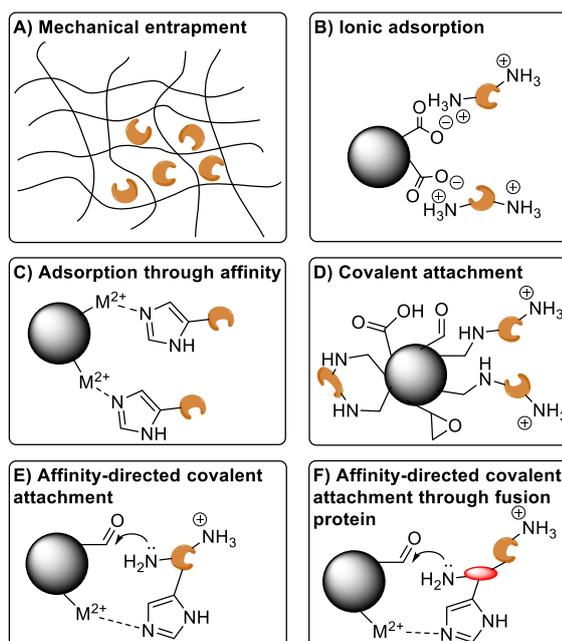
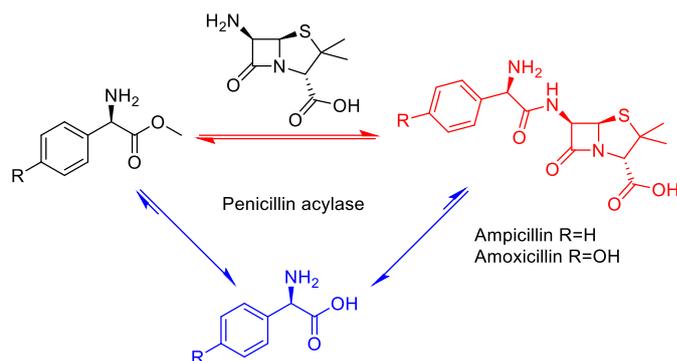


Figure 1-14: Examples of enzyme immobilization strategies. **A)** Mechanical entrapment restricts the diffusion of the enzyme. **B)** Adsorption through ionic interactions, offering little control over the orientation of the enzyme. **C)** Adsorption through affinity, in this case His-tag-metal coordination, allowing control over the orientation of the enzyme through the tag placement. **D)** Covalent attachment, offering little control over the orientation of the enzyme. Multipoint attachment can lead to irreversible deformation of the enzyme shape. Common functional groups for covalent attachment are carboxylic acids, aldehydes, and epoxides—using amide formation, reductive amination, and ring opening, respectively. **E)** Affinity-directed covalent immobilization orients the enzyme prior to covalent attachment. **F)** By fusing a (small) protein to the enzyme, covalent immobilization and any shape disruption can be localized to that fusion protein, reducing the effect on the enzyme. However, such an enzyme is more exposed to the environment and stability benefits from immobilization may be diminished. Not shown are covalent crosslinking of enzymes (e.g. using a dialdehyde), and the inherent different properties of supports, with respect to e.g. their size, pore-size, hydrophilicity/hydrophobicity, etc.

The problem of leaching can be fully avoided using covalent immobilization (**Figure 1-14D**). However, it often results in severe distortions of the enzyme and loss of activity, although this is highly dependent on the enzyme and support and often unpredictable. Leaching may also still occur for multimeric proteins if not all subunits are covalently attached. Moreover, once the enzyme has degraded the support cannot be reused whereas with adsorption the enzyme may be desorbed and replaced with fresh enzyme.²¹⁹ The orientation of the enzyme may be controlled by initially adsorbing the enzyme onto the support using tags, followed by the formation of the covalent attachment (**Figure 1-14E**). The distortion of the enzyme can be alleviated using small protein tags, with the covalent attachment points being on that protein rather than the enzyme itself (**Figure 1-14F**).^{224,225} However, as the enzyme is more exposed to solvent, the stabilizing effect of immobilization is reduced. One elegant tag directed covalent immobilization strategy is the use of the SpyTag/SpyCatcher system, a peptide and protein that spontaneously form an isopeptide bond when coming together.^{226,227}



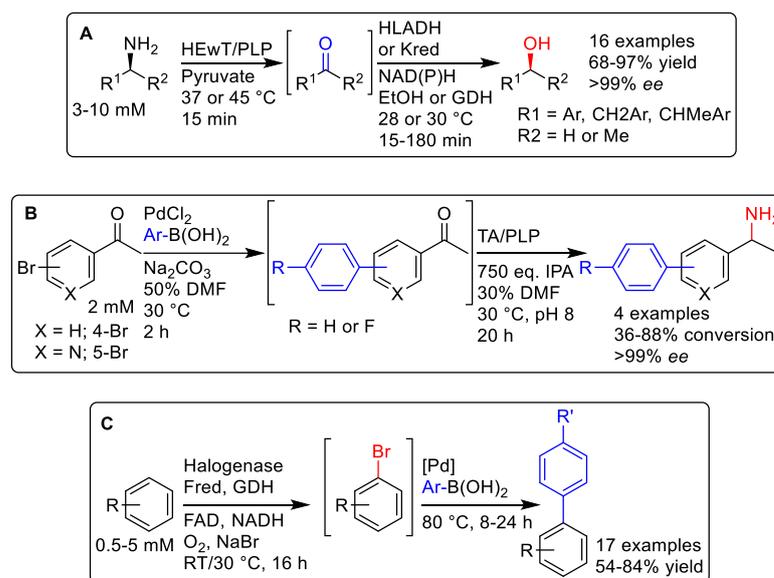
Scheme 1-5: Competing acyl transfer (red) and hydrolysis reactions (blue) catalysed by Penicillin acylase. By tuning the characteristics of the support, synthesis can be kinetically favoured over the thermodynamically favoured hydrolysis reaction.

Enzymes may also be covalently cross-linked into cross-linked enzyme aggregates (CLEAs), an immobilization without a support. However, the poorly defined properties (such as particle size) of these aggregates often render them unsuitable for flow applications. However, adsorption of these aggregates onto a support with more defined properties can alleviate this issue. Glucose isomerase, immobilized in this way, is being used for the production of HFCs in flow on a 10 million tons per year scale, using 500 tons of the immobilized catalyst.^{77,194,228} Regardless of the immobilization technique used, tuning the characteristics of the support with respect to e.g. their size, pore-size, hydrophilicity/hydrophobicity, etc. is also key. For example, tuning the composition of the support reduced the synthesis/hydrolysis (S/H) ratio of covalently immobilized penicillin acylase (**Scheme 1-5**). This was a key step in rendering it suitable for the kinetically controlled synthesis of various semi-synthetic β -lactam antibiotics.^{71,79,132}

In some cases, subunit dissociation followed by leaching can significantly reduce the operational stability of covalently immobilized biocatalysts. In those cases, coating the biocatalysts either before or after immobilization with other polymers, such as polyethylenimine (PEI) or activated dextran can help maintain the quaternary structure.²¹⁶ However, as with any immobilization, excessive rigidification of the enzyme can result in a loss of catalytic efficiency if structural rearrangements necessary for catalysis are impeded. In addition, the increased complexity of more sophisticated immobilization techniques often outweighs any benefits bestowed on the catalyst. In general, the simplest catalysts that can meet the process requirements is the preferred one.

Synthetic cascades, without the need for purification of intermediates are very attractive due to the reduced amount of waste that is produced. In addition, intermediates that are too unstable to be isolated may be telescoped to the next step, offering alternative

routes to classical synthesis. Flow chemistry, being inherently modular by design, is a very important platform for such cascades. Sequential reactions can be compartmentalized, avoiding incompatibility between reagents as well as allowing the conditions to be fine-tuned for each reaction.²¹⁵ Cascades involving biocatalysis may either be multiple enzymatic reactions combined in sequence, or chemo(catalytic) reactions combined with enzymatic reactions.²²⁹



Scheme 1-6: Three examples of enzymatic cascades. **A)** Transformation of amines into alcohols, using an immobilized transaminase and either an ADH or Kred in flow.²³⁰ **B)** A Suzuki cross-coupling to produce a bi-aryl ketone which is then aminated using a transaminase catalyst. The transaminase had to tolerate 30% DMF carried over from the cross-coupling, as well Pd catalyst, excess base, and unreacted boronic acid.²³¹ **C)** halogenation of aromatic compounds using a halogenase, followed by a Suzuki coupling. The enzyme had to be separated, either by ultrafiltration, immobilization, or compartmentalization from the Pd catalyst.²³²

A nice example of the former was demonstrated by Contente and Paradisi,²³⁰ who developed a cascade in flow converting amines into alcohols, employing a transaminase and either an alcohol dehydrogenase (ADH) or ketoreductase (Kred) that had been immobilized covalently on epoxide functionalized methacrylate beads (**Scheme 1-6A**). By compartmentalizing both catalysts, reaction temperatures and times were optimized independently for each step. In addition, in-line purification steps allowed the removal of product. Recycling of the aqueous phase containing co-factors and buffer salts was also demonstrated, reducing the overall amount of the cofactors required (from 1:100 to 1:2000) while also eliminating the aqueous waste stream. Uwe Bornscheuer's group demonstrated a Suzuki-Miyaura coupling in batch to produce a biaryl ketone substrate for a subsequent transaminase-catalysed amination in flow (**Scheme 1-6B**).²³¹ Here, the ability of the transaminase reaction to tolerate 30% (v/v) DMF as well as salts and palladium from the first reaction step was key. Compatibility between palladium and enzyme catalysts can be a

problem, as was the case for a halogenase-Suzuki-coupling cascade in batch reported by Latham *et al.* (**Scheme 1-6C**).²³² Using free enzyme, ultrafiltration or compartmentalization with a semi-permeable membrane had to be used to physically separate the enzyme and Pd catalyst. Alternatively, immobilization of the halogenase into CLEAs was also successful.

In another collaboration between Codexis and Merck, a three-step, nine-enzyme cascade to synthesize the HIV drug islatravir was developed (**Figure 1-15**).²³³ This involved engineering of five enzymes to accept unnatural substrates, as well as enzyme immobilization to simplify the final purification. For this, a cost-effective affinity immobilization using a His-tag was chosen for the first two steps, while the last step used free enzymes. Enzymes from seven organisms were used, and each step involved enzymes from either two or three organisms and one or three evolved enzymes. Thus, by bringing together the right enzymes from the right organisms (a testament to the vast number of genome sequences that are available) and applying directed evolution only were needed, the number of steps in the synthesis of islatravir was cut more than in half (from 12-18 steps), and the overall yield was almost doubled (51% vs 37% previously reported²³⁴). Atom economy was improved, overall waste was reduced, and hazardous reagents (such as a Birch reduction in Fukuyama *et al.*'s synthesis²³⁴) were avoided.

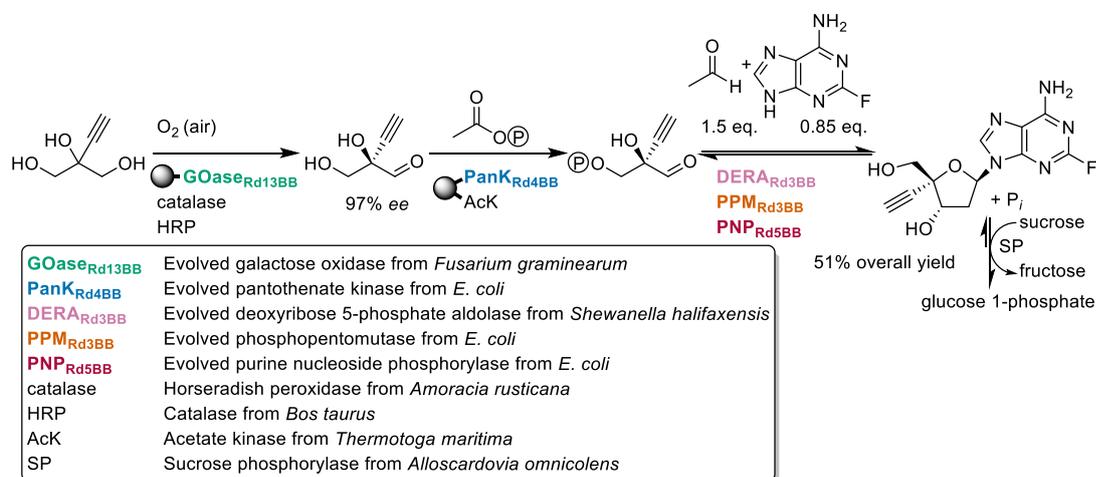


Figure 1-15: 9-enzyme cascade to produce the HIV drug islatravir. Five enzymes had to be evolved. Compared to a chemical synthesis, steps were reduced by more than half and yield was almost doubled. No purification of intermediates was necessary. Immobilized enzymes shown attached to spheres. Figure adapted from Huffman *et al.*²³³

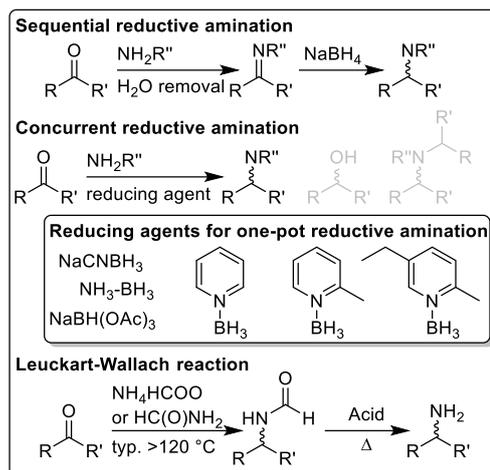
1.2 Amines

The amino group is a key functional group found in many industrially relevant compounds particularly in pharmaceuticals and agrochemicals.^{193,235} Estimates have been made that 30-40% of active pharmaceutical ingredients (APIs) contain chiral amino groups.^{236,237} Thus, the synthesis of chiral amines is clearly a key reaction, and a round table of several pharmaceutical companies concluded in 2007 that efficient syntheses of chiral amines from carbonyl groups was a key area in which green reactions needed to be developed.²³⁸ One challenging aspect of amine synthesis is the increasing reactivity of amines with their substitution. For example, simple nucleophilic substitutions using amine nucleophiles often results in the formation of quaternary ammonium ions.

1.2.1 Chemical synthesis of chiral amines

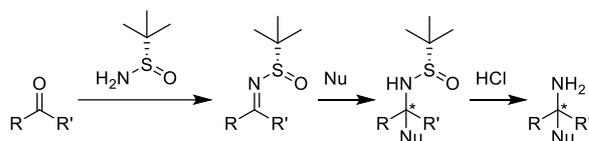
The most important reaction to form amines is by reductive amination of carbonyl compounds (**Scheme 1-7**), usually proceeding via imines or enamines, which are prone to hydrolysis. Water-removal often needs to be used to drive the complete formation of the imine, followed by subsequent reduction. Strategies also exist to form the imine and reduce it concurrently, traditionally using cyanoborohydride which selectively reduces the more reactive iminium ion over ketones. However, the use cyanoborohydride is disfavoured due to the potential for hydrogen cyanide formation. Several alternative reducing agents have been developed, such as pyridine borane and derivatives (picoline borane and 5-Ethyl-2-methylpyridine borane), or sodium triacetoxyborohydride. However, problems with over-substitution and solvent incompatibilities do exist and, in particular, primary amines can be challenging to synthesise in this way.²³⁹⁻²⁴¹ Amines can also be synthesized using the Leuckart-Wallach reaction, using ammonium formate or formamides and high reaction temperatures (typically 120 °C or higher).^{242,243}

Resolution of racemic amines using chiral carboxylic acids is often very effective at producing high enantiomeric excesses.¹⁹³ However, this process is necessarily limited to a yield of 50% and not very attractive from a modern green chemistry point of view, and thus the asymmetric synthesis of chiral amines is generally preferred.



Scheme 1-7: Common strategies for the synthesis of amines. Reductive amination of a carbonyl, via the imine. Water removal is often necessary to assure complete imine formation prior to reduction. Concurrent imine formation and reduction relies on the use of reducing agents that selectively reduce iminium ions but not ketones. Over substitution can be a problem. The case of secondary amines is shown, although primary and tertiary amines can also be produced. The Leuckart-Wallach reaction: Ammonium formate or formamide reacts with a carbonyl to form a substituted formamide, which may then be hydrolysed. Secondary and tertiary amines are also accessible.²⁴⁴

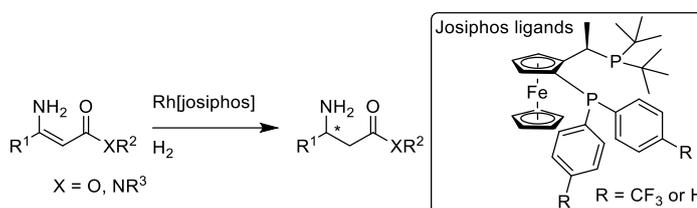
One of the most common syntheses of primary chiral amines is in reductive aminations using Ellman's sulfinamide as a chiral auxiliary (**Scheme 1-8**).²⁴⁵ Other chiral amines, such as methylbenzylamine, have also been used.^{193,246} Usually, the imine is pre-formed prior to reduction, followed by deprotection, requiring three steps. Clearly, the low atom economy from using an auxiliary group and the need for three steps are not desirable, yet due to high yields and often good to excellent enantiomeric excess for wide variety of substrates, it remains a valuable tool.^{247,248} Recycling strategies for the auxiliary have also been proposed.²⁴⁹



Scheme 1-8: Asymmetric synthesis of primary amines using Ellman's sulfinamide, showing a general nucleophile which may also be a hydride from a reducing agent. Water removal is often necessary to form the sulfinyl imine. Redrawn from Xu *et al.*²⁴⁸

In general, catalytic methods are preferred. These usually rely on transition-metal catalysed hydrogenation with chiral ligands.²⁵⁰ For example, a rhodium based catalyst using chiral ferrocene based ligands was developed at Merck for the asymmetric reduction of enamines from β -ketoesters/-amides in 2004 (**Scheme 1-9**).²⁵¹ This is the process that was initially used for the sitagliptin synthesis mentioned above.²⁰⁸ A more general variant that does not require the dicarbonyl and forms the imine *in-situ* has also been developed, using

a ruthenium catalyst with a different ferrocene based ligand, and catalytic amounts of ammonium iodide.²⁵² Several similar systems exist for the asymmetric synthesis of secondary amines, using either iridium, ruthenium, rhodium, palladium, and titanium catalysts, either in step-wise or concurrent fashion.^{253–255} However, while these syntheses often produce high enantiomeric excess, they frequently do not meet requirements for enantiopurity needed for pharmaceutical ingredients, necessitating further purification, offering opportunities for biocatalysis. In addition, the high pressures needed for hydrogenation, the required removal of toxic transition metal impurities, and the unrenewable nature of transition metal catalysts are also undesirable.



Scheme 1-9: Rhodium[josiphos] catalyzed asymmetric reduction of β -enamine esters/amides, developed by Merck.²⁵¹ The toxic catalyst needs to be carefully removed and the enantiomeric excess of the amine often needs to be further improved, e.g. by crystallization.

1.2.2 Enzymatic synthesis of amines

As already mentioned, the BASF process is a highly efficient kinetic resolution of amines (**Scheme 1-4**). It is particularly suitable for the production of chiral amine building blocks, due to the ease of isolating both enantiomers. Additionally, should only one enantiomer be desired, recycling of the undesirable enantiomer is possible.^{193,194} However, this necessarily introduces additional steps and generates additional waste. The use of an acylating agent also reduces atom economy; however, space-time yields (due to solvent-less reactions) and catalyst longevity are almost unprecedented.¹⁹⁴

Monoamine oxidases (MAO) catalyze the oxidation of amines to imines, using molecular oxygen (air). Coupled with a suitable reducing agent, such as ammonia borane, they have been applied in the deracemization of chiral amines since the early 2000s.^{237,256–258} Using extensive directed evolution, Nick Turner's group have been able to significantly widen the substrate scope of MAO-N (from *Aspergillus niger*) and applied it to increasingly large and unnatural substrates including fused heterocycles and tertiary amines (**Figure 1-16**).^{259–262}

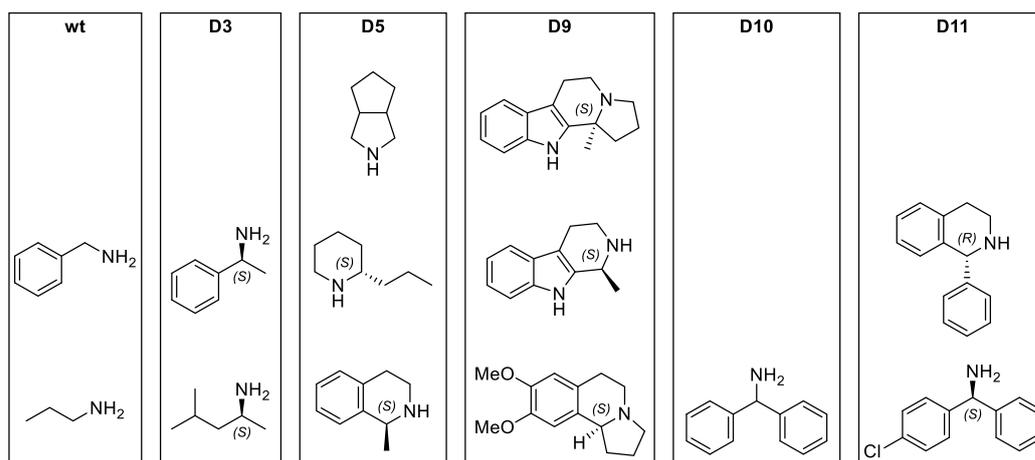
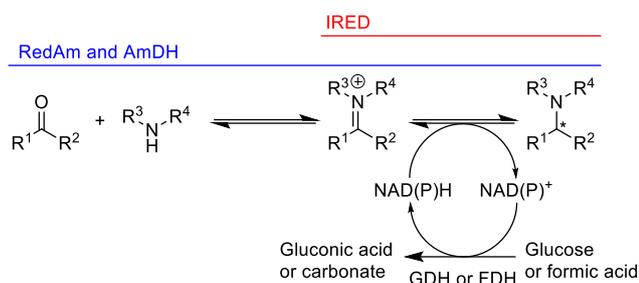


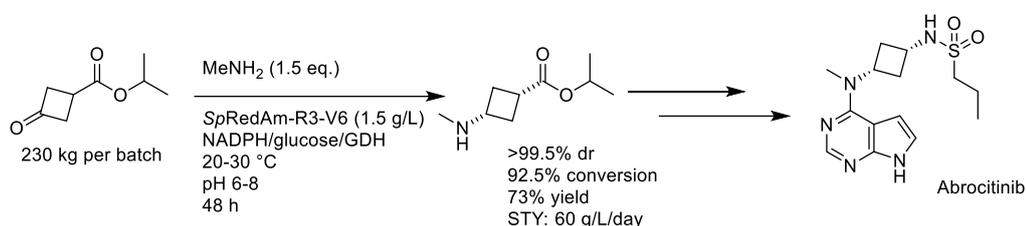
Figure 1-16: Diversity in substrate scope of evolved variants of MAO-N.²³⁷

Amine dehydrogenases (AmDHs), imine reductases (IREDs), and reductive aminases (RedAms) are an emerging classes of enzymes^{237,263} that are very attractive due to their ability to produce chiral amines directly from ammonia (to produce primary amines) or amines (to produce secondary or tertiary amines) (**Scheme 1-10**). This sets them apart from transaminases which can only synthesize primary amines using an amino donor, forming a carbonyl by-product (**Section 1.2.2.1**). However, unlike transaminases they require co-factor recycling systems due to their dependence on NAD(P)H. AmDHs are engineered amino acid dehydrogenases, originally developed by the Bommarius group (using a CASTing methodology), removing the structural requirement of the carboxylic acid group.^{264–266} The discovery of natural AmDHs has recently been reported from metagenomes.²⁶⁷ They catalyse the formation of primary amines from ammonia, on a variety of substrates,^{268–270} although promiscuity for the formation of secondary amines has been reported by Francesco Mutti's group.²⁷¹ IREDs and RedAms are very similar in their structure, and apparent differences in their activity (reduction of imines vs catalytic formation of imine followed by reduction) are still not fully understood.^{272,273} They are primarily applied for the synthesis of secondary and tertiary amines, although synthesis of primary amines has also been reported.²⁷⁴ Due to the unfavourable equilibrium for imine formation in water, cyclic imines are preferred substrates of IREDs and excesses of the amine are usually required for the synthesis of non-cyclic amine products.^{275–279}



Scheme 1-10: Reductive amination using AmDHs, IREDS, or RedAms. RedAms and IREDS are structurally very similar, while AmDHs originate from amino acid dehydrogenases. RedAms and AmDHs catalyse both imine formation and reduction, whereas IREDS only catalyse imine reduction. While IREDS and RedAms have been predominantly applied to the synthesis of primary or tertiary amines and AmDHs for the synthesis of primary amines, this distinction has been blurred in recent years.

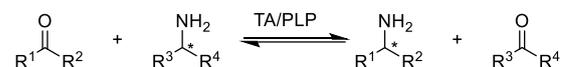
A scale-up of an IRED catalysed reductive amination process has been reported by Johnson Matthey,²⁸⁰ as well as the synthesis of the cancer-drug GSK2879552 mentioned above (**Figure 1-12**).²⁰⁹ Recently, the synthesis of an intermediate in the route to Abrocitinib (a JAK1 inhibitor currently in the final stages of approval) using an engineered bacterial RedAm has been reported by Pfizer.²⁸¹ After three rounds of engineering, introducing only four mutations, the enzyme was used to produce a total of > 3.5 tons of the succinate salt of the intermediate (**Scheme 1-11**), further highlighting their great potential and industrial applicability.



Scheme 1-11: RedAm catalysed synthesis of the *N*-methylamino cyclobutyl intermediate used in the route to Abrocitinib. After three rounds of engineering, the RedAm from *Streptomyces purpureus* could be employed in 230 kg batches, giving the intermediate in good yield with excellent dr.²⁸¹

1.2.2.1 *Transaminases*

In 1937, Braunstein and Kritzmann reported the discovery of the reversible transfer of an amino group from glutamic acid to pyruvic acid that was catalysed by muscle pulp.⁵⁴ Soon thereafter, similar “Umaminierungen” (trans-aminations) were discovered, both in other tissues, as well as involving other substrates. Initially, the name aminopherase was suggested for the enzymes catalysing these reactions, but this quickly started being replaced with the current names transaminase (TA) and aminotransferase (**Scheme 1-12**).²⁸² As already alluded to, from the 1940s, the mechanism (**Figure 1-17**) of these enzymes was being developed, but it took until 1960 for the mechanism to be fully established.^{47,68,69,283–288}



Scheme 1-12: Generalized reaction catalysed by transaminases: an amine group is transferred from a (potentially chiral) primary amine (amine donor) to a (potentially prochiral) carbonyl (amine acceptor).

These first transaminases that were discovered required the carbonyl/amine functional groups of the substrate to be α to a carboxylic acid group and are referred to as α -transaminases. However, in the early 1960s transaminases that did not have this constraint started to emerge.^{289–292} These can be classified as β -, γ -, etc. transaminases to describe the precise position of the carbonyl/amine relative to the acid group. However, these enzymes can often accept multiple substrates with varying positions of the carbonyl/amine group, and the general name ω -transaminase (ω TA) became common to distinguish them from α -transaminases.^{293,294} In cases where an acid group is not required to be present at all, these are often further distinguished as amine transaminases (ATAs).²⁹⁵ ω TA and ATA are often used synonymously.

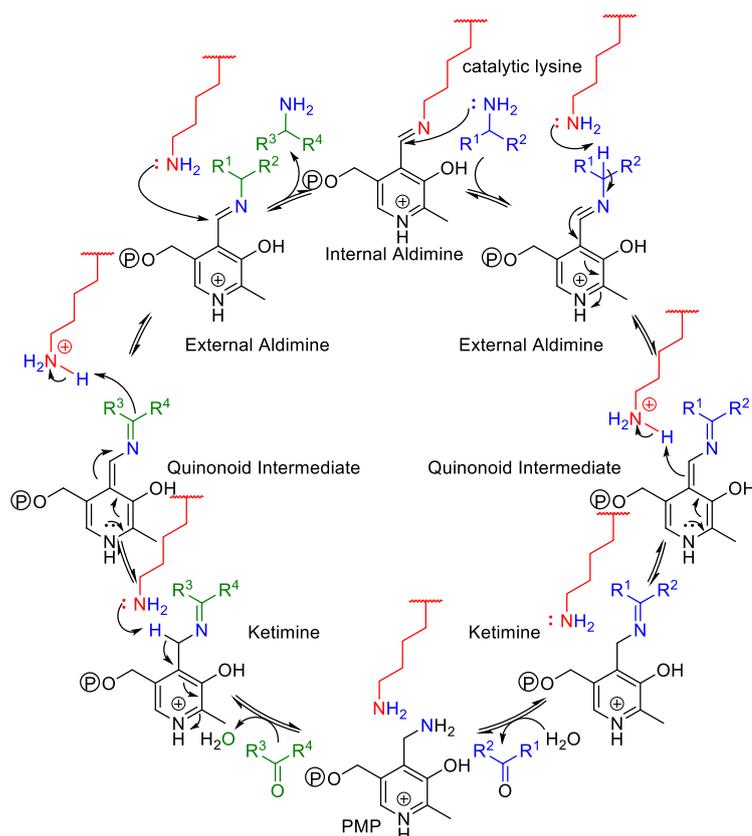


Figure 1-17: Mechanism of transamination. For clarity, the individual steps of aldimine/ketimine formation and hydrolysis as well as transamination are not shown. The mechanism is symmetric—referred to as a “ping-pong” bi-bi or shuttle mechanism—and fully reversible. Note: the ketimine intermediates are a second aldimine if one of the R-groups is a hydrogen. Catalytic lysine: red; amine donor: blue; ketone acceptor: green.

Another classification system groups transaminases into six classes, based on sequence similarity.^{296–298} Confusingly, not all enzymes included in each class are transaminases, in particular for class II. Class III is typically used synonymously with ω TAs,²⁹⁶ however both (*S*)-selective and (*R*)-selective ATAs (STAs and RTAs) exist with only the former being included in the aminotransferase class III. A clearer distinction is based on fold-types of all PLP dependent enzymes, of which there are seven.²⁹⁹ Here, all transaminases belong to fold types I and IV. Other PLP-dependent enzymes include tryptophan synthases (fold type II), alanine racemase (fold type III), glycogen phosphorylase (fold type V), D-lysine 5,6-aminomutase (fold type VI), and lysine 2-3-aminomutase (fold type VII).^{296,300} Transaminases that belong to fold type IV are D-amino acid transaminases (DATAs), branched chain aminotransferases (BCATs), and RTAs. The remaining transaminases belong to fold type I. A more specific nomenclature refers to the natural substrates of the enzymes, e.g. an aspartate: α -ketoglutarate transaminase catalyses the transfer of an amino group from L-aspartic acid to α -ketoglutarate, producing oxaloacetate and L-glutamic acid. It is the orientation of the aldimine intermediate, as determined by the active site, that regulates the type of reaction catalysed by the enzyme. The group that is abstracted to form the quinonoid

intermediate is oriented perpendicular to the ring (H in the case of transaminases), its σ -orbital being aligned with the π^* -orbitals of the double bond network (Dunathan's hypothesis, **Figure 1-18**).^{298,301,302}

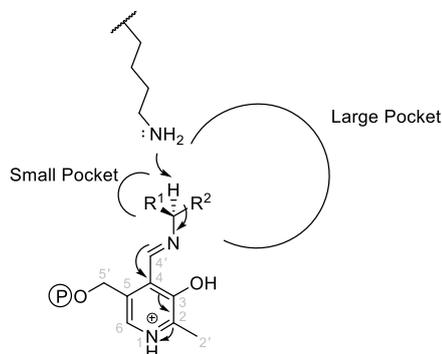
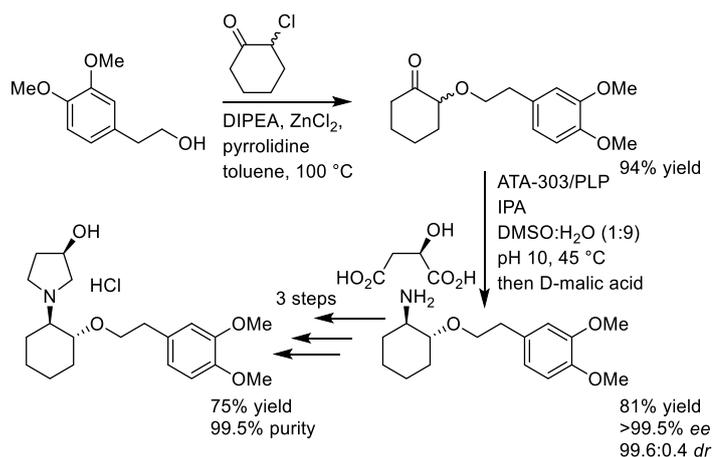


Figure 1-18: Prior to formation of the quinonoid intermediate, the proton to be removed must be oriented perpendicular to the ring system of PLP (Dunathan's rule). A small and a large binding pocket in the active site determine the enantioselectivity of the TA, with proton abstraction occurring on the *si*-face (shown) in STAs or alternatively on the *re*-face in RTAs (*si*- and *re*- are defined with respect to C4'). The small binding pocket can only accommodate groups that are not significantly larger than a methyl group.^{301,303}

The ATAs are of particular interest, as they catalyse the formation of general amines. Between 1990 and 1994 Stirling *et al.* based at Celgene Corporation filed three patents^{304–306} on the use of ω TAs for both the chiral resolution and enantioselective synthesis of optically active amines, the first reported use of ω TAs for this purpose.³⁰⁷ However, the substrate scope of wild-type ATAs is severely limited by the architecture of the active site, entailing a small and a large pocket with groups even slightly larger than a methyl group being rejected from the small pocket (**Figure 1-18**). The enzyme engineering of an RTA, ATA-117, to accept the bulky pro-sitagliptin ketone²⁰⁸ (which has been outlined above) opened up the possibility of freeing transaminases from that restriction. The use of an engineered RTA, by Merck, for the synthesis of vernakalant was another highlight (**Scheme 1-13**), showing the possibility of a dynamic asymmetric transamination in which two stereocenters were set simultaneously.³⁰⁸

The Bornscheuer group then demonstrated that STAs could also be engineered to accept bulky-bulky ketone substrates.^{309–314} However, significant challenges remain. Mutations that are beneficial in one scaffold may have no effect or be deleterious in structurally highly similar scaffolds; however, some hotspots have emerged.



Scheme 1-13: Synthesis of vernakalant (antiarrhythmic agent), involving a dynamic asymmetric transamination. ATA-303 had to be evolved over three rounds to improve diastereoselectivity, as well as increase tolerance of the (for an enzyme) harsh conditions (pH 10, 45 °C) needed for racemization of the ether.³⁰⁸

Determining the appropriate scaffold may also be challenging.³¹² Changing two residues in an STA from *Ruegeria sp.* TM1040 (3FCR), an enzyme with fairly low activity in general, to those found in the more active and extensively studied STAs from *Chromobacterium violaceum* (CvSTA) and *Vibrio fluvialis* (VfSTA) resulted in increased activity toward a set of bulky substrates, including one that neither CvSTA nor VfSTA can accept (**Figure 1-19**). This theme continued and additional mutations, most of which replicated the native residues in CvSTA or VfSTA, further enhanced 3FCR's activity toward these bulky substrates. Additionally, the set of mutations was transferable to another STA which shared sufficient similarity to 3FCR (3GJU, 72% sequence identity). Wild-type STAs bearing the motif with >70% sequence identity to 5FCR also showed good activity toward that set of ketones. The authors proposed that both CvSTA and VfSTA are optimized for their substrates, whereas 3FCR (and homologues), being rather poor STAs ("low-fitness valley", **Figure 1-13**), are more promiscuous and thus the uphill climb to good activity with the bulky ketones is more easily achievable. An alternative explanation might be that 5FCR was already a very specialized transaminase for those bulky ketones and simply handicapped by uncommon residues in the active site. From this point of view, reverting to more common residues simply restored the function of the enzyme.

Screening of transaminase libraries is also challenging. While several medium to high-throughput screens have been developed, they require moving away from process conditions which significantly reduces their utility. In general, one may screen in either the forward (synthesis) or the reverse (resolution) direction, with the former being more desirable. The most general screens are spectrophotometric³¹⁵ and usually rely on the use of methylbenzylamine (MBA) as the donor, forming acetophenone which absorbs at 245 nm

(Scheme 1-14). However, the absorbance of acetophenone cannot be used where the desired starting ketone also absorbs at that wavelength. In these cases, screening in the reverse direction may be possible (using the desired amine product as the donor in place of MBA).³¹² However, this has sometimes identified improved variants where the improvement did not transfer to the forward reaction.^{309,311} Conductometric assays and systems coupled to other enzymes (LDH or LDH/GDH, following consumption of NADH or a pH change, respectively; also see Scheme 1-16) have also been developed, but are less general and more complex to set-up.²⁹⁶

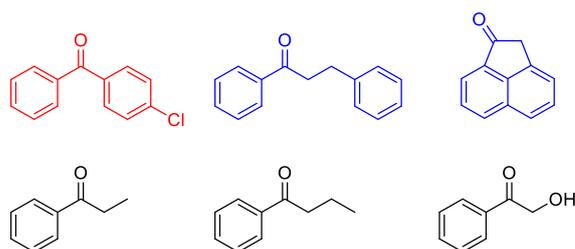
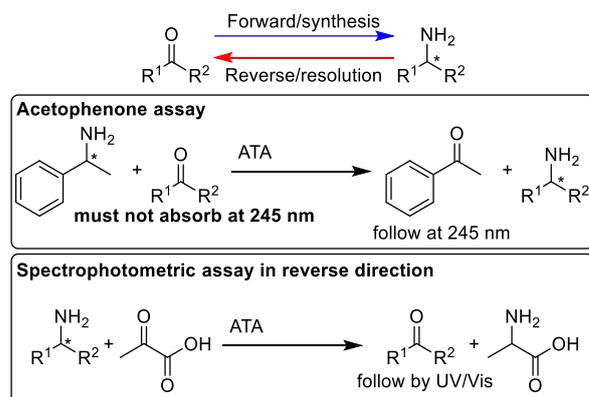


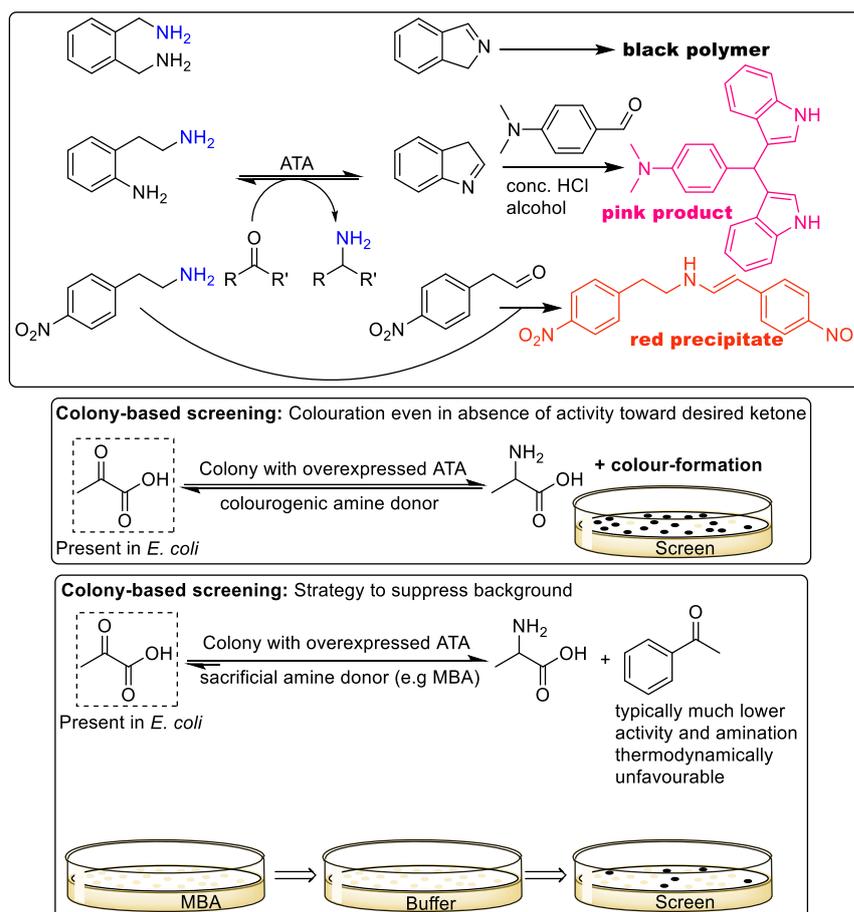
Figure 1-19: Bulky ketones accepted by the 5FCR mutant. Shown in blue, substrates to which V_fSTA, CvSTA, and 5FCR wt had traces of activity. In red, a compound to which neither of the three enzymes had activity. However, with four mutations (most of which already present in V_fSTA and CvSTA) 5FCR's activity significantly increased for all 6 substrates.

MBA may also not be the desired donor for the process due to poorer atom economy and the chiral nature of the substrate (usually only (*R*)MBA or (*S*)MBA are accepted). As no evolutionary pressure is being applied for activity toward the desired amine donor, activity toward it can often be seen to decrease. Reactions also need to be conducted in relatively dilute conditions, so as to not saturate the spectrometer, moving even further away from process conditions.

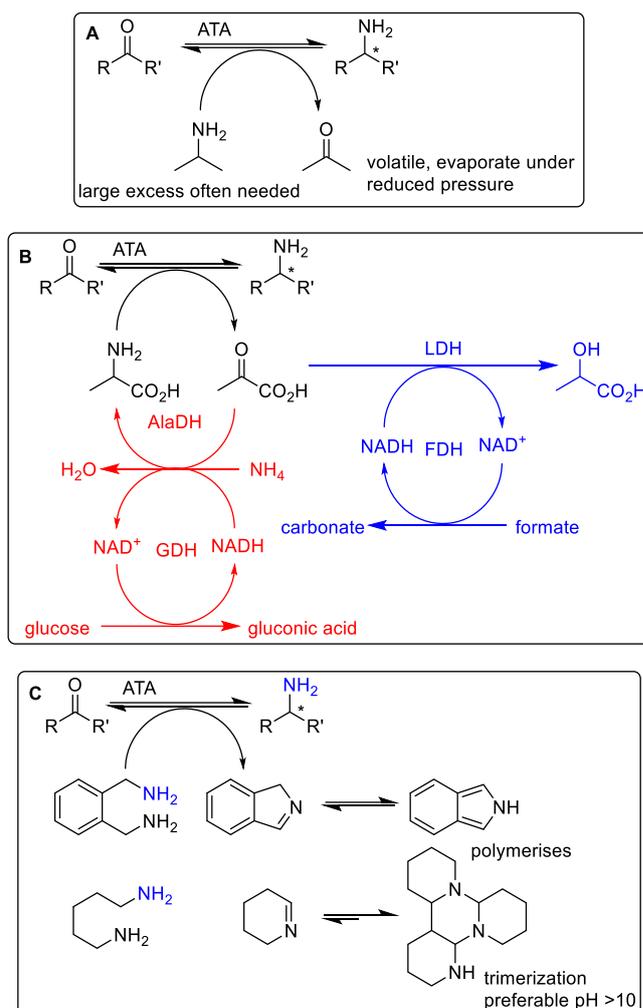


Scheme 1-14: Spectrophotometric medium-throughout assay for transaminases. In the forward direction, MBA is used as the amino donor producing acetophenone with a distinct absorption at 245 nm. However, this cannot be used with substrates that also absorb at 245 nm. In those cases, a reverse screen may be possible, using the desired amine product as the donor. However, beneficial mutation identified in this screen are not necessarily transferable to the forward direction.

Several colourimetric assays have also been developed, both screening for amine donors^{316,317} and ketone acceptors (**Scheme 1-15**).^{318–320} These also rely on using an amine donor (or a keto acceptor) that is unlikely to be used in the final application. Notwithstanding, these screens have been applied to identify variants with activity toward specific substrates^{321,322} and helped identify transaminases from metagenomic libraries.³²³ However, in these cases the screen had to be used on cell extracts. When applying any of the ketone acceptor screens on colonies, another problem arises. Most ATAs have excellent activity toward pyruvate, and thus colour formation is observed even in the absence of any keto-substrate. While a method for suppressing this background has been developed by the Paradisi group,³²⁴ and been applied to screening a random library for enhanced activity toward multiple substrates, the method for pyruvate removal relies on the enzyme one is seeking to evolve. This introduces additional variables that can lead to problems as encountered in **chapter 3**.



Scheme 1-15: Colourimetric screens in the forward direction.^{318–320} These can also be applied to colonies; however, endogenous pyruvate is usually accepted by ATAs, causing colour-formation even in the absence of any ketone substrate, causing indiscriminate colour-formation. This may be reduced by depleting the pyruvate using the transaminase and a non-colourogenic amine donor, allowing colonies expressing variants with activity toward the target ketone to be distinguished.³²⁴



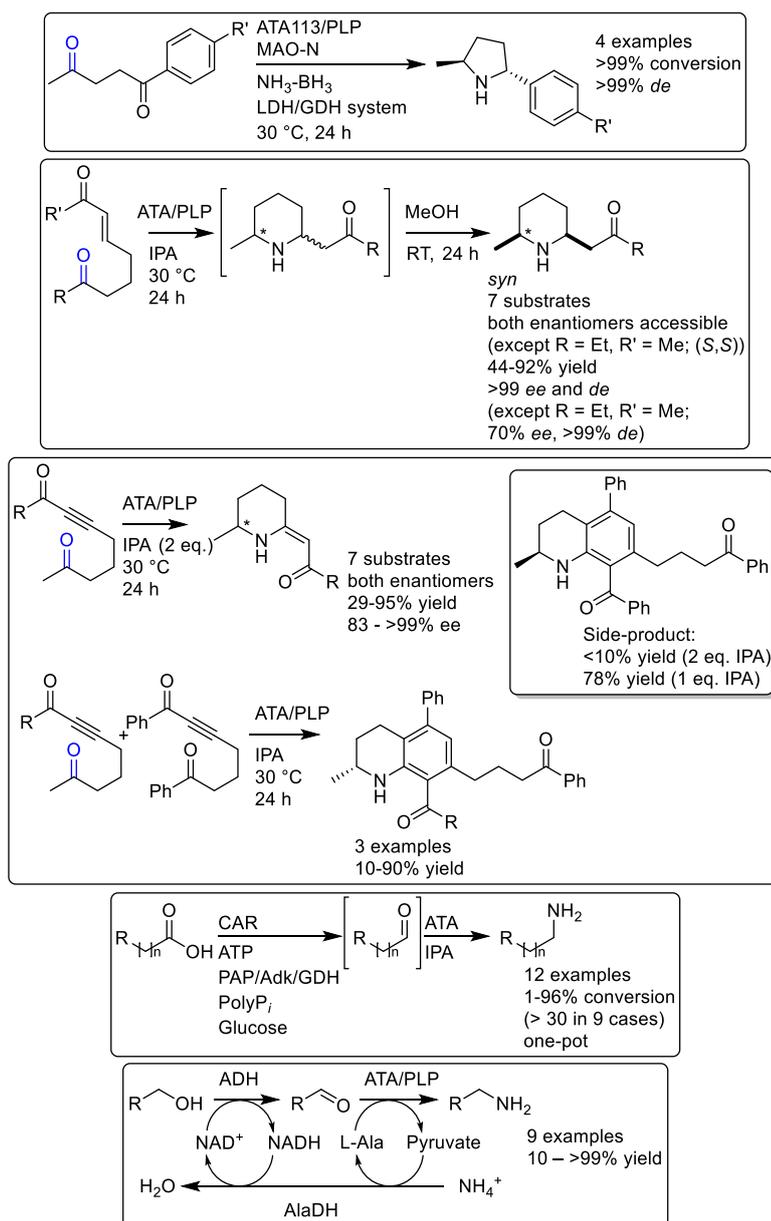
Scheme 1-16: Common strategies to shift the equilibrium of TA catalysed reactions. **A)** IPA is cheap and atom-efficient. It is completely miscible with water, allowing a large excess to be used. Acetone is volatile and may be removed by evaporation, provided the ketone starting material has a sufficiently high boiling point (amines are protonated at typical ATA conditions (pH \approx 8)). **B)** Alanine may be used when coupled with a pyruvate removal system, such as lactate dehydrogenase (LDH, blue) or Alanine dehydrogenase (AlaDH, red). In either case, NADH has to be recycled (using either glucose dehydrogenase (GDH) or formate dehydrogenase (FDH)). AlaDH produces L-Ala, so in the case of STAs it is recycling pyruvate rather than removing it, resulting in increased efficiency. **C)** “Smart” amine donors, such as *o*-xylylenediamine (*o*XDA) and cadaverine, are design in such a way that the by-product reacts further, removing itself from the equilibrium. The by-product from *o*XDA undergoes a hydride shift, forming isoindole which spontaneously polymerizes. This can lead to clogging of reactors on process scale and complicates the work-up. Trimerization of 1-piperidine, the by-product from cadaverine, was shown to be most effective at pH > 10, limiting its applicability to ATAs that can tolerate such a high pH.

The reversibility and symmetry of the reaction catalysed by transaminases introduces a further problem, the equilibrium is often unfavourable, necessitating the use of the amine donor in excess (reducing efficiency) and product removal. Strategies for product removal either focus on the by-product, or the desired product. For the latter, in-situ product crystallization is being developed by the Hülsewede *et al.*^{325,326} however appears to be limited to certain substrate-amine donor combinations, as well as requiring extensive process engineering. More common is the removal of the by-product (**Scheme 1-16**). This may be

done enzymatically, through evaporation, or by designing further reactivity into the by-product (“smart” amine donor). Using alanine as the donor, the formed pyruvate may be removed using either pyruvate decarboxylase,³²⁷ lactate dehydrogenase (LDH),³²⁸ or alanine dehydrogenase (AlaDH, which rather than removing pyruvate, recycles it into alanine).³²⁹ The last two examples require NAD(P)H as the cofactor, which may be recycled by glucose dehydrogenase (GDH) and glucose, or formate dehydrogenase (FDH) and formate (**Scheme 1-16B**).

Using isopropylamine (IPA) as the donor, the produced acetone may be evaporated, provided the desired product is not volatile. IPA is small, cheap, very soluble in water (allowing large excesses to be used) and the acetone removal requires no additional enzymes or chemicals, making it the preferred choice for industry (both the Sitagliptin and vernakalant examples employed it). Lastly, “smart” amine donors^{318,330} are designed in such a way that the ketone product further reacts with itself, removing itself from the equilibrium. Specifically, these are diamines that cyclize and then either oligomerize or polymerize (**Scheme 1-16C**). However, they are relatively expensive, and purification or scale-up might be further complicated.²⁹⁵

Transaminases have also found extensive use in cascade reactions. Two examples have already been mentioned, the conversion of amines to (chiral) alcohols using an immobilized transaminase from *Halomonas elongata* (HEwT) and an ADH or KRed,²³⁰ and a Suzuki-Miyaura-ATA cascade.²³¹ Other examples (**Scheme 1-17**) include ATA triggered intramolecular aza-Michael cyclizations of ketoenones³³¹ and ketonones,³³² ATA and MAO-N catalysed formation of 2,5-disubstituted pyrrolidines,³³³ the synthesis of amines from fatty acids using a carboxylic acid reductase-transaminase cascade,³³⁴ and the synthesis of amines from alcohols using an ADH-ATA cascade.³³⁵ In the last example, pyruvate was recycled by an AlaDH, whose cofactor NADH was in turn recycled by the first half reaction of the ADH, an example of a hydrogen-borrowing cascade.



Scheme 1-17: Examples of cascade reactions using transaminases. Highlighted in blue: the carbonyl being transaminated (where ambiguous). PAP/Adk: ATP recycling system.

1.3 Aims

As described above, chiral amines are key synthetic targets; and biocatalytic strategies can often result in greener and more effective syntheses. However, limited substrate scopes and challenges around intensification and scale-up result in a high cost of developing and implementing such processes and are key challenges facing biocatalysis. Transaminases in particular are a key class of enzymes in the synthesis of chiral amines. Previously successful implementations, such as the sitagliptin synthesis mentioned above, have focussed on bulky high-value compounds, where the profit margins can tolerate higher costs. However, the synthesis of smaller chiral building blocks is also attractive as they could then feed into the synthesis of several compounds.

Therefore, the use of transaminases in the synthesis of such (small) chiral amine building blocks is the central theme in this thesis. Transaminases will be investigated for the synthesis of some of the smallest chiral amines, such as 2-aminobutane, where the small and similar size of the substituents on the ketone may pose challenges to the usually excellent enantioselectivity exhibited by this class of enzyme. By using the techniques described above, in particular site-directed mutagenesis and enzyme immobilization, a suitable system for its synthesis on a multi-gram scale will be developed. The suitability of transaminases for the synthesis of other small chiral amines will also be investigated, in particular 2,2-dimethylhexan-3-amine and tetrahydrofuran-3-amine.

To further aid in the development of these processes, new (*R*)-selective transaminases will be recruited to the lab and characterized. Using extremophilic organisms as the source, these enzymes should have enhanced properties (enhanced stability) compared to mesophilic counterparts which make them more suitable for use in biocatalysis. Structural analysis and rational mutagenesis will be used to further enhance the stability.

Lastly, the combination of transaminases with Buchwald-Hartwig amination (BHA) to create chiral *N*-arylamine building blocks in a sequential cascade will be investigated. Here, the compatibility of the reaction environment of the transaminase and palladium catalyst, the interference of the amine donor in the BHA, as well as the different substrate concentrations for each reaction pose challenges that will need to be overcome for such a process to be efficient and effective.

1.4 References

1. Armstrong, E. F. Enzymes: A Discovery and its Consequences. *Nature* **131**, 535–537 (1933).
2. Büchner, E. Eduard Büchner – Nobel Lecture: Cell-free fermentation. *Nobel lectures: chemistry 1901-1921*. <https://www.nobelprize.org/prizes/chemistry/1907/buchner/lecture/> (1907).
3. Kühne, W. Über das Verhalten verschiedener organisirter und sog. ungeformter Fermente. in *Verhandlungen des naturhistorisch-medicinischen Vereins zu Heidelberg* 190–193 (1877).
4. Robertson, A. J. B. The Early History of Catalysis. *Platinum Met. Rev.* **19**, 64–69 (1975).
5. Northrop, J. H. John H. Northrop – Nobel Lecture: The preparation of pure enzymes and virus proteins. *Nobel Lectures, Chemistry 1942-1962*. <https://www.nobelprize.org/prizes/chemistry/1946/northrop/lecture/> (1946).
6. Manchester, K. L. Louis Pasteur (1822-1895) - chance and the prepared mind. *Trends Biotechnol.* **13**, 511–515 (1995).
7. Kohler, R. The Background to Eduard Buchner's Discovery of Cell-Free Fermentation. *J. Hist. Biol.* **4**, 35–61 (1971).
8. Hein, G. E. The Liebig-Pasteur Controversy Vitality without vitalism. *J. Chem. Educ.* **38**, 614–619 (1961).
9. Sumner, J. B. . James B. Sumner – Nobel Lecture: The chemical nature of enzymes. *Nobel Lectures, Chemistry 1942-1962* <https://www.nobelprize.org/prizes/chemistry/1946/sumner/lecture/> (1946).
10. Von Loesecke, H. W. A review of information on: Mycological citric acid production. *Chem. Eng. News* **23**, 1952–1959 (1945).
11. Tauber, H. *The chemistry and technology of enzymes*. (Wiley, 1949).
12. Hildebrandt, G. & Klavehn, W. Manufacture of laevo-1-phenyl-2-methylaminopropanol-1. (1934).
13. Maroney, K. A. N., Culshaw, P. N., Wermuth, U. D. & Cresswell, S. L. Investigation of the l-phenylacetylcarbinol process to substituted benzaldehydes of interest. *Forensic Sci. Int.* **235**, 52–61 (2014).
14. Breuer, M., Pohl, M., Hauer, B. & Lingen, B. High-throughput assay of (R)-phenylacetylcarbinol synthesized by pyruvate decarboxylase. *Anal. Bioanal. Chem.* **374**, 1069–1073 (2002).
15. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* **27**, 2985–2993 (1894).
16. Johnson, K. A. & Goody, R. S. The original Michaelis constant: Translation of the 1913 Michaelis-Menten Paper. *Biochemistry* **50**, 8264–8269 (2011).
17. Pauling, L. Nature of forces between Large molecules of Biological interest. *Nature* **161**, 707–709 (1948).
18. Eyeing, H. The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **17**, 65–77 (1935).
19. Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **44**, 98–104 (1958).
20. Koshland, D. E. Enzyme flexibility and enzyme action. *J. Cell. Comp. Physiol.* **54**, 245–258 (1959).
21. Fischer, E. Synthese von Polypeptiden. XVII. *Berichte der Dtsch. Chem. Gesellschaft* **40**, 1754–1767 (1907).
22. Fruton, J. S. Early Theories of Protein Structure. *Ann. N.Y. Acad. Sci.* **325**, 1–20 (1979).
23. Sanger, F. & Tuppy, H. The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *Biochem. J* **49**, 463–481 (1951).

24. Sanger, F. & Tuppy, H. The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates. *Biochem. J* **49**, 481–490 (1951).
25. Pauling, L. & Corey, R. B. Two hydrogen-bonded spiral configurations of the polypeptide chain. *J. Am. Chem. Soc.* **72**, 5349 (1950).
26. Pauling, L. & Corey, R. B. Configurations of Polypeptide Chains with Favored Orientations Around Single Bonds: Two New Pleated Sheets. *Proc. Natl. Acad. Sci. U. S. A.* **37**, 729–740 (1951).
27. Pauling, L. Linus Pauling – Nobel Lecture: Modern structural chemistry. *Nobel Lectures, Chemistry 1942-1962* <https://www.nobelprize.org/prizes/chemistry/1954/pauling/lecture/> (1954).
28. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **37**, 205–211 (1951).
29. Perutz, M. F. New X-Ray Evidence on the Configuration of Polypeptide Chains: Polypeptide Chains in Poly- γ -benzyl-L-glutamate, Keratin and Hæmoglobin. *Nature* **167**, 1053–1054 (1951).
30. Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666 (1958).
31. Kendrew, J. C. *et al.* Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **185**, 422–427 (1960).
32. Perutz, M. F. *et al.* Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **185**, 416–422 (1960).
33. Perutz, M. F. X-ray Analysis of Hemoglobin. *Science* **140**, 863–870 (1963).
34. Kendrew, J. C. Myoglobin and the structure of proteins. *Science* **139**, 1259–1266 (1963).
35. de Chadarevian, S. John Kendrew and myoglobin: Protein structure determination in the 1950s. *Protein Sci.* **27**, 1136–1143 (2018).
36. Bragg, L. First stages in the x-ray analysis of proteins. *Rep. Prog. Phys.* **28**, 1–14 (1965).
37. Jaskolski, M., Dauter, Z. & Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J.* **281**, 3985–4009 (2014).
38. Blake, C. C. F. *et al.* The three-dimensional structure of hen eggwhite lysozyme. *Nature* **206**, 757–761 (1965).
39. Johnson, L. N. & Phillips, D. C. Structure of Some Crystalline Lysozyme-Inhibitor Complexes Determined by X-Ray Analysis At 6 Å Resolution. *Nature* **206**, 761–763 (1965).
40. Reeke, G. N. *et al.* The structure of carboxypeptidase a, vi. Some results at 2.0-Å resolution, and the complex with glycyl-tyrosine at 2.8-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **58**, 2220–2226 (1967).
41. Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. Three-dimensional structure of tosyl- α -chymotrypsin. *Nature* **214**, 652–656 (1967).
42. Blow, D. M., Birktoft, J. J. & Hartley, B. S. Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* **221**, 337–340 (1969).
43. Henderson, R. Structure of crystalline α -chymotrypsin. IV. The structure of indoleacryloyl- α -chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme. *J. Mol. Biol.* **54**, 341–354 (1970).
44. Blow, D. M. Structure and Mechanism of Chymotrypsin. *Acc. Chem. Res.* **9**, 145–152 (1976).
45. PDB Statistics: Overall Growth of Released Structures Per Year. <https://www.rcsb.org/stats/growth/growth-released-structures>.
46. Bugg, T. D. H. The development of mechanistic enzymology in the 20th century. *Nat. Prod. Reports* **18**, 465–493 (2001).

Chapter 1 — Introduction

47. Westheimer, F. H. The Discovery of the Mechanisms of Enzyme Action, 1947–1963. *Adv. Phys. Org. Chem.* **21**, 1–35 (1985).
48. Warburg, O. & Christian, W. Pyridin, der wasserstoffübertragende Bestandteil von Gärungsfermenten. *Helv. Chim. Acta* **19**, E79–E88 (1936).
49. Williams, R. R. & Cline, J. K. Synthesis of vitamin B1. *J. Am. Chem. Soc.* **58**, 1504–1505 (1936).
50. Snell, E. E., Guirard, B. M. & Williams, R. J. Occurrence in Natural Products of a Physiologically Active Metabolite of Pyridoxine. *J. Biol. Chem.* **143**, 519–531 (1942).
51. Snell, E. E. The Vitamin B6 group. I. Formation of Additional Members from Pyridoxine and Evidence Concerning their Structure. *J. Am. Chem. Soc.* **66**, 2082–2088 (1944).
52. Harris, S. A., Heyl, D. & Folkers, K. The Vitamin B6 group. II. The Structure and Synthesis of Pyridoxamine and Pyridoxal. *J. Am. Chem. Soc.* **66**, 2088–2092 (1944).
53. Braunstein, A. E. & Kritzmann, M. G. Co-Aminopherase, Co-Decar-Boxylase and Pyridoxal. *Nature* **158**, 102–103 (1946).
54. Braunstein, A. E. & Kritzmann, M. G. Formation and breakdown of Aminoacids by intermolecular transfer of the amino group. *Nature* **140**, 503–504 (1937).
55. Haynes, L. J. & Todd, A. R. Codehydrogenases. Part I. The synthesis of dihydronicotinamide-D-ribofuranoside [N-D-ribofuranosidyl-1 : 2(or 6)-dihydronicotinamide]. *J. Chem. Soc.* 303–308 (1950) doi:<https://doi.org/10.1039/JR9500000303>.
56. Haynes, L. J., Hughes, N. A., Kenner, G. W. & Todd, A. Codehydrogenases. Part II. A synthesis of nicotinamide nucleotide. *J. Chem. Soc.* 3727–3732 (1957) doi:<https://doi.org/10.1039/JR9570003727>.
57. N. A. Hughes, Kenner, G. W. & Todd, A. Codehydrogenases. Part III. A synthesis of diphosphopyridine nucleotide (cozymase), and some observations on the synthesis of triphosphopyridine nucleotide. *J. Chem. Soc.* 3733–3738 (1957) doi:<https://doi.org/10.1039/JR9570003733>.
58. Baddiley, J. & Todd, A. R. Nucleotides. Part I. Muscle adenylic acid and adenosine diphosphate. *J. Chem. Soc.* 648–651 (1947) doi:<https://doi.org/10.1039/JR9470000648>.
59. Baddiley, J., Michelson, A. M. & Todd, A. R. Nucleotides. Part II. A synthesis of adenosine triphosphate. *J. Chem. Soc.* 582–586 (1949) doi:<https://doi.org/10.1039/JR9490000582>.
60. Christie, S. M. H., Kenner, G. W. & Todd, A. R. Total Synthesis of Flavin-adenine-dinucleotide. *Nature* **170**, 924 (1952).
61. Christie, S. M. H., Kenner, G. W. & Todd, A. R. Nucleotides. Part XXV. A synthesis of flavin-adenine dinucleotide. *J. Chem. Soc.* 46–52 (1954) doi:<https://doi.org/10.1039/JR9540000046>.
62. Hodgkin, D. C. *et al.* Structure of vitamin B12: The crystal structure of the hexacarboxylic acid derived from B12 and the molecular structure of the vitamin. *Nature* **176**, 325–328 (1955).
63. Hodgkin, D. C. *et al.* Structure of Vitamin B12. *Nature* **178**, 64–66 (1956).
64. Loewus, F. A., Westheimer, F. H. & Vennessland, B. Enzymatic Synthesis of the Enantiomorphs of Ethanol-1-d. *J. Am. Chem. Soc.* **75**, 5018–5023 (1953).
65. Westheimer, F. H., Fisher, H. F., Conn, E. E. & Vennessland, B. The enzymatic transfer of hydrogen from alcohol to dpn. *J. Am. Chem. Soc.* **73**, 2403 (1951).
66. Breslow, R. Rapid Deuterium Exchange in Thiazolium Salts. *J. Am. Chem. Soc.* **79**, 1762–1763 (1957).
67. Breslow, R. On the Mechanism of Thiamine Action. IV. Evidence from Studies on Model Systems. *J. Am. Chem. Soc.* **80**, 3719–3726 (1958).
68. Snell, E. E. The Vitamin B6 Group. V. The Reversible Interconversion of Pyridoxal and Pyridoxamine by Transamination Reactions. *J. Am. Chem. Soc.* **67**, 194–197 (1945).

69. Metzler, D. E., Ikawa, M. & Snell, E. E. A General Mechanism for Vitamin B6-catalyzed Reactions. *J. Am. Chem. Soc.* **76**, 648–652 (1954).
70. Snell, E. E. The Vitamin Activities of 'Pyridoxal' and 'Pyridoxamin'. *J. Biol. Chem.* **154**, 313–314 (1944).
71. Bruggink, A. *Synthesis of β -lactam antibiotics*. *Synthesis of β -lactam antibiotics* (2001). doi:10.1007/978-94-010-0850-1.
72. Grubhofer, N. & Schleith, L. Modifizierte Ionenaustauscher als spezifische Adsorbentien. *Naturwissenschaften* **40**, 508 (1953).
73. Michee, F. & Dinkloh, E. Kohlenhydrat-Eiweißverbindungen aus Proteinen und 1-Fluor-d-glucose (Kohlenhydrat-Eiweißverbindungen, VIII. Mitteil.). *Chem. Ber.* **84**, 210–212 (1951).
74. Manecke, G., Singer, S. & Gillert, K.-E. Serologisch spezifische Adsorbentien. *Naturwissenschaften* **45**, 440–441 (1958).
75. Bornscheuer, U. T. & Buchholz, K. Highlights in biocatalysis - Historical landmarks and current trends. *Eng. Life Sci.* **5**, 309–323 (2005).
76. Elander, R. P. Industrial production of β -lactam antibiotics. *Applied Microbiology and Biotechnology* vol. 61 385–392 (2003).
77. Messing, R. A. & Filbert, A. M. Immobilized Glucose Isomerase for the Continuous Conversion of Glucose to Fructose. *J. Agric. Food. Chem.* **23**, 920–923 (1975).
78. Hanover, M. L. & White, J. S. Manufacturing, composition, and applications of fructose. *Am. J. Clin. Nutr.* **58**, 724S-732S (1993).
79. Bruggink, A., Roos, E. C. & De Vroom, E. Penicillin acylase in the industrial production of β -lactam antibiotics. *Org. Process Res. Dev.* **2**, 128–133 (1998).
80. Buchholz, K. A breakthrough in enzyme technology to fight penicillin resistance — industrial application of penicillin amidase. *Appl Microbiol Biotechnol* **100**, 3825–3839 (2016).
81. Hsiao, H. Y., Walter, J. F., Anderson, D. M. & Hamilton, B. K. Enzymatic production of amino acids. *Biotechnol. Genet. Eng. Rev.* **6**, 179–219 (1988).
82. Heckmann, C. M. Construction of Expression Vectors for Selective His-tag Removal. (University of Nottingham, 2017).
83. Dawes, H. The quiet revolution. *Curr. Biol.* **14**, R605–R607 (2004).
84. Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids. *Nature* vol. 171 737–738 (1953).
85. Anfinsen, C. B., Haber, E., Sela & White, F. H. the Kinetics of Formation of Native Ribonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1309–1314 (1961).
86. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
87. Gros, F. *et al.* Unstable ribonucleic acid revealed by pulse labelling of Escherichia coli. *Nature* **190**, 581–585 (1961).
88. Brenner, S., Jacob, F. & Meselson, M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**, 576–581 (1961).
89. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* **192**, 1227–1232 (1961).
90. Nirenberg, M. W. & Matthaei, J. H. The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1588–1602 (1961).
91. Nirenberg, M. W., Matthaei, J. H. & Jones, O. W. An intermediate in the biosynthesis of polyphenylalanine directed by synthetic template RNA. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 104–109 (1962).

Chapter 1 — Introduction

92. Leder, P. & Nirenberg, M. Rna Codewords and Protein Synthesis. Ii. Nucleotide Sequence of a Valine RNA Codeword. *Proc. Natl. Acad. Sci. U. S. A.* **52**, 420–427 (1964).
93. Lengyl, P., Speyer, J. F. & Ochoa, S. Synthetic polynucleotides and the amino acid code. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1936–1942 (1961).
94. Speyer, J. F., Lengyl, P., Basilio, C. & Ochoa, S. Synthetic polynucleotides and the amino acid code. II. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 63–68 (1962).
95. Söll, D. *et al.* Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinnucleotides and a survey of codon assignments for 20 amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **54**, 1378–1385 (1965).
96. Nirenberg, M. Historical review: Deciphering the genetic code - A personal account. *Trends Biochem. Sci* **29**, 46–54 (2004).
97. Marshall, R. E., Caskey, C. T. & Nirenberg, M. Fine Structure of RNA Codewords Recognized by Bacterial, Amphibian, and Mammalian Transfer RNA. *Science* **155**, 820–826 (1967).
98. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).
99. Rosenblum, B. *et al.* New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* **25**, 4500–4504 (1997).
100. Lee, L. G. *et al.* New energy transfer dyes for DNA sequencing. *Nucleic Acids Res.* **25**, 2816–2822 (1997).
101. Lederberg, J. Plasmid (1952 – 1997). *Plasmid* **39**, 1–9 (1998).
102. Lederberg, J. Cell genetics and hereditary symbiosis. *Physiol. Rev.* **32**, 403–430 (1952).
103. Luria, S. E. & Human, M. L. A nonhereditary, host-induced variation of bacterial viruses. *J. Bacteriol.* **64**, 557–569 (1952).
104. Bertani, G. & Weigle, J. J. Host controlled variation in bacterial viruses. *J. Bacteriol.* **65**, 113–121 (1953).
105. Arber, W. Host-Controlled Modification of Bacteriophage. *Annu. Rev. Microbiol.* **19**, 365–378 (1965).
106. Linn, S. & Arber, W. Host specificity of DNA produced by Escherichia coli, X. In vitro restriction of phage fd replicative form. *Proc. Natl. Acad. Sci. U. S. A.* **59**, 1300–1306 (1968).
107. Arber, W. & Lin, S. DNA Modification and Restriction. *Annu. Rev. Biochem* **38**, 467–500 (1969).
108. Meselson, M. & Yuan, R. DNA Restriction Enzyme from E Coli. *Nature* **217**, 1110–1114 (1968).
109. Nathans, D. & Smith, H. O. Restriction endonucleases in the analysis and restructuring of dna molecules. *Annu. Rev. Biochem* **44**, 273–293 (1975).
110. Smith, H. O. & Welcox, K. W. A Restriction enzyme from Hemophilus influenzae I. *J. Mol. Biol.* **51**, 379–391 (1970).
111. Kelly, T. J. & Smith, H. O. A restriction enzyme from Hemophilus influenzae II. *J. Mol. Biol.* **51**, 393–409 (1970).
112. Roberts, R. J. How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 5905–5908 (2005).
113. Danna, K. & Nathans, D. Specific cleavage of simian virus 40 DNA by restriction endonuclease of Hemophilus influenzae. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 2913–2917 (1971).
114. Mertz, J. E. & Davis, R. W. Cleavage of DNA by R 1 restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3370–3374 (1972).
115. Jackson, D. A., Symons, R. H. & Berg, P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose

- operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 2904–2909 (1972).
116. Brown, T. A. *Gene Cloning and DNA Analysis: An Introduction*. (John Wiley & Sons Ltd, 2016).
 117. Lobban, P. E. & Kaiser, A. D. Enzymatic end-to-end joining of DNA molecules. *J. Mol. Biol.* **78**, 453–471 (1973).
 118. Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 3240–3244 (1973).
 119. Bolivar, F., Rodriguez, R. L., Betlach, M. C. & Boyer, H. W. Construction and Characterization of New Cloning Vehicles I. Ampicillin-Resistant Derivatives of the Plasmid pMB9. *Gene* **2**, 75–93 (1977).
 120. Bolivar, F. *et al.* Construction and Characterization of New Cloning Vehicles II. a Multipurpose Cloning System. *Gene* **2**, 95–113 (1977).
 121. Vieira, J. & Messing, J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268 (1982).
 122. Yanisch-Perron, C., Vieira, J. & Messing, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119 (1985).
 123. Rosenberg, A. H. *et al.* Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125–135 (1987).
 124. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89 (1990).
 125. Studier, F. W. & Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. *J. Mol. Biol.* **189**, 113–130 (1986).
 126. Novy, R. & Morris, B. Use of glucose to control basal expression in the pET System. *Innovations* **13**, 13–15 (2001).
 127. Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S. & Kim, J. F. Tracing Ancestors and Relatives of *Escherichia coli* B, and the Derivation of B Strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 634–643 (2009).
 128. Terpe, K. Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **72**, 211–222 (2006).
 129. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **5**, 172 (2014).
 130. Johnson, I. S. Human insulin from recombinant DNA technology. *Science* **219**, 632–637 (1983).
 131. Gualandi-Signorini, A. M. & Giorgi, G. Insulin formulations - A review. *Eur. Rev. Med. Pharmacol. Sci.* **5**, 73–83 (2001).
 132. Marsden, S. R., Mestrom, L., McMillan, D. G. G. & Hanefeld, U. Thermodynamically and Kinetically Controlled Reactions in Biocatalysis – from Concepts to Perspectives. *ChemCatChem* **12**, 426–437 (2020).
 133. Harris, T. J. R. *et al.* Molecular cloning and nucleotide sequence of cDNA coding for calf preprochymosin. *Nucleic Acids Res.* **10**, 2177–2187 (1982).
 134. Emtage, J. S. *et al.* Synthesis of calf prochymosin (prorennin) in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3671–3675 (1983).
 135. Hicks, C. L., O’Leary, J. & Bucy, J. Use of Recombinant Chymosin in the Manufacture of Cheddar and Colby Cheese. *J. Dairy Sci.* **71**, 1127–1131 (1988).
 136. Kumar, A., Grover, S., Sharma, J. & Batish, V. K. Chymosin and other milk coagulants: sources and biotechnological interventions. *Crit. Rev. Biotechnol.* **30**, 243–58 (2010).
 137. Johnson, M. E. & Lucey, J. A. Major technological advances and trends in cheese. *J. Dairy Sci.* **89**, 1174–

Chapter 1 — Introduction

- 1178 (2006).
138. Bornscheuer, U. T. *et al.* Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
 139. Carlson, E. A. H.J. Muller's contributions to mutation research. *Mutat. Res. - Rev. Mutat. Res.* **752**, 1–5 (2013).
 140. Witkin, E. M. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* **40**, 869–907 (1976).
 141. Muller, H. J. Artificial transmutation of the gene. *Science* **66**, 84–87 (1927).
 142. Crow, J. F. & Abrahamson, S. Seventy years ago: Mutation becomes experimental. *Genetics* **147**, 1491–1496 (1997).
 143. Lederberg, J. & Lederberg, E. M. Replica plating and indirect selection of bacterial mutants. *J. Bacteriol.* **63**, 399–406 (1952).
 144. Adrio, J. L. & Demain, A. L. Genetic improvement of processes yielding microbial products. *FEMS Microbiol. Rev.* **30**, 187–214 (2006).
 145. Lerner, S. A., Wu, T. T. & Lin, E. C. C. Evolution of a catabolic pathway in bacteria. *Science* **146**, 1313–1315 (1964).
 146. Parekh, S., Vinci, V. A. & Strobel, R. J. Improvement of microbial strains and fermentation processes. *Appl. Microbiol. Biotechnol.* **54**, 287–301 (2000).
 147. Hutchison, C. A. *et al.* Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **253**, 6551–6560 (1978).
 148. Khorana, H. G. Nucleic acid synthesis. *Pure Appl. Chem.* **17**, 349–382 (1968).
 149. Gillam, S., Rottman, F., Jahnke, P. & Smith, M. Enzymatic synthesis of oligonucleotides of defined sequence : Synthesis of a segment of yeast iso-i-cytochrome c gene Biochemistry. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 96–100 (1977).
 150. Matteucci, M. D. & Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J. Am. Chem. Soc.* **103**, 3185–3191 (1981).
 151. Beaucage, S. L. & Iyer, R. P. Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* **48**, 2223–2311 (1992).
 152. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. Redesigning enzyme structure by site-directed mutagenesis: Tyrosyl tRNA synthetase and ATP binding. *Nature* **299**, 756–758 (1982).
 153. Brannigan, J. A. & Wilkinson, A. J. Protein engineering 20 years on. *Nat. Rev. Mol. cell Biol.* **3**, 964–70 (2002).
 154. Estell, D. A., Graycar, T. P. & Wells, J. A. Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.* **260**, 6518–6521 (1985).
 155. Quax, W. J. *et al.* Enhancing the thermostability of glucose isomerase by protein engineering. *Bio/Technology* **9**, 738–742 (1991).
 156. *Methods in Molecular Biology: PCR Protocols.* vol. 226 (Humana Press Inc., 2003) doi: 10.1385/1592593844.
 157. Saiki, R. K. *et al.* Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350–1354 (1985).
 158. Saik, R. K. *et al.* Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science* **239**, 487–491 (1988).
 159. Arnold, F. H. Innovation by Evolution: Bringing New Chemistry to Life (Nobel Lecture). *Angew. Chemie Int. Ed.* **58**, 14420–14426 (2019).

160. Chen, K. & Arnold, F. H. Enzyme engineering for nonaqueous solvents: Random mutagenesis to enhance activity of subtilisin E in Polar Organic Media. *Bio/Technology* **9**, 1073–1077 (1991).
161. Chen, K. & Arnold, F. H. Tuning the activity of an enzyme for unusual environments: Sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5618–5622 (1993).
162. Stemmer, W. P. C. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391 (1994).
163. Stemmer, W. P. C. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10747–10751 (1994).
164. Moore, J. C. & Arnold, F. H. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467 (1996).
165. Cherry, J. R. *et al.* Directed evolution of a fungal peroxidase. *Nat. Biotechnol.* **17**, 379–384 (1999).
166. Zhao, H., Giver, L., Shao, Z., Affholter, J. A. & Arnold, F. H. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**, 258–261 (1998).
167. Ness, J. E. *et al.* DNA shuffling of subgenomic sequences of subtilisin. *Nat. Biotechnol.* **17**, 893–896 (1999).
168. Giver, L., Gershenson, A., Freskgard, P. O. & Arnold, F. H. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12809–12813 (1998).
169. Taguchi, S., Ozaki, A. & Momose, H. Engineering of a cold-adapted protease by sequential random mutagenesis and a screening system. *Appl. Environ. Microbiol.* **64**, 492–495 (1998).
170. Kumamaru, T., Suenaga, H., Mitsuoka, M., Watanabe, T. & Furukawa, K. Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat. Biotechnol.* **16**, 663–666 (1998).
171. Yano, T., Oue, S. & Kagamiyama, H. Directed evolution of an aspartate aminotransferase with new substrate specificities. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5511–5515 (1998).
172. May, O., Nguyen, P. T. & Arnold, F. H. Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. *Nat. Biotechnol.* **18**, 317–320 (2000).
173. Liebeton, K. *et al.* Directed evolution of an enantioselective lipase. *Chem. Biol.* **7**, 709–718 (2000).
174. Arnold, F. H. Design by directed evolution. *Acc. Chem. Res.* **31**, 125–131 (1998).
175. Arnold, F. H. & Volkov, A. A. Directed evolution of biocatalysts. *Curr. Opin. Chem. Biol.* **3**, 54–59 (1999).
176. Reetz, M. T. & Jaeger, K.-E. Superior Biocatalysts by Directed Evolution. *Top. Curr. Chem.* **200**, 31–57 (1999).
177. Arnold, F. H. Combinatorial and computational challenges for biocatalyst design. *Nature* **409**, 253–257 (2001).
178. Cobb, R. E., Chao, R. & Zhao, H. Directed evolution: Past, present, and future. *AIChE J.* **59**, 1432–1440 (2013).
179. Schmid, A. *et al.* Industrial biocatalysis today and tomorrow. *Nature* **409**, 258–268 (2001).
180. Schoemaker, H. E., Mink, D. L. & Wubbolts, M. G. Dispelling the myths - Biocatalysis in industrial synthesis. *Science* **299**, 1694–1697 (2003).
181. Murphy, M. A. Early Industrial Roots of Green Chemistry and the history of the BHC Ibuprofen process invention and its Quality connection. *Found. Chem.* **20**, 121–165 (2018).
182. Anastas, P. T., Kirchhoff, M. M. & Williamson, T. C. Catalysis as a foundational pillar of green chemistry. *Appl. Catal. A Gen.* **221**, 3–13 (2001).

Chapter 1 — Introduction

183. Trost, B. M. The atom economy - A search for synthetic efficiency. *Science* **254**, 1471–1477 (1991).
184. Anastas, P. T. & Kirchhoff, M. M. Origins, current status, and future challenges of green chemistry. *Acc. Chem. Res.* **35**, 686–694 (2002).
185. Linthorst, J. A. An overview: Origins and development of green chemistry. *Found. Chem.* **12**, 55–68 (2010).
186. Ni, Y., Holtmann, D. & Hollmann, F. How green is biocatalysis? to calculate is to know. *ChemCatChem* **6**, 930–943 (2014).
187. Sheldon, R. A. & Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **118**, 801–838 (2018).
188. Straathof, A. J. J., Panke, S. & Schmid, A. The production of fine chemicals by biotransformations. *Curr. Opin. Biotechnol.* **13**, 548–556 (2002).
189. Sheldon, R. A., Brady, D. & Bode, M. L. The Hitchhiker’s Guide to Biocatalysis: Recent Advances in the Use of Enzymes in Organic Synthesis. *Chem. Sci.* **11**, 2587–2605 (2020).
190. Woodley, J. M. Accelerating the implementation of biocatalysis in industry. *Appl. Microbiol. Biotechnol.* **103**, 4733–4739 (2019).
191. Balkenhohl, F., Ditrich, K., Hauer, B. & Aktiengesellschaft, B. Optisch aktive Amine durch Lipase-katalysierte Methoxyacetylierung. *J. für Prakt. Chemie* **339**, 381–384 (1997).
192. Reetz, M. T. Biocatalysis in organic chemistry and biotechnology: Past, present, and future. *J. Am. Chem. Soc.* **135**, 12480–12496 (2013).
193. Breuer, M. *et al.* Industrial methods for the production of optically active intermediates. *Angew. Chemie Int. Ed.* **43**, 788–824 (2004).
194. Basso, A. & Serban, S. Industrial applications of immobilized enzymes—A review. *Mol. Catal.* **479**, 110607 (2019).
195. Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L. & Arnold, F. H. Protein building blocks preserved by recombination. *Nat. Struct. Biol.* **9**, 553–558 (2002).
196. Bommarius, A. S., Blum, J. K. & Abrahamson, M. J. Status of protein engineering for biocatalysts: How to design an industrially useful biocatalyst. *Curr. Opin. Chem. Biol.* **15**, 194–200 (2011).
197. Cobb, R. E., Si, T. & Zhao, H. Directed evolution: An evolving and enabling synthetic biology tool. *Curr. Opin. Chem. Biol.* **16**, 285–291 (2012).
198. Moore, J. C. *et al.* ‘site and Mutation’-Specific Predictions Enable Minimal Directed Evolution Libraries. *ACS Synth. Biol.* **7**, 1730–1741 (2018).
199. Dalby, P. A. Strategy and success for the directed evolution of enzymes. *Curr. Opin. Struct. Biol.* **21**, 473–480 (2011).
200. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angew. Chemie Int. Ed.* **44**, 4192–4196 (2005).
201. Reetz, M. T., Wang, L. W. & Bocola, M. Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space. *Angew. Chemie Int. Ed.* **45**, 1236–1241 (2006).
202. Reetz, M. T. *et al.* Expanding the substrate scope of enzymes: Combining mutations obtained by CASTing. *Chem. Eur. J* **12**, 6031–6038 (2006).
203. Reetz, M. T. What are the Limitations of Enzymes in Synthetic Organic Chemistry? *Chem. Rec.* **16**, 2449–2459 (2016).
204. Vavra, O. *et al.* CaverDock: A molecular docking-based tool to analyse ligand transport through protein tunnels and channels. *Bioinformatics* **35**, 4986–4993 (2019).

205. Yang, K. K., Wu, Z. & Arnold, F. H. Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **16**, 687–694 (2019).
206. Fox, R. Directed molecular evolution by machine learning and the influence of nonlinear interactions. *J. Theor. Biol.* **234**, 187–199 (2005).
207. Fox, R. J. *et al.* Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344 (2007).
208. Savile, C. K. *et al.* Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **329**, 305–310 (2010).
209. Schober, M. *et al.* Chiral synthesis of LSD1 inhibitor GSK2879552 enabled by directed evolution of an imine reductase. *Nat. Catal.* **2**, 909–915 (2019).
210. Roiban, G. D. *et al.* Efficient Biocatalytic Reductive Aminations by Extending the Imine Reductase Toolbox. *ChemCatChem* **9**, 4475–4479 (2017).
211. Kan, S. B. J., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome c for carbon–silicon bond formation: Bringing silicon to life. *Science* **354**, 1048–1051 (2016).
212. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
213. Jiang, J., Chen, X., Zhang, D., Wu, Q. & Zhu, D. Characterization of (R)-selective amine transaminases identified by in silico motif sequence blast. *Appl. Microbiol. Biotechnol.* **99**, 2613–2621 (2015).
214. Steffen-Munsberg, F. *et al.* Connecting Unexplored Protein Crystal Structures to Enzymatic Function. *ChemCatChem* **5**, 150–153 (2013).
215. Britton, J., Majumdar, S. & Weiss, G. A. Continuous flow biocatalysis. *Chem. Soc. Rev.* **47**, 5891–5918 (2018).
216. Romero-Fernández, M. & Paradisi, F. General overview on immobilization techniques of enzymes for biocatalysis. in *Catalyst Immobilization: Methods and Applications* (eds. Benaglia, M. & Puglisi, A.) 409–435 (2019). doi:10.1002/9783527817290.ch12.
217. Thompson, M. P., Peñafiel, I., Cosgrove, S. C. & Turner, N. J. Biocatalysis Using Immobilized Enzymes in Continuous Flow for the Synthesis of Fine Chemicals. *Org. Process Res. Dev.* **23**, 9–18 (2019).
218. *Immobilization of Enzymes and Cells. Methods in Molecular Biology* vol. 1051 (Humana Press Inc., 2013) doi:10.1007/978-1-62703-550-7.
219. Sheldon, R. A. Enzyme immobilization: The quest for optimum performance. *Adv. Synth. Catal.* **349**, 1289–1307 (2007).
220. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* **6**, 1321–1325 (1988).
221. Engelmark Cassimjee, K. *et al.* A general protein purification and immobilization method on controlled porosity glass: Biocatalytic applications. *Chem. Commun.* **50**, 9134–9137 (2014).
222. Böhmer, W., Volkov, A., Engelmark Cassimjee, K. & Mutti, F. G. Continuous Flow Bioamination of Ketones in Organic Solvents at Controlled Water Activity using Immobilized ω -Transaminases. *Adv. Synth. Catal.* **362**, 1858–1867 (2020).
223. Böhmer, W. *et al.* Highly efficient production of chiral amines in batch and continuous flow by immobilized ω -transaminases on controlled porosity glass metal-ion affinity carrier. *J. Biotechnol.* **291**, 52–60 (2019).
224. Planchestainer, M., Padrosa, D. R., Contente, M. L. & Paradisi, F. Genetically fused T4L acts as a shield in covalent enzyme immobilisation enhancing the rescued activity. *Catalysts* **8**, 40 (2018).
225. Padrosa, D. R., Vitis, V. De, Contente, M. L., Molinari, F. & Paradisi, F. Overcoming water insolubility in flow: Enantioselective hydrolysis of Naproxen Ester. *Catalysts* **9**, 232 (2019).

Chapter 1 — Introduction

226. Anderson, G. P. *et al.* Oriented Immobilization of Single-Domain Antibodies Using SpyTag/SpyCatcher Yields Improved Limits of Detection. *Anal. Chem.* **91**, 9424–9429 (2019).
227. Wong, J. X., Gonzalez-Miro, M., Sutherland-Smith, A. J. & Rehm, B. H. A. Covalent Functionalization of Bioengineered Polyhydroxyalkanoate Spheres Directed by Specific Protein-Protein Interactions. *Front. Bioeng. Biotechnol.* **8**, 1–19 (2020).
228. Jensen, V. J. & Rugh, S. Industrial-Scale Production and Application of Immobilized Glucose Isomerase. *Methods Enzymol.* **136**, 356–370 (1987).
229. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial Biocatalytic Linear Cascades for Preparation of Organic Molecules. *Chem. Rev.* **118**, 270–348 (2018).
230. Contente, M. L. & Paradisi, F. Self-sustaining closed-loop multienzyme-mediated conversion of amines into alcohols in continuous reactions. *Nat. Catal.* **1**, 452–459 (2018).
231. Dawood, A. W. H., Bassut, J., de Souza, R. O. M. A. & Bornscheuer, U. Combination of the Suzuki-Miyaura-cross coupling reaction with engineered transaminases. *Chem. Eur. J* **24**, 16009–16013 (2018).
232. Latham, J. *et al.* Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C-H activation. *Nat. Commun.* **7**, 1–8 (2016).
233. Huffman, M. A. *et al.* Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019).
234. Fukuyama, K., Ohru, H. & Kuwahara, S. Synthesis of EFdA via a diastereoselective aldol reaction of a protected 3-keto furanose. *Org. Lett.* **17**, 828–831 (2015).
235. Vitaku, E., Smith, D. T. & Njardarson, J. T. Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals. *J. Med. Chem.* **57**, 10257–10274 (2014).
236. Bommarius, A. S. Biocatalysis: A Status Report. *Annu. Rev. Chem. Biomol. Eng.* **6**, 319–345 (2015).
237. Ghislieri, D. & Turner, N. J. Biocatalytic approaches to the synthesis of enantiomerically pure chiral amines. *Top. Catal.* **57**, 284–300 (2014).
238. Constable, D. J. C. *et al.* Key green chemistry research areas - A perspective from pharmaceutical manufacturers. *Green Chem.* **9**, 411–420 (2007).
239. Burkhardt, E. R. & Coleridge, B. M. Reductive amination with 5-ethyl-2-methylpyridine borane. *Tetrahedron Lett.* **49**, 5152–5155 (2008).
240. Abdel-Magid, A. F., Carson, K. G., Harris, B. D., Maryanoff, C. A. & Shah, R. D. Reductive Amination of Aldehydes and Ketones with Sodium Triacetoxyborohydride. Studies on Direct and Indirect Reductive Amination Procedures. *J. Org. Chem.* **61**, 3849–3862 (1996).
241. Baxter, E. W. & Reitz, A. B. Reductive Aminations of Carbonyl Compounds with Borohydride and Borane Reducing Agents. in *Organic Reactions* (John Wiley & Sons, Inc., 2004).
242. Alexander, E. R. & Wildman, R. B. Studies on the Mechanism of the Leuckart Reaction. *J. Am. Chem. Soc.* **70**, 1187–1189 (1948).
243. Leuckart, R. Ueber eine neue Bildungsweise von Tribenzylamin. *Ber. Dtsch. Chem. Ges.* **18**, 2341–2344 (1885).
244. Wang, Z. Leuckart Reaction. in *Comprehensive Organic Name Reactions and Reagents* (ed. Wang, Z.) 1737–1742 (John Wiley & Sons, Inc., 2010).
245. Cogan, D. A., Liu, G. & Ellman, J. Asymmetric synthesis of chiral amines by highly diastereoselective 1,2-additions of organometallic reagents to N-tert-butanefulfinyl imines. *Tetrahedron* **55**, 8883–8904 (1999).
246. Moss, Neil; Gauthier, Jean; Ferland, J.-M. The enantioselective Synthesis of neopentylamine derivatives. *Synlett* 142–144 (1995).

247. Robak, M. T., Herbage, M. A. & Ellman, J. A. Synthesis and applications of tert -butanesulfinamide. *Chem. Rev.* **110**, 3600–3740 (2010).
248. Xu, H. C., Chowdhury, S. & Ellman, J. A. Asymmetric synthesis of amines using tert-butanesulfinamide. *Nat. Protoc.* **8**, 2271–2280 (2013).
249. Wakayama, M. & Ellman, J. A. Recycling the tert-butanesulfinyl group in the synthesis of amines using tert-butanesulfinamide. *J. Org. Chem.* **74**, 2646–2650 (2009).
250. Ryan, J. Transaminase Triggered Aza-Michael Approach for the Enantioselective Synthesis of Chiral Alkaloids. (Manchester Metropolitan University, 2018).
251. Hsiao, Y. *et al.* Highly efficient synthesis of β -amino acid derivatives via asymmetric hydrogenation of unprotected enamines. *J. Am. Chem. Soc.* **126**, 9918–9919 (2004).
252. Gallardo-Donaire, J. *et al.* Direct Asymmetric Ruthenium-Catalyzed Reductive Amination of Alkyl-Aryl Ketones with Ammonia and Hydrogen. *J. Am. Chem. Soc.* **140**, 355–361 (2018).
253. Xie, J. H., Zhu, S. F. & Zhou, Q. L. Transition metal-catalyzed enantioselective hydrogenation of enamines and imines. *Chem. Rev.* **111**, 1713–1760 (2011).
254. Tang, W. & Xiao, J. Asymmetric hydrogenation of imines via metal-organo cooperative catalysis. *Synth.* **46**, 1297–1302 (2014).
255. Nugent, T. C. Asymmetric Reductive Amination. in *Chiral Amine Synthesis: Methods, Developments and Applications* (ed. Nugent, T. C.) 225–245 (Wiley-VCH, 2010). doi:10.1002/9783527629541.ch7.
256. Alexeeva, M., Enright, A., Dawson, M. J., Mahmoudian, M. & Turner, N. J. Deracemization of a - Methylbenzylamine Using an Enzyme Obtained by In Vitro Evolution. *Angew. Chem. Int. Ed.* **41**, 3177–3180 (2002).
257. Bailey, K. R., Ellis, A. J., Reiss, R., Snape, T. J. & Turner, N. J. A template-based mnemonic for monoamine oxidase (MAO-N) catalyzed reactions and its application to the chemo-enzymatic deracemisation of the alkaloid (\pm)-crispine A. *Chem. Commun.* **44**, 3640–3642 (2007).
258. Atkin, K. E. *et al.* The Structure of Monoamine Oxidase from *Aspergillus niger* Provides a Molecular Context for Improvements in Activity Obtained by Directed Evolution. *J. Mol. Biol.* **384**, 1218–1231 (2008).
259. Rowles, I., Malone, K. J., Etchells, L. L., Willies, S. C. & Turner, N. J. Directed evolution of the enzyme monoamine oxidase (mao-n): Highly efficient chemo-enzymatic deracemisation of the alkaloid (\pm)-crispine A. *ChemCatChem* **4**, 1259–1261 (2012).
260. Ghislieri, D. *et al.* Engineering an enantioselective amine oxidase for the synthesis of pharmaceutical building blocks and alkaloid natural products. *J. Am. Chem. Soc.* **135**, 10863–10869 (2013).
261. Köhler, V. *et al.* Enantioselective biocatalytic oxidative desymmetrization of substituted pyrrolidines. *Angew. Chemie Int. Ed.* **49**, 2182–2184 (2010).
262. Znabet, A. *et al.* A highly efficient synthesis of telaprevir by strategic use of biocatalysis and multicomponent reactions. *Chem. Commun.* **46**, 7918–7920 (2010).
263. Grogan, G. Synthesis of chiral amines using redox biocatalysis. *Curr. Opin. Chem. Biol.* **43**, 15–22 (2018).
264. 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.* **51**, 3969–3972 (2012).
265. Abrahamson, M. J., Wong, J. W. & Bommarius, A. S. The evolution of an amine dehydrogenase biocatalyst for the asymmetric production of chiral amines. *Adv. Synth. Catal.* **355**, 1780–1786 (2013).
266. Bommarius, B. R., Schürmann, M. & Bommarius, A. S. A novel chimeric amine dehydrogenase shows altered substrate specificity compared to its parent enzymes. *Chem. Commun.* **50**, 14953–14955 (2014).
267. Caparco, A. A. *et al.* Metagenomic Mining for Amine Dehydrogenase Discovery. *Adv. Synth. Catal.* **362**, 2427–2436 (2020).

Chapter 1 — Introduction

268. Knaus, T., Böhmer, W. & Mutti, F. G. Amine dehydrogenases: Efficient biocatalysts for the reductive amination of carbonyl compounds. *Green Chem.* **19**, 453–463 (2017).
269. Scrutton, N. S., Turner, N. J., Breuer, M., Mutti, F. G. & Knaus, T. Conversion of alcohols to enantiopure amines through dual-enzyme hydrogen-borrowing cascades. *Science* **349**, 1525–1529 (2015).
270. Tseliou, V., Knaus, T., Masman, M. F., Corrado, M. L. & Mutti, F. G. Generation of amine dehydrogenases with increased catalytic performance and substrate scope from ϵ -deaminating L-Lysine dehydrogenase. *Nat. Commun.* **10**, (2019).
271. Tseliou, V., Masman, M. F., Böhmer, W., Knaus, T. & Mutti, F. G. Mechanistic Insight into the Catalytic Promiscuity of Amine Dehydrogenases: Asymmetric Synthesis of Secondary and Primary Amines. *ChemBioChem* **20**, 800–812 (2019).
272. Sharma, M. *et al.* A Mechanism for Reductive Amination Catalyzed by Fungal Reductive Aminases. *ACS Catal.* **8**, 11534–11541 (2018).
273. Aleku, G. A. *et al.* A reductive aminase from *Aspergillus oryzae*. *Nat. Chem.* **9**, 961–969 (2017).
274. Mangas-Sanchez, J. *et al.* Asymmetric synthesis of primary amines catalyzed by thermotolerant fungal reductive aminases. *Chem. Sci.* **11**, 5052–5057 (2020).
275. Lenz, M., Borlinghaus, N., Weinmann, L. & Nestl, B. M. Recent advances in imine reductase-catalyzed reactions. *World J. Microbiol. Biotechnol.* **33**, 1–10 (2017).
276. Sharma, M., Mangas-Sanchez, J., Turner, N. J. & Grogan, G. NAD(P)H-Dependent Dehydrogenases for the Asymmetric Reductive Amination of Ketones: Structure, Mechanism, Evolution and Application. *Adv. Synth. Catal.* **359**, 2011–2025 (2017).
277. Wetzl, D. *et al.* Asymmetric Reductive Amination of Ketones Catalyzed by Imine Reductases. *ChemCatChem* **8**, 2023–2026 (2016).
278. Matzel, P., Gand, M. & Höhne, M. One-step asymmetric synthesis of (R)- and (S)-rasagiline by reductive amination applying imine reductases. *Green Chem.* **19**, 385–389 (2017).
279. Mangas-Sanchez, J. *et al.* Imine reductases (IREDs). *Curr. Opin. Chem. Biol.* **37**, 19–25 (2017).
280. Bornadel, A. *et al.* Technical Considerations for Scale-Up of Imine Reductase Catalyzed Reductive Amination: A Case Study. *Org. Process Res. Dev.* **23**, 1262–1268 (2019).
281. Kumar, R. *et al.* Biocatalytic Reductive Amination-Discovery to Commercial Manufacturing applied to Abrocitinib JAK1 inhibitor. *Preprint (version 1) available at Research Square* (2021) doi:10.21203/rs.3.rs-244144/v1.
282. Green, D. E., Leloir, L. F. & Nocito, V. Transaminases. *J. Biol. Chem.* **161**, 559–582 (1945).
283. Torchinsky, Y. M. Transamination: its discovery, biological and chemical aspects (1937-1987). *Trends Biochem. Sci.* **12**, 115–117 (1987).
284. Jenkins, W. T., Yphantis, D. A. & Sizer, I. W. Glutamic Aspartic Transaminase: I. Assay, Purification, and General Properties. *J. Biol. Chem.* **234**, 51–57 (1959).
285. Jenkins, W. T. & Sizer, I. W. Glutamic Aspartic Transaminase: IV. the Mechanism of Transamination. *J. Biol. Chem.* **235**, 620–624 (1960).
286. Lis, H., Fasella, P., Turano, C. & Vecchini, P. On the Mechanism of Action of Glutamic-Aspartic Transaminase: Intermediate Steps in the Reaction. *Biochim. Biophys. Acta* **45**, 529–536 (1960).
287. Lichstein, H. C., Gunsalus, I. C. & Umbreit, W. W. Function of the Vitamin B6 Group: Phosphate (Codecarboxylase) in Transamination. *J. Biol. Chem.* **161**, 311–320 (1945).
288. Meister, A., Sober, H. A. & Peterson, E. A. Studies On The Coenzyme Activation Glutamic-Aspartic Apotransaminase. *J. Biol. Chem.* **206**, 89–100 (1954).
289. Hayaishi, O., Nishizuka, Y., Tatibana, M., Takeshita, M. & Kuno, S. Enzymatic Studies on the Metabolism

- of β -Alanine. *J. Biol. Chem.* **236**, 781–790 (1961).
290. Kim, K. & Tchen, T. T. Putrescine — α -Ketoglutarate transaminase in *E. coli*. *Biochem. Biophys. Res. Commun.* **9**, 99–102 (1962).
 291. Yamada, H., Kimura, T., Tanaka, A. & Ogata, K. Amine Transaminase. *Biochem. Biophys. Res. Commun.* **28**, 443–450 (1964).
 292. Hasse, K. & Neumann, W. Amin- α -Ketosäure-transaminasen in Extrakten tierischer Organe. *Naturwissenschaften* **51**, 243–243 (1964).
 293. Yonaha, K. & Toyama, S. Monoamine Transamination Catalyzed By (ω -Amino Acid: Pyruvate Aminotransferase Of *Pseudomonas* Sp. F-126. *Agric. Biol. Chem.* **42**, 2363–2367 (1978).
 294. Malik, M. S., Park, E. & Shin, J. Features and technical applications of ω -transaminases. *Appl. Microbiol. Biotechnol.* **94**, 1163–1171 (2012).
 295. Gomm, A. & O'Reilly, E. Transaminases for chiral amine synthesis. *Curr. Opin. Chem. Biol.* **43**, 106–112 (2018).
 296. Slabu, I., Galman, J. L., Lloyd, R. C. & Turner, N. J. Discovery, Engineering, and Synthetic Application of Transaminase Biocatalysts. *ACS Catal.* **7**, 8263–8284 (2017).
 297. Grishin, N. V, Phillips, M. A. & Goldsmith, E. J. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci.* **4**, 1291–1304 (1995).
 298. Steffen-Munsberg, F. *et al.* Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. *Biotechnol. Adv.* **33**, 566–604 (2015).
 299. 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 Bioinforma. Vol.* **10**, 273 (2009).
 300. Liang, J., Han, Q., Tan, Y., Ding, H. & Li, J. Current Advances on Pyridoxal 5'-Phosphate-Dependent Structure-Function Relationships of Enzymes. *Front. Mol. Biosci.* **6**, 4 (2019).
 301. Dunathan, H. C. Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **55**, 712–716 (1966).
 302. Slabu, I. *et al.* n-Butylamine as an alternative amine donor for the stereoselective biocatalytic transamination of ketones. *Catal. Today* **36**, 96–101 (2018).
 303. Toney, M. D. Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim. Biophys. Acta J.* **1814**, 1407–1418 (2011).
 304. Stirling, D. I., Zeitlin, A. L. & Matcham, G. W. Enantiomeric enrichment and stereoselective synthesis of chiral amines. (1990).
 305. Stirling, D. I., Zeitlin, A. L., Matcham, G. W. & James D. Rozzell, J. Enantiomeric enrichment and stereoselective synthesis of chiral amines. (1992).
 306. Stirling, D. I., Matcham, G. W. & Zeitlin, A. L. Enantiomeric enrichment and stereoselective synthesis of chiral amines. (1994).
 307. Kelly, S. A. *et al.* Application of ω -Transaminases in the Pharmaceutical Industry. *Chem. Rev.* **118**, 349–367 (2018).
 308. Limanto, J. *et al.* A highly efficient asymmetric synthesis of vernakalant. *Org. Lett.* **16**, 2716–2719 (2014).
 309. Nobili, A. *et al.* Engineering the active site of the amine transaminase from *Vibrio fluvialis* for the asymmetric synthesis of aryl-alkyl amines and amino alcohols. *ChemCatChem* **7**, 757–760 (2015).
 310. Genz, M. *et al.* Engineering the Amine Transaminase from *Vibrio fluvialis* towards Branched-Chain Substrates. *ChemCatChem* **8**, 3199–3202 (2016).

Chapter 1 — Introduction

311. Weiß, M. S. *et al.* Protein-engineering of an amine transaminase for the stereoselective synthesis of a pharmaceutically relevant bicyclic amine. *Org. Biomol. Chem.* **14**, 10249–10254 (2016).
312. Pavlidis, I. V. *et al.* Identification of (S)-selective transaminases for the asymmetric synthesis of bulky chiral amines. *Nat. Chem.* **8**, 1076–1082 (2016).
313. Weiß, M. S. *et al.* Amine Transaminase Engineering for Spatially Bulky Substrate Acceptance. *ChemBioChem* **18**, 1022–1026 (2017).
314. Voss, M. *et al.* In Silico Based Engineering Approach to Improve Transaminases for the Conversion of Bulky Substrates. *ACS Catal.* **8**, 11524–11533 (2018).
315. Schätzle, S., Höhne, M., Redestad, E., Robins, K. & Bornscheuer, U. T. Rapid and sensitive kinetic assay for characterization of ω -transaminases. *Anal. Chem.* **81**, 8244–8248 (2009).
316. Weiß, M. S., Pavlidis, I. V., Vickers, C., Hohne, M. & Bornscheuer, U. T. Glycine oxidase based high-throughput solid-phase assay for substrate profiling and directed evolution of (R)- and (S)-selective amine transaminases. *Anal. Chem.* **86**, 11847–11853 (2014).
317. Willies, S. C., Galman, J. L., Slabu, I. & Turner, N. J. A stereospecific solid-phase screening assay for colonies expressing both (R)- and (S)-selective ω -aminotransferases. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**, 20150084 (2016).
318. 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.* **53**, 10714–10717 (2014).
319. Baud, D., Ladkau, N., Moody, T. S., Ward, J. M. & Hailes, H. C. A rapid, sensitive colorimetric assay for the high-throughput screening of transaminases in liquid or solid-phase. *Chem. Commun.* **51**, 17225–17228 (2015).
320. Cairns, R., Gomm, A., Peel, C., Sharkey, M. & O'Reilly, E. A Comprehensive Quantitative Assay for Amine Transaminases. *ChemCatChem* **11**, 4738–4743 (2019).
321. Cairns, R. *et al.* Conversion of Aldoses to Valuable ω -Amino Alcohols Using Amine Transaminase Biocatalysts. *ACS Catal.* **9**, 1220–1223 (2019).
322. Gomm, A. *et al.* Application of “Smart” Amine Donors for Rapid Screening and Scale-Up of Transaminase-Mediated Biotransformations. *Eur. J. Org. Chem.* **2018**, 5282–5284 (2018).
323. Leipold, L. *et al.* The identification and use of robust transaminases from a domestic drain metagenome. *Green Chem.* **21**, 75–86 (2019).
324. Planchestainer, M., Hegarty, E., Heckmann, C. M., Gourlay, L. J. & Paradisi, F. Widely applicable background depletion step enables transaminase evolution through solid-phase screening. *Chem. Sci.* **10**, 5952–5958 (2019).
325. Hülsewede, D. *et al.* Development of an in situ-Product Crystallization (ISPC)-Concept to Shift the Reaction Equilibria of Selected Amine Transaminase-Catalyzed Reactions. *Eur. J. Org. Chem.* **2018**, 2130–2133 (2018).
326. Hülsewede, D., Dohm, J. N. & von Langermann, J. Donor Amine Salt-Based Continuous in situ-Product Crystallization in Amine Transaminase-Catalyzed Reactions. *Adv. Synth. Catal.* **361**, 2727–2733 (2019).
327. Höhne, M., Kühl, S., Robins, K. & Bornscheuer, U. T. Efficient asymmetric synthesis of Chiral amines by combining transaminase and pyruvate decarboxylase. *ChemBioChem* **9**, 363–365 (2008).
328. Shin, J. S. & Kim, B. G. Asymmetric synthesis of chiral amines with omega-transaminase. *Biotechnol. Bioeng.* **65**, 206–11 (1999).
329. Koszelewski, D. *et al.* Formal Asymmetric Biocatalytic Reductive Amination. *Angew. Chemie Int. Ed.* **47**, 9337–9340 (2008).
330. Gomm, A., Lewis, W., Green, A. P. & O'Reilly, E. A New Generation of Smart Amine Donors for Transaminase-Mediated Biotransformations. *Chem. Eur. J* **22**, 12692–12695 (2016).

331. Ryan, J. *et al.* Transaminase Triggered Aza-Michael Approach for the Enantioselective Synthesis of Piperidine Scaffolds. *J. Am. Chem. Soc.* **138**, 15798–15800 (2016).
332. Taday, F. *et al.* Asymmetric Construction of Alkaloids by Employing a Key ω -Transaminase Cascade. *Chem. Eur. J* **26**, 3729–3732 (2020).
333. O'Reilly, E. *et al.* A regio- and stereoselective ω -transaminase/monoamine oxidase cascade for the synthesis of chiral 2,5-disubstituted pyrrolidines. *Angew. Chemie Int. Ed.* **53**, 2447–2450 (2014).
334. Citoler, J., Derrington, S. R., Galman, J. L., Bevinakatti, H. & Turner, N. J. A biocatalytic cascade for the conversion of fatty acids to fatty amines. *Green Chem.* **21**, 4932–4935 (2019).
335. Sattler, J. H. *et al.* Redox self-sufficient biocatalyst network for the amination of primary alcohols. *Angew. Chemie Int. Ed.* **51**, 9156–9159 (2012).

2 Materials and Methods

2.1 Materials

All cells used were propagated from lab stocks. Unstained Protein Standard, Broad Range from New England Biolabs® was the weight marker used for SDS–PAGE. Plasmids were purified using the NucleoSpin® Plasmid kit from Macherey–Nagel. Centrifugation: for small volumes the accuSpin™ Micro 17R centrifuge from Fisher Scientific™ and for large volumes the Heraeus™ Multifuge™ X3R from Thermo Scientific were used. Activity assays as well as DNA and protein quantification were carried out using the Epoch 2 Microplate Spectrophotometer from BioTek Instruments. Cells were lysed using a Fisherbrand™ Model 120 Sonic Dismembrator. Affinity purification was carried out on the Äkta™ Start from GE Healthcare Life Sciences using a Qiagen Ni-NTA Superflow column (5 mL). Protein dialysis was carried out using Dialysis tubing cellulose membrane (MWCO 14,000 Da) purchased from Sigma–Aldrich®. Flow biocatalysis was carried out using an R-series modular flow chemistry system from Vapourtec. Relizyme resins for protein immobilization were kindly donated by Residion. Reagents were purchased from Arcos organics, Sigma Aldrich, Thermo Fisher, Alfa Aesar, Apollo Scientific, or Fluorochem and used without further purification. Restriction enzymes, polymerases and ligases were purchased from New England Biolabs. Enzymes were prepared in-house (see electronic supporting information) or provided by Johnson Matthey (RTA-40; RTA-57; *RTA-43; RTA-25; STA-1; STA-2 STA-13; STA-14, where the * indicates a pre-commercial enzyme and the numbers refer to the product code). MALDI-TOF MS was carried out using ground steel target plates on a Bruker ultraFlex III MALDI-TOF mass spectrometer. NMR spectra were obtained using a Bruker 400 MHz NMR spectrometer (Bruker AV3400 or AV3400HD) and referenced relative to the residual protonated solvent peak. ESI-MS data were obtained on a Bruker MicroTOF spectrometer.

2.2 Media

2.2.1 LB medium and LB agar

Fisher BioReagents™ LB broth, Miller (composition: 10 g tryptone, 5 g yeast extract, 10 g sodium chloride) or LB agar, Miller (composition: 10 g tryptone, 5 g yeast extract, 10 g sodium chloride, 15 g agar) were added to distilled water (25 g and 40 g per 1 L, respectively) and autoclaved.

2.2.2 *SOC medium*

Peptone (4 g), yeast extract (1 g), and sodium chloride (0.1 g) were dissolved in distilled water (198 mL). Potassium chloride (250 mM, 2 mL) was added, and the pH adjusted to 7.0. After autoclaving, magnesium chloride (1 M, 100 μ L), magnesium sulfate (1 M, 100 μ L), and glucose (20% w/v, 180 μ L) were added to 10 mL aliquots before use.

2.2.3 *ZYP-5052 auto-induction medium*¹

N-Z-amine (3 g), yeast extract (1.5 g), potassium dihydrogen phosphate (50 mM), potassium monohydrogen phosphate (50 mM), and ammonium sulfate (25 mM) were dissolved in 300 mL of distilled water and sterilized by autoclaving. After autoclaving, the following filter-sterilized solutions were added: magnesium sulfate (1M, 0.6 mL), trace element solution (0.6 mL), and 5052 solution (6 mL).

Trace element solution: Iron(II) chloride (50 mM), calcium(II) chloride (20 mM), manganese(II) chloride (10mM), zinc(II) sulfate (10 mM), cobalt(II) chloride (2 mM), copper(II) chloride (2mM), nickel(II) chloride (2mM), hydrogen chloride (60 mM), sodium molybdate (2 mM), sodium selenate (2 mM), boric acid (2 mM).

5052 solution: Glycerol (25 g), glucose (2.5 g), α -lactose monohydrate (10 g) in 100 mL of distilled water.

2.2.4 *TB-lac medium*

Tryptone (12 g), yeast extract (24 g), glycerol (4 mL), potassium dihydrogen phosphate (3.2 g), and potassium monohydrogen phosphate (9.4 g) were dissolved in distilled water (1 L) and sterilized by autoclaving. Sterile-filtered (0.2 μ m) lactose solution (20 \times) was added (final concentration of 5 g/L).

2.3 *Preparation of chemically competent cells*

A single colony of *E. coli* was inoculated into LB medium (2 \times 50 mL) and grown at 37 $^{\circ}$ C, 170 rpm to an OD₆₀₀ of ca. 0.5. The cultures were then placed on ice for 1 h and subsequently harvested by centrifugation (3,000 g, 15 min, 4 $^{\circ}$ C). Each pellet was resuspended in 10 mL of an ice-cold solution of magnesium chloride (100 mM), combined, and harvested by centrifugation (3,000 g, 15 min, 4 $^{\circ}$ C). The pellet was resuspended in 20 mL of an ice-cold solution of calcium chloride (100 mM) and incubated on ice for ca. 30 min. After harvesting by centrifugation (3,000 g, 15 min, 4 $^{\circ}$ C), pellets were resuspended in 10 mL of an ice-cold solution of calcium chloride (85 mM) and glycerol (15 % v/v). Cells were harvested (3,000 g,

10 min, 4 °C), resuspended in 500 µL of the previous solution, aliquoted (40 µL) into 1.5 mL microfuge tubes, and frozen at -80 °C.

2.4 Preparation of electrocompetent cells

A single colony of *E. Coli* XL10-Gold was inoculated into LB medium (2 × 50 mL) and incubated at 37 °C, 170 rpm until an OD₆₀₀ of ca. 0.5 was reached. Cells were chilled on ice for ca. 30 min and then harvested by centrifugation (2,000 g, 4 °C, 15 min). Each pellet was then resuspended in 50 mL of ice-cold autoclaved water, followed by centrifugation (4,700 g, 4 °C, 20 min). This was repeated using 25 mL of a 10% (v/v) glycerol solution per pellet, followed by 25 mL the same solution on the combined pellets. The final pellet was then resuspended in 0.5 mL of the solution and aliquoted (50 µL) into 1.5 mL microfuge tubes and frozen at -80 °C.

2.5 Transformation

2.5.1 Routine heat shock

Chemically competent cells (20–80 µL) with added plasmid–DNA (0.1–2 µL, 1–100 ng) were incubated on ice for 20 min, then heat–shocked at 42 °C for 90 seconds and immediately placed back on ice for 5 min. LB medium (500 µL) was added to the cells followed by outgrowth at 37 °C, 180 rpm, 45 min. Appropriate amounts were then plated on selective LB agar plates and incubated overnight (37 °C).

2.5.2 Higher efficiency heat shock

Chemically competent cells (60–80 µL) were incubated on ice with β–ME (25 mM) for 10 min. Plasmid–DNA was added, and samples were incubated on ice for 30 min, heat–shocked and placed on ice as above. 1 mL of warm (37 °C) SOC medium was added. Cells were outgrown at 37 °C, 180 rpm, 60 min and appropriate amounts were plated on selective LB agar plates and incubated overnight (37 °C).

2.5.3 Electroporation

Electrocompetent cells (50 µL) with added plasmid–DNA (1–5 µL) were placed into ice-cold electroporation cuvettes. Cells were electroporated at 1.8 kV for ca. 5 ms, and 0.5 mL of warm SOC medium (37 °C) were added immediately. Cells were outgrown at 37 °C for 60 mins and appropriate amounts were plated on selective LB agar plates and incubated overnight (37 °C).

2.6 *Molecular biology techniques*

2.6.1 *DNA agarose gel electrophoresis*

Agarose gels were prepared by dissolving 0.75 or 1%v (w/v) agarose in hot TAE buffer (Tris (40 mM), acetic acid (20 mM), EDTA (1 mM)). After cooling to approx. 60 °C, SYBR Safe (1 µL per 10 mL) was added, and the gel allowed to set. Gels were run at 75 V, 150 mA for approx. 50 min, and visualized using blue LEDs.

2.6.2 *Discovery of novel RTA sequences*

Protein BLAST searches were performed with the sequences of several reported RTAs against the NCBI non-redundant protein sequence database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) as well as the now defunct fungal genomics database (<https://genome.fungalgenomics.ca/>) of the Genozymes project. Candidate sequences were searched for sequences from extremophilic or extremotolerant organisms which were then inspected for the consensus sequence described by Höhne, *et al.*² Multiple sequence alignments against several reported RTAs were performed using MUSCLE³ to assess sequence identities.

2.6.3 *Cloning*

The codon-optimized synthetic gene of *TsRTA* was purchased from GeneArt in the cloning vector pMA, flanked by BamHI and HindIII restriction sites. Restriction digests of pMAT-*TsRTA* and pCH93b-*CvSTA* were set up in separate vials as follows: pMAT-*TsRTA* (700 ng) or pCH93b-*CvSTA* (1500 ng), CutSmart (10x) buffer (3 µL), BamHI HF (2 µL), HindIII HF (2 µL) and nuclease free water to 30 µL followed by incubation at 37 °C for 30 min. The backbone pCH93b and the insert *TsRTA* were purified from an agarose gel (1%, 150 mA, 75 V, 50 min) using the GeneJet gel purification kit (excised backbone and insert were combined prior to purification). DNA was eluted in nuclease free water (43 µL, 45 °C). 10x T4 ligase buffer (5 µL) as well as T4 ligase (2 µL) were added. The reaction was incubated at 16 °C overnight. DNA was isolated by ethanol precipitation and transformed into electrocompetent *E. coli* XL10-gold and, following outgrowth, plated onto selective LB-agar plates (ampicillin (amp) 100 µg/mL), and incubated at 37 °C overnight. The presence of the insert and absence of mutations were verified by sequencing. *HoRTA* was cloned analogously into both pMP89b and pCH93b.

The synthetic gene of *AtRTA* was kindly provided by Johnson Matthey. The gene was amplified (Q5 polymerase 2x Master Mix (12.5 µL), *AtRTA*_fwd (0.5 µM), *AtRTA*_rev (0.5 µM),

template (ca 1-10 ng), final volume 25 μ L. Cycling conditions: 98 °C, 30 s; 30x (98 °C, 10 s, 67 °C 30 s, 72 °C 30 s); 72 °C 5 min, 4 °C. The mix was purified (GeneJet PCR purification kit), digested (PCR product (1.4 μ g) or pCH93b (2.4 μ g) (23 μ L), SacI HF (2 μ L), HindIII HF (2 μ L), CutSmart (10x) buffer (3 μ L); incubation at 37 °C for 90 min), gel purified as above (elution in 25 μ L diluted kit elution buffer (25% v/v in nuclease free water, 65 °C), and ligated (sample (25 μ L), 10x T4 ligase buffer (3 μ L), T4 ligase (2 μ L), incubation at 25 °C for 30 min, followed by 65 °C for 10 min). Electroporation was then carried out as above, using 5 μ L of the ligation mix.

AtRTA_fwd: 5'- ACAGATAGAGCTCCATGGCCAGCATGGACAAAG -3'

AtRTA_rev: 5'- ACAGATAAAGCTTATTACGCTCGTTATAGTCGATTTCAAACG -3'

2.6.4 Site-directed mutagenesis

Mutants were prepared using the QuikChange Lightning multi site-directed mutagenesis kit from Agilent, following the manufacturer protocol: 10x QuikChange Lightning Multi reaction buffer (2.5 μ L), vector containing the target gene (ca. 100 ng), primer(s) (0.4 μ M), dNTP mix (1 μ L), QuikSolution (0-1 μ L), QuikChange Lightning Multi enzyme blend (1 μ L), nuclease free water (to 25 μ L). Cycling conditions: 95 °C, 2 min; 30x (95 °C, 20 s, 55 °C 30 s, 65 °C 30s/kbp); 65 °C 5 min, 4 °C. DpnI (1 μ L) was added and the mix incubated at 37 °C for 1h. 4 μ L of the mix were then transformed into the provided XL10-gold cells and, following outgrowth in SOC medium (37 °C, 180 rpm 1h), plated onto selective LB-agar plates (amp 100 μ g/mL), and incubated at 37 °C overnight. The mutations were then verified by in-house Sanger sequencing.

The following primers were used:

AtRTA_G207C 5'-CAACCTATCCATTTTTAACGGATTGTGATGCCCATTTAACGG-3'

TsRTA_G205C 5'-ATCCGTTTCTGACCGATTGTGATGCCAATCTGAC-3'

HEwT_F18W 5'-CCTTGAAATCGGTCCAGGGATGCAGGTGATGGGCG-3'

HEwT_F84W 5'-GGATGCGTGGTCTTGAACCAAGGTGTTGTAGTACGGC-3'

HoRTA_Δ925 5'-CCGGTGTTCGTCATGGTGAT-AAGCTTGCGGC-3'

Mutation in red.

2.7 Substrate Docking

Substrates were prepared as the quinonoid intermediate using IQmol and minimized using the UFF force-field. Waters, small molecules, and PLP were removed from the crystal structure (or homology model) using the Chimera Dock Prep tool,⁴ and the lysine residue was

re-protonated. Substrates were docked using the Autodock/Vina⁵ plugin in PyMOL,⁶ with a grid centred around the active site as estimated from the PLP position in the crystal structure. All conjugated bonds in the substrate were set to rigid using the AutoDockTools package. Docked poses were accepted or rejected based by comparison to the position of PLP in the crystal structure.

2.8 Directed Evolution

2.8.1 CASTing library generation

The protocol for site-directed mutagenesis was followed, with the following modifications: where multiple degenerate primers were used to mutate the same position, primers were combined in a ratio corresponding to the number of amino acids encoded. The DpnI digestion was carried out for 2h. Following transformations, 10 µL and 50 µL of the mix were plated on selective LB-agar plates, the remainder was grown in LB medium (amp 100 µg/mL; 5 mL) at 37 °C, 170 rpm overnight. The library size was estimated from LB-agar plates, and the transformation was repeated until a satisfactory library size was reached (>3x theoretical number of variants).

The following primers were used:

HEwT_F18:NHC	5'-CCCATCACCTGCATCCC NHC ACCGATTTCAAGGCGTTG-3'
HEwT_F84:NHC	5'-CTGCCGTACTACAACAC NHC TTCAAGACCACGCATCC-3'
HEwT_F317:NHC	5'-GGCGAGTTCTTCCACGGT NHC ACCTACTCGGGGCATC-3'
HEwT_L55:KST	5'-ACGGCATGGCCGGG KST TGGTGCGTGAATCTCGG-3'
HEwT_L55:VTA	5'-ACGGCATGGCCGGG VTA TGGTGCGTGAATCTCGG-3'
HEwT_N82:NHC	5'-GAGCAACTGCCGTACTAC NHC ACCTTCTTCAAGACCACG-3'
HEwT_W56:KST+C57:NYC	5'-GACGGCATGGCCGGGCTT KSTNYC GTGAATCTCGGCTATGGTCG-3'
HEwT_W56:KST+C57:KGC	5'-GACGGCATGGCCGGGCTT KSTKGC GTGAATCTCGGCTATGGTC-3'
HEwT_W56:TGG+C57:NYC	5'-GCATGGCCGGGCTTTGG NYC GTGAATCTCGGCTATGGTC-3'
HEwT_W56:TGG+C57:KGC	5'-CATGGCCGGGCTTTGG KGC GTGAATCTCGGCTATG-3'
HEwT_I258:NYC	5'-GTCGCCGACGAGGTG NYC TGCGGCTTCGGGC-3'
HEwT_A227:KST	5'-GAACCGGTGCAGGG KST GGGGGGCCATCATG-3'

Mutation in red.

2.8.2 Colony-based screening

E. coli NEB21 (DE3) (HEwT) or NEB10-beta (STA X14) were transformed with a mutant library and grown on nitrocellulose membranes placed on selective LB-agar plates (carb (100 µg/mL) (HEwT) or kan (50 µg/mL) (STA-X14)) at 30 °C overnight. The membranes were

transferred to selective LB–agar plates supplemented with IPTG (1 mM) and incubated for a further 7 hours. For the background depletion, membranes were transferred to dialysis plates (agarose (2% w/v), PLP (0.1 mM) in Tris/HCl (10 mM, pH 8)) and incubated at 4 °C overnight. The dialysed membranes were moved to filter paper soaked in a solution of the depleting amine (10–50 mM in potassium phosphate buffer (50 mM, pH 8), 1% DMSO; 1 mL), and incubated at room temperature for 90 min. The membranes were then moved to filter paper soaked in phosphate buffer (50 mM, pH 8; 1 mL) and incubated at RT or 4 °C for 90 min. Membranes were moved to filter paper soaked in either *p*-nitrophenethylamine (10 mM) or *o*-xylylenediamine (10 mM), and DMHone (50 mM) (in phosphate buffer (50 mM, pH 8)). Plates were incubated at room temperature and colonies were observed for colour formation. Colonies exhibiting faster colour formation than a wild–type control were picked and grown in a 96-DWP (0.5 mL LB + antibiotic, 37 °C, 900 rpm) overnight. Glycerol stocks were prepared by diluting 150 µL of the culture with 50 µL of glycerol (80% v/v) and stored at -80 °C.

2.8.3 DWP screening

ZYP-AI medium (500 µL) with carb (100 µg/mL) (HEwT) or TB-lac with kan (50 µg/mL) (STA-X14) in a 96-DWP were inoculated with either the glycerol stock (10 µL) or the LB-culture (5 µL) from **Section 2.8.2** and incubated at 25 or 30 °C, 900 rpm for 24 h. Cells were harvested by centrifugation, 3800 g, 4 °C, 10 min and the supernatant was discarded. Plates were stored at -20 °C. To screen, plates were thawed, and pellets resuspended in 100 µL of buffer (potassium phosphate (50 mM), PLP (0.1 mM), pH 8). To prepare cfe, cells were lysed in a chilled (4 °C) QSonica cup horn sonicator (100% amplitude, 5 s on, 10 s off, for 3×3 min) and the debris removed by centrifugation (3800 g, 4 °C, 30 min). Biotransformations were set up by transferring 95 µL of the cfe to the reaction mixture, containing DMHone or acetophenone (50 mM), IPA (250 mM) or SMBA (100 mM), PLP (0.1 mM), DMSO (10% (v/v)) in potassium phosphate buffer (50 mM), pH 8 (final concentrations). The volume of the reaction was 300 or 500 µL. Reactions were incubated at 30 °C, 900 rpm for 24 h. 10 µL of the reaction was diluted with 90 µL of potassium phosphate buffer (50 mM), pH 8, Fmoc derivatized (DMHone only) (**Section 2.16.1**) and analysed by UPLC (**Section 2.16.5**).

2.9 Enzyme preparation and characterization

2.9.1 Expression of HEwT

E. coli BL21 (DE3) or BL21 STAR™ (DE3) were transformed with pMP89a–HEwT and grown on selective LB–agar plates (amp 100 µg/mL) at 37 °C overnight. ZYP–AI medium¹ (300 mL;

amp 100 µg/mL) was inoculated with a single colony of transformed *E. coli* and incubated for 20 h at 37 °C, 180 rpm (19 mm throw). Cells were harvested (4500 g, 20 min, 4 °C) and stored at -20 °C either as pellets or in resuspension buffer.

2.9.2 Expression of *TsRTA*, *AtRTA*, **RTA-43_His*, *LDH-4*, and *GDH*

The plasmid containing the synthetic gene was transformed into BL21 STAR (DE3) *E. coli* cells. 300 mL of TB-lac (amp (100 µg/mL) or kan (50 µg/mL)) were inoculated with a single colony and incubated at 37 °C, with shaking (180 rpm, 19 mm throw) for 4 h followed by 25 °C with shaking (180 rpm, 19 mm throw) for 20 h. Cells were harvested by centrifugation (4500 g, 15 min, 4 °C) and stored at -20 °C either as pellets or in resuspension buffer.

2.9.3 Enzyme purification

Enzymes were purified using the Äkta Start system: Pellets were resuspended in loading buffer (potassium phosphate (50 mM), sodium chloride (100 mM), imidazole (30 mM), PLP (0.1 mM), pH 8.0; 3:1 v:w), lysed on ice by sonication ($1/4''$ probe ($1/8''$ for <2 mL), 50% amplitude, 5 s on, 5 s off, for 12 min), and cell debris was removed by centrifugation (22,800 or 25,500 g, 4 °C, 45 or 60 min). The filtered (0.45 µm) supernatant was loaded (1 mL/min) onto a Qiagen Ni-NTA Superflow column (5 mL), washed with loading buffer (4 CV), followed by a step with 10% elution buffer (loading buffer containing imidazole (300 mM); 5 CV). The protein was then eluted with 100% elution buffer in fractions of 2.5 mL. Fractions containing protein were combined and dialysed against ice-cold dialysis buffer (potassium phosphate (50 mM), PLP (0.1 mM), pH 8.0; 800 mL); the buffer was renewed after the first 2 hours of dialysis. Enzyme concentration was estimated by the absorbance at 280 nm (non-denatured protein), using predicted extinction coefficients (HEwT: 54214.57 Da, $\epsilon = 62340 \text{ M}^{-1} \text{ cm}^{-1}$, *TsRTA*: 40493.38 Da, $53860 \text{ M}^{-1} \text{ cm}^{-1}$, *AtRTA*: 39860.55 Da, $50880 \text{ M}^{-1} \text{ cm}^{-1}$; <https://web.expasy.org/protparam/>). Protein samples were stored at 4 °C until further use.

For *TsRTA* and *AtRTA* (and mutants), pellets were optionally resuspended in loading buffer with increased PLP (1 mM) (resuspension buffer), as the concentration of TA exceeded 0.1 mM. For *TsRTA_G205C* and *AtRTA_G207C*, β -mercaptoethanol (10 mM) was optionally included in the resuspension, loading, and elution buffers.

2.9.4 Enzyme lyophilization

Pellets were resuspended in buffer (potassium phosphate (50 mM), PLP (0.1 mM (HEwT) or 1 mM (*Ts*RTA and *At*RTA)), pH 8.0; 3:1 v:w), and the cfe was obtained as in **Section 2.9.3**. The cfe was filtered (0.2 μ m), frozen, and lyophilized to obtain a beige powder, which was stored at -20 °C.

2.9.5 Enzyme immobilization

Based on the technique previously applied to the STA HEwT,^{7,8} resin (Relizyme HFA403/S, EP403/S, or epoxy-agarose) was incubated in modification buffer (2 mL/g_{resin}; sodium borate (100 mM), saturated (approx. 2 M) iminodiacetic acid, pH 8.5) at ambient temperature with gentle shaking for 2h. The resin was removed by vacuum filtration, washed with dH₂O, and incubated in metal buffer (5 mL/g_{resin}; sodium chloride (1 M), cobalt(II) chloride hexahydrate (5 mg/mL)) at ambient temperature with gentle shaking for 2h. Following filtration and washing, the resin was suspended in storage buffer (potassium phosphate (50 mM), pH 8, PLP (0.1 mM)) containing the enzyme (varying concentrations and ratios with respect to resin, to achieve the desired loading) and gently shaken at ambient temperature. The incubation time was determined from test immobilizations and the remaining protein concentration and activity in the supernatant, typically 3-5 h. The resin was filtered and washed, suspended in desorption buffer (4 mL/g_{resin}; EDTA (50 mM), sodium chloride (500 mM), potassium phosphate (20 mM), pH 7.4), filtered and washed again, and suspended in blocking buffer (4 mL/g_{resin}; saturated (approx. 3 M) glycine, PLP (0.1 mM), pH 8.5). Following incubation at ambient temperature with gentle shaking for 20 h the resin was filtered off, washed, and stored in storage buffer at 4 °C. Samples of each washing step were taken and analysed by SDS-PAGE. Resins containing immobilized enzyme were boiled in SDS-loading buffer, and the supernatant was analysed by SDS-PAGE. The activity of the resin was determined by shaking an appropriate amount of resin in standard assay buffer, taking samples periodically and measuring the increase in UV absorbance at 245 nm (samples were re-added after measurement).

2.9.6 Gel filtration

Gel filtration chromatography was carried on a Superdex 200 10/300 GL column (GE Healthcare), using a mobile phase of potassium phosphate buffer (50 mM), sodium chloride (100 mM), PLP (0.1 mM), pH 8. Injection volume: 100 μ L, flow rate 0.75 mL/min. Samples were prepared in potassium phosphate buffer (50 mM), PLP (0.1 mM), pH 8 (with or without sodium chloride (100 mM)), to a final protein concentration of ca 2 mg/mL. Samples were

either injected directly after preparation or incubated at ambient temperatures for varying amounts of time to follow the interconversion of different quaternary states. A calibration curve was generated using the Sigma Aldrich Gel Filtration Markers Kit for Protein Molecular Weights 12,000-200,000 Da (MWGF200).

2.9.7 SDS-PAGE

Gels (12% acrylamide) were cast as follows: distilled water (3.125 mL), Tris-HCl buffer (1.5 M, pH 8.8; 1.95 mL), acrylamide (40% 37.5:1; 2.25 mL), SDS (10%; 75 μ L), APS (10%, 75 μ L), and TEMED (7.5 μ L) were mixed and poured into the mould. Once set, the stacking gel (distilled water (1.4 mL), Tris-HCl buffer (1 M, pH 6.8; 0.25 mL), acrylamide (40% 37.5:1; 0.33 mL), SDS (10%; 20 μ L), APS (10%, 20 μ L), and TEMED (3 μ L)) was poured on top and the comb for the wells placed. Samples were mixed with 2x loading buffer (Bromophenol blue (10 mg), Tris-HCl (1 M, pH 6.8; 1 mL), distilled water (2.5 mL), SDS (0.4 g), glycerol (4 mL), EDTA (0.15 g), and β -ME (0.5 mL)) and boiled for ca. 10 min. Gels were run at 300 V, 30 mA for 60-90 min (Running buffer: glycine (14 g), Tris-base (3.1 g), SDS (10%, 10 mL), total volume 1 L). Gels were stained with InstantBlue™ Protein Stain.

2.9.8 Stability, activity, and kinetic assays

Activity assays were based on the method by Schätzle *et al.*⁹ as applied in Cerioli *et al.*¹⁰, using (*R*)-methylbenzylamine (RMBA) in place of SMBA for RTAs, in UV-free 96-well plates using the EPOCH 2 plate reader at 30 °C unless otherwise stated (pathlength 0.84 cm, calculated according to $\frac{A_{977}-A_{900}}{0.18}$); $\epsilon = 12.6 \text{ mM}^{-1} \text{ cm}^{-1}$).

2.9.8.1 Temperature, pH, and co-solvent stability assays

The purified enzyme in potassium phosphate buffer (50 mM, pH 8, PLP (0.1 mM)) was diluted to a final conc. of 0.30–0.38 mg/mL with either phosphate buffer (temperature stability assays), universal buffer adjusted to the desired pH (pH stability assays)¹⁰, or phosphate buffer containing 10% or 20% ((v/v) final concentration) co-solvents (co-solvent stability assays). Aliquots (30 μ L) were stored at 35-60 °C for the temperature stability assays, 4 °C for pH stability, and 25 °C for the co-solvent stability assays. For each time point, one aliquot was briefly centrifuged in a microfuge to ensure the complete sample was collected at the bottom of the tube, incubated on ice for ca. 10 min and the standard activity assay described above was performed in triplicate on each sample, using 3 μ L of the enzyme sample in each reaction. The sample was then discarded.

2.9.8.2 Temperature, pH, and co-solvent activity assays

Modified activity assays were performed as follows: for temperature activity assays, 3 μL of enzyme solution, stored on ice at the appropriate dilution (0.07-0.09 mg/mL), and assay buffer (300 μL), pre-heated to the desired temperature, were combined. The assay was performed in a 96-well plate in a pre-heated plate reader. For pH activity assays, assay buffer was prepared in universal buffer¹¹ at the desired pH. For the co-solvent activity assays, assay buffer containing the desired co-solvent was prepared. To initiate the assay, 3 μL of enzyme solution, stored on ice at the appropriate dilution (0.37 mg/mL), and assay buffer (300 μL), were combined and the assay performed as usual.

2.9.8.3 Kinetic studies

Activity assays were performed by either fixing the pyruvate concentration at 10 mM (1% (v/v) DMSO) while the RMBA concentration was varied (0.005 mM to 10 mM) or with the RMBA concentration fixed at 2.5 mM (0.25% (v/v) DMSO) and the pyruvate concentration varied (0.005 mM to 10 mM), using purified enzyme at the appropriate dilution (0.0030–0.0038 mg/mL, concentration in assay). Assays at each concentration were performed in triplicate. Kinetic parameters k_{cat} , K_m , and K_i were obtained by fitting a substrate

inhibition curve $\left(v = \frac{k_{cat}}{1 + \frac{K_m + [S]}{[S]} + \frac{[S]}{K_i}} \right)$ using GraphPad Prism.¹²

2.9.9 MALDI-TOF MS

Protein samples (20 μL) in potassium phosphate buffer (50 mM; PLP (0.1 mM), pH 8) were diluted with an equal volume 1% TFA. For preparation of reduced samples, TCEP (10 μL of a 200 mM stock) was added to the sample, followed by incubation at 70 °C for 15 min. Samples were then desalted using C4 ZipTip® pipette tips (Merck Millipore) as follows: The sample was bound to the tip under saturating conditions, the resin washed (5% methanol, 0.1% TFA in water; 20 \times 10 μL) and the protein eluted in 5 μL elution buffer (80% acetonitrile, 0.1% TFA in water). The desalted samples (2 μL) were mixed with 2 μL of a solution of sinapic acid (20 mg/mL; in elution buffer) of which 1 μL was spotted onto a ground steel target plate, dried, and coated with 1 μL of the sinapic acid solution. The spots were then analysed using a Bruker ultraFlex III MALDI-TOF mass spectrometer (laser amplitude 60%). External calibration relative to HSA (66440 kDa).

2.10 TsRTA crystallization and data collection

The following was carried out by our collaborator Dr Louise Gourley (University of Milan), and has been included for completeness:

Crystallization trials of TsRTA (10 mg/ml; 20 mM Tris-HCl pH 8.0; 0.1 mM PLP) were carried out using an Orxy 4 crystallization robot (Douglas Instruments) and flat-bottomed, Greiner CrystalQuick 96 well sitting drop plates (Greiner Bio-one). TsRTA microcrystals grown over 2-3 days at 20°C in a 500 nl drop, containing 30% protein and PACT screen (Molecular Dimensions) condition G3 (sodium iodide (0.2 M), Bis-Tris Propane (0.1 M), pH 7.5, PEG 3350 (20% (w/v))), were used to prepare a seed stock. Seeds were prepared using the Seed Bead Kit (Hampton Research), crushing the crystals in 50 µl well solution by vortexing for 2 min. 0.15 µl seed stock was used to seed a second PACT screen, preparing 0.8 µl drops at diverse protein concentrations (31.25%, 50% and 68.75%). For microseeding, the protein concentration was halved to 5 mg/ml. Crystals were harvested from PACT condition C12 (zinc chloride (0.01 M), HEPES (0.1 M), pH 7.0, PEG 6K (20% (w/v))) from a drop containing 68.75% protein. For cryoprotection, crystals were soaked in a solution comprising 18.75% (v/v) PACT condition G3 (PEG6K (20% (w/v)), HEPES (0.1M), pH 7.0, ZnCl (0.002 M) and ethylene glycol (28% (w/v))).

X-ray diffraction data were collected on a single TsRTA crystal at 2.2 Å resolution on the XDR2 beamline at the ELETTRA synchrotron facility (Trieste, Italy). Two TsRTA chains (Chains A and B) were present in the asymmetric unit, with an estimated Matthew's coefficient of 2.7 Å³/Da (54.4 % solvent content). Data reduction was carried out using Mosflm and assigned to the body-centred monoclinic space group I₁2₁ using POINTLESS and scaled with AIMLESS.^{13,14} Molecular replacement was carried out using MOLREP and chain A of the omega transaminase from *Aspergillus terreus* (AtRTA; PDB entry 4ce5; 82% sequence identity over 321 residues) as a search model.^{15,16} All programs are available under the CCP4 suite.¹⁷ The structure was manually built and refined to convergence using coot and phenix.refine and structure geometry was validated by Molprobity in the PHENIX platform (**Table 2-1**).¹⁸

For data collection and refinement parameters see (**Table 2-1**). Atomic coordinates and structure factors are available for download from the RCSB Protein Data Bank (www.rcsb.org) under accession code 6XWB.

Table 2-1: Data collection statistics and refinement parameters for TsRTA.^a

	TsRTA
Data collection	
Space group	I ₁ 2 ₁
Cell dimensions	
<i>a, b, c</i> (Å)	67.8, 98.0, 117.3
α, β, γ (°)	90.0, 91.2, 90.0
Resolution (Å)	59-2.2 (2.2-2.27)
^a <i>R</i> _{merge}	0.112 (0.526)
<i>I</i> / σ <i>I</i>	10.2 (2.7)
Completeness (%)	100 (100)
Redundancy	6.1(6.2)
^b CC ^{1/2}	100 (91.2)
Refinement	
Resolution (Å)	2.2-59.0
No. unique reflections	38887
^c <i>R</i> _{work} / ^d <i>R</i> _{free}	20.7/23.4
No. atoms	
Protein	2482(A)2496(B)
Ethylene glycol	114
PLP	30
Sodium ion	2
Water	165
<i>B</i> -factors (Å ²)	
Protein	29 (A) 31 (B)
Water	27 (A) 26 (B)
Ethylene glycol	32
PLP	23
Sodium ion	26
Water	25
RMSD:	
Bond lengths (Å)	0.003
Bond angles (°)	0.598
Ramachandran Plot (%)	
Allowed Regions	100

^aData were collected AT 2.2Å on a single TsRTA crystal. Parentheses indicate parameters related to the high-resolution cell (2.2 – 2.27 Å). ^a*R*_{merge} = $\sum |1 - \langle I \rangle| / \sum I \times 100$, where *I* is the intensity of a reflection and $\langle I \rangle$ is the average intensity. ^bCC^{1/2} is the correlation between random half-sets of data. ^c*R*_{work} = $\sum |F_o - F_c| / \sum F_o \times 100$; ^dFor cross-validation, 10 % experimental reflections were randomly selected to calculate the *R*_{free}.

2.11 Batch biotransformations

2.11.1 Substrate scope analysis *TsRTA*

Reactions were set up containing RMBA (10 mM) (IPA (50 mM) for acetophenone) and a carbonyl acceptor (10 mM), or benzaldehyde (10 mM) (pyruvate (10 mM) for MBA) and an amino donor (10 mM (20 mM for racemates)), purified *TsRTA* (wild type and mutant, 0.5 mg/mL), PLP (0.1 mM), and DMSO (5 or 10% v/v) in potassium phosphate buffer (50 mM pH 8), in a final volume of either 600 μ L or 1 mL. Reactions were incubated at 37 °C, 180 rpm in triplicate. 100 μ L samples were quenched in with acetonitrile and aq. HCl (0.2%) (900 μ L; 1:1 v/v) and analysed by reverse-phase HPLC. Enantioselectivity was determined by chiral GC-FID or chiral reverse-phase HPLC (see **Section 2.16** for details on all analytical techniques). For the intensification studies, reactions were set up analogously but with an increasing amount of RMBA and phenoxyacetone, and the ratio of enzyme to the substrate was kept constant (0.1 mg/mL for a 10 mM reaction or 0.025 mol%). Reactions were quenched analogously but with proportionally increasing dilution factors.

2.11.2 Batch Biotransformations chapter 3—*IPA method*

Reactions were set up containing ketone (10-300 mM) and IPA (5 eq.), lyophilized TA cfe (50 mg/mmol for butanone and pinacolone; 1500 mg/mmol for DMHone), PLP (1 mM), and DMSO (0 to 5% v/v) in potassium phosphate buffer (50 mM pH 8), in a final volume of 500 μ L to 1 mL. Reactions were incubated at 30 °C, with gentle shaking in duplicate. Samples were removed, diluted appropriately (total amine content <25 mM), FMOC-derivatized and analysed by LC-MS or reverse-phase HPLC. Enantioselectivity was determined by chiral GC-FID or chiral reverse-phase HPLC.

2.11.3 Batch Biotransformations chapter 3—*Alanine method*

Reactions were set up containing ketone (10-300 mM) and alanine (1.2 eq. L-Ala (STAs) or 2.4 eq. DL-Ala (RTAs)), TA lyophilized cfe (50 mg/mmol), ammonium chloride (2.4 eq.), lyophilized AlaDH cfe (25 mg/mmol), D-glucose (1.2 eq.), lyophilized GDH cfe (25 mg/mmol), PLP (1 mM), NAD⁺ (1 mM), and DMSO (0 to 5% v/v) in potassium phosphate buffer (300 mM pH 8), in a final volume of 500 μ L to 1 mL. Reactions were incubated at 30 °C, with gentle shaking in duplicate. Samples were removed, diluted appropriately (total amine content <25 mM), FMOC-derivatized and analysed by LC-MS or reverse-phase HPLC. Enantioselectivity was determined by chiral GC-FID or chiral reverse-phase HPLC.

2.11.4 Batch biotransformations chapter 4

Reactions were set up containing ketone (10-300 mM), amine donor (SMBA or isopropylamine in varying amounts as indicated), PLP (0.1-1 mM), DMSO (0-11 % (v/v)) and HEwT (either lyophilized crude or purified), and additives (were indicated), in potassium phosphate buffer (50-100 mM), pH 8. Ketones were added from a 10-fold concentrated stock either in buffer (tetrahydrofuran-3-one) or DMSO (tetrahydrothiophene-3-one), IPA was added from a pH adjusted stock in buffer, and SMBA was added from a 10-fold concentrated pH adjusted stock containing 10% (v/v) DMSO. Additives (*i*PrOH, ammonium chloride, sodium chloride) were added from pH adjusted stocks in buffer. Reactions were incubated at 30 or 37 °C for the indicated amount of time. Conversions and enantiomeric excess were determined either by extraction into ethyl acetate, acetylation, and analysis by GC-FID, based on the relative peak areas of SMBA and acetophenone; or by FMOC-derivatization and analysis by chiral RP-HPLC, determining the concentration of product using a calibration curve (as indicated).

2.12 Flow biotransformations TsRTA

The packed-bed reactor was prepared as follows: an Omnifit glass column (6.6 mm i.d. × 150 mm length) was packed with immobilized transaminase (up to 3 g_{resin} for a packed bed volume of 3.87 mL). The packed-bed reactor was equilibrated with storage buffer (potassium phosphate (50 mM), pH 8, PLP (0.1 mM)) for 5-7 CV at a flow-rate of 0.4-0.5 mL/min.

2.12.1 Operational stability tests

Two-fold concentrated stock solutions of pyruvate (or butanone) and RMBA were prepared separately in potassium phosphate buffer (50 mM, pH 8, 0.1 mM PLP, containing the appropriate amount of DMSO). Using an R-series modular flow chemistry system from Vapourtec, both solutions were mixed and passed through the packed-bed reactor at the desired residence time and temperature. The reactor was equilibrated for 2 column volumes (CV), after which 1 fraction per column volume was collected. Conversions were assessed by reverse-phase HPLC, following the depletion of RMBA (acetophenone is retained to some extent by the resin).

2.13 Scaled-up 2-aminobutane synthesis in flow

Two Omnifit glass columns (6.6 mm i.d. × 150 mm length) were employed in series, giving a combined packed-bed volume of 7.73 mL (using 6 g_{resin}). The reactor was equilibrated with storage buffer (potassium phosphate 50 mM, pH 8, 0.1 mM PLP) for 5 CV at a flowrate of 0.5 mL/min. For reactor start-up, the reaction mixture (butanone (300 mM), IPA (1.5 M)

and PLP (1 mM) in potassium phosphate buffer (50 mM, pH 8)) was passed through the reactor at a flow-rate of 0.5 mL/min at ambient temperature until it reached the outlet. At this point, the output was fed back into the vessel containing the reaction mixture and the reaction was run in circulation mode (see **Figure 2-1**) using a residence time of 25 min per pass, at 30 °C. The reaction mixture was circulated until 4 h of contact time between the reaction mixture and resin (corresponding to 9.6 passes) was reached, at which point the inlet was moved to a fresh reaction mixture. After the volume remaining in the reactor and tubing was collected, the outlet was also moved to the fresh reaction mixture which was then circulated as before. The swapping of the reaction mixture was carried out every 1-4 days (corresponding to 6-24 CV) and the reactor was left running for 165 h total (1 week). For reactor shut-down, the inlet was placed in storage buffer and the remaining reaction mixture was flushed out of the tubing at a flow-rate of 0.5 mL/min at ambient temperature. All processed reaction mixtures were combined, acidified using hydrochloric acid (pH <1) and concentrated (to approximately 60% of the initial weight) *in-vacuo*, removing the ketones. The remaining aqueous solution was alkalized with an excess of potassium hydroxide (approx. 33 g) and distilled under atmospheric pressure. Three cuts (b.p. ≤38 °C, 38–50 °C, and 46–65 °C) were obtained that contained both IPA and 2-aminobutane, with the proportion of the latter increasing in the higher b.p. cuts (as determined by ¹H-NMR, **Table 3-4**). These cuts were then redistilled over potassium hydroxide employing a fractional distillation set-up using a four-ball Snyder column, obtaining four cuts (b.p. ≤32 °C, 32–34 °C, 34–62 °C, and 62 °C). Cuts one and two were virtually pure IPA, cut three was a mixture of IPA and 2-aminobutane, and cut four was 2-aminobutane with ≤1% IPA as determined by ¹H-NMR (**Table 3-5**), giving 2.49 g of (*S*)-2-aminobutane (35% yield, ≥99% purity (¹H-NMR), >99.5% *ee*) and 1.97 g of (*R*)-2-aminobutane (28% yield, ≥99% purity (¹H-NMR), 98.7% *ee*), from the HEwT_F84W and *RTA-43 flow reactions, respectively.

(*S*)-2-aminobutane: ¹H-NMR (400 MHz, CDCl₃) δ 2.79 (1 H, h, *J* 6.3, CHN), 1.37–1.26 (2 H, m, CH₂), 1.15 (2 H, br s, NH₂), 1.02 (3 H, d, *J* 6.3, CHNCH₃), 0.87 (3 H, t, *J* 7.4, CH₂CH₃); ¹³C-NMR (101 MHz, CDCl₃) δ 48.4 (CHN), 32.9 (CH₂), 23.4 (CHNCH₃), 10.6 (CH₂CH₃); in agreement with lit.¹⁹ *m/z* [M+H]⁺ calculated: 74.0964, found: 74.0956.

(*R*)-2-aminobutane: ¹H-NMR (400 MHz, CDCl₃) δ 2.77 (1 H, h, *J* 6.3, CHN), 1.38–1.26 (2 H, m, CH₂), 1.18 (2 H, br s, NH₂), 1.03 (3 H, d, *J* 6.3, CHNCH₃), 0.88 (3 H, t, *J* 7.4, CH₂CH₃); ¹³C-NMR (101 MHz, CDCl₃) δ 48.4 (CHN), 32.8 (CH₂), 23.4 (CHNCH₃), 10.6 (CH₂CH₃); in agreement with lit.¹⁹ *m/z* [M+H]⁺ calculated: 74.0964, found: 74.0958.

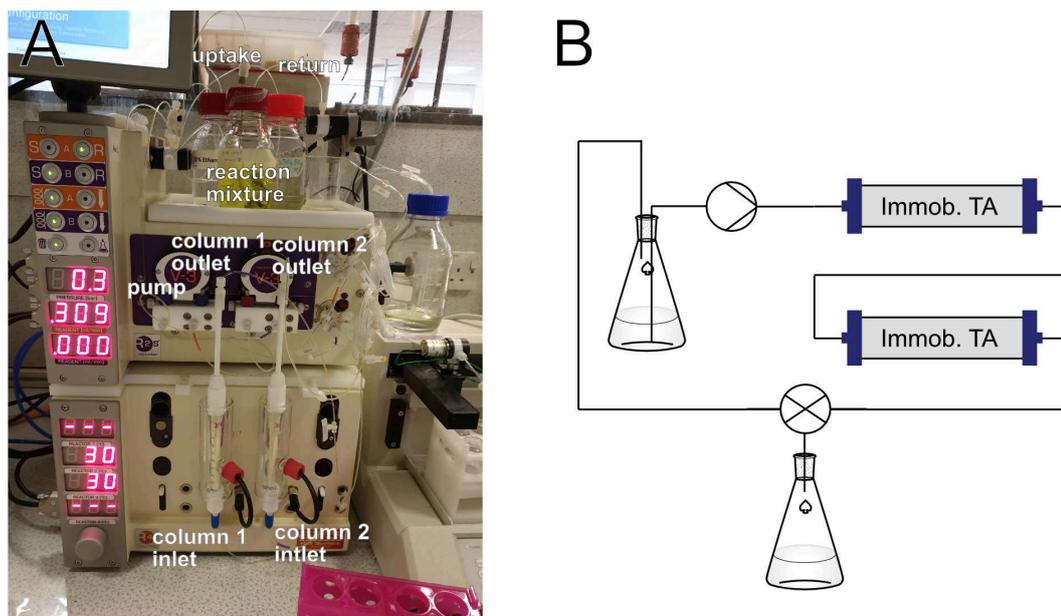
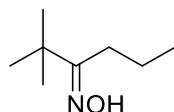


Figure 2-1: A: picture of the flow-set up, in recirculation mode. B: scheme of the flow-set-up, showing the recirculation path as well as the option for collection at the end of the reaction.

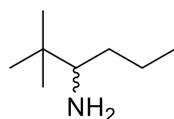
2.14 Synthesis of primary amine standards

2.14.1 2,2-dimethylhexan-3-oxime



2,2-Dimethylhexan-3-one (500 mg, 3.90 mmol) and hydroxylamine hydrochloride (540 mg, 7.77 mmol) were dissolved in 2.6 mL of an aqueous solution of sodium acetate (3 M) and 2 mL of methanol were added. The mixture was stirred for 72 h at ambient temperature during which time a white precipitate formed. The precipitate was obtained by vacuum filtration and washed with ice-cold water, giving 2,2-dimethylhexan-3-oxime as glistening white crystals in quantitative yield. $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 8.66 (1 H, br s, OH), 2.29 – 2.20 (2 H, m, C(=NOH)CH_2), 1.65–1.52 (2 H, m, CH_2CH_3), 1.12 (9 H, s, $\text{C(CH}_3)_3$), 0.98 (3 H, t, J 7.4, CH_2CH_3); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 167.5 (C(=NOH)), 37.4 ($\text{C(CH}_3)_3$), 28.1 (C(=NOH)CH_2), 27.7 ($\text{C(CH}_3)_3$), 19.9 (CH_2CH_3), 15.0 (CH_2CH_3). m/z $[\text{M}+\text{H}]^+$ calculated: 144.1383, found: 144.1397.

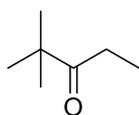
2.14.2 2,2-dimethylhexan-3-amine hydrochloride²⁰



2,2-Dimethylhexan-3-oxime (3.9 mmol) and nickel(II) chloride hexahydrate (1.9 g, 7.8 mmol) were dissolved in ice-cold methanol (20 mL). Sodium borohydride (3.0 g, 78 mmol) was added slowly over 90 min with vigorous stirring at 0 °C, and a black precipitate formed. The reaction was stirred at 0 °C for an additional 30 min and then allowed to warm to RT over 1 h under continuous stirring. The reaction mixture was quenched by addition of hydrochloric acid (6 M, ca 13 mL) and the black precipitate was removed by gravity filtration. The filtrate was basified with

sodium hydroxide (pH >10) and EDTA (ca. 1 g) was added. The filtrate was extracted with DCM (3 × 30 mL). Hydrochloric acid (6 M, 2 mL) was added to the extract, and the solvents removed *in vacuo*, giving crude 2,2-dimethylhexan-3-amine hydrochloride as a white powder (435 mg). This was recrystallized from isopropanol/hexane (1:1 v/v), giving pure 2,2-dimethylhexan-3-amine hydrochloride as a white powder (241 mg, 37%). ¹H-NMR (400 MHz, DMSO-d₆) δ 7.88 (3 H, br s, NH₃), 2.71 (1 H, dd, *J* 9.5, 2.2, CH(NH₃)), 1.62 – 1.53 (1 H, m, CH_aH_bCH₃), 1.53 – 1.47 (1 H, m, CH(NH₃)CH_aH_b), 1.45 – 1.34 (1 H, m, CH(NH₃)CH_aH_b), 1.34 – 1.20 (1 H, m, CH_aH_bCH₃), 0.94 (9 H, s, C(CH₃)₃), 0.88 (3 H, t, *J* 7.0, CH₂CH₃); in agreement with lit.²¹ ¹³C-NMR (101 MHz, DMSO-d₆) δ 59.7 (CH(NH₃)), 33.1 (C(CH₃)₃), 30.5 (CH(NH₃)CH₂), 25.8 (C(CH₃)₃), 19.3 (CH₂CH₃), 13.7 (CH₂CH₃). *m/z* [M+H]⁺ calculated: 130.1590, found: 130.1606.

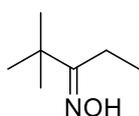
2.14.3 2,2-dimethylpentan-3-one



2,2-dimethylpentan-3-ol (1.04 g, 8.95 mmol) was added to a suspension of magnesium sulfate (4 g), pyridinium chlorochromate (3.71 g, 17.2 mmol) in dichloromethane (DCM, 20 mL) and stirred at ambient temperature for 16 h.

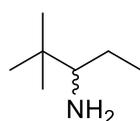
The reaction mixture was filtered through a silica/magnesium sulfate plug and solvent removed *in vacuo*, giving 2,2-dimethylpentan-3-one as a clear colourless oil (841 mg, 7.36 mmol, 82% yield). ¹H-NMR (400 MHz, CDCl₃) δ 1.03 (3 H, t (*J* 7.2), CH₂CH₃), 1.14 (9 H, s, C(CH₃)₃), 2.51 (2 H, q (*J* 7.2), CH₂); ¹³C-NMR (100 MHz, CDCl₃) δ 8.2 (CH₂CH₃), 26.5 (C(CH₃)₃), 29.6 (CH₂), 44.0 (C(CH₃)₃), 216.7 (C=O). *m/z* [M+H]⁺ calculated: 115.1116, found: 115.1117.

2.14.4 2,2-dimethylpentan-3-oxime

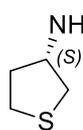


2,2-dimethylpentan-3-one (223 mg, 1.95 mmol), hydroxylamine hydrochloride (272 mg, 3.91 mmol) and sodium acetate (320 mg, 3.90 mmol) were dissolved in 2 mL of water and 1 mL of methanol was added. The mixture was stirred for

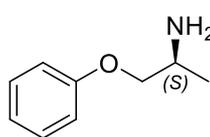
72 h at ambient temperature during which time a white precipitate formed. The precipitate was obtained by vacuum filtration and washed with ice-cold water, giving 2,2-dimethylpentan-3-oxime as white fine crystals (180 mg, 1.39 mmol, 71% yield). ¹H-NMR (400 MHz, CDCl₃) δ 1.14 (9 H, s, C(CH₃)₃), 1.15 (3 H, t (*J* 7.5), CH₂CH₃), 2.34 (2 H, q (*J* 7.5), CH₂), 8.57 (1 H, br s, OH); ¹³C-NMR (100 MHz, CDCl₃) δ 11.1 (CH₂CH₃), 18.9 (CH₂), 27.6 (C(CH₃)₃), 37.5 (C(CH₃)₃), 168.5 (C=NOH). *m/z* [M+H]⁺ calculated: 130.1226, found: 130.1232.

2.14.5 2,2-dimethylpentan-3-amine hydrochloride²⁰

2,2-Dimethylpentan-3-oxime (180 mg, 1.39 mmol) and nickel(II) chloride hexahydrate (661 mg, 2.7 mmol) were dissolved in ice-cold methanol (8 mL). Sodium borohydride (1.05 g, 27.8 mmol) was added slowly over 60 min with vigorous stirring at 0 °C, and a black precipitate formed. The reaction was stirred at 0 °C for an additional 30 min and then allowed to warm to RT over 1 h under continuous stirring. The reaction mixture was quenched by addition of hydrochloric acid (6 M, ca 4 mL) and the black precipitate was removed by gravity filtration. The filtrate was basified with sodium hydroxide (pH >10) and EDTA (ca. 1 g) was added. The filtrate was extracted with DCM (3 × 20 mL). Hydrochloric acid (6 M, 0.5 mL) was added to the extract, and the solvents removed *in vacuo*, giving crude 2,2-dimethylhexan-3-amine hydrochloride as an off-white powder (121 mg, 0.798 mmol, 57%). ¹H-NMR (400 MHz, CDCl₃) δ 1.09 (9 H, s, C(CH₃)₃), 1.21 (3 H, t (*J* 7.5), CH₂CH₃), 1.56 – 1.71 (1 H, m, CH(NH₃)CH_aH_b), 1.72 – 1.84 (1 H, m, CH(NH₃)CH_aH_b), 2.72 – 2.83 (1H, m, CH(NH₃)), 8.28 (3 H, br s, NH₃) ¹³C-NMR (101 MHz, CDCl₃) δ 12.2 (CH₂CH₃), 22.2 (CH₂), 26.5 (C(CH₃)₃), 33.8 (C(CH₃)₃), 64.1 (CNH₃). *m/z* [M+H]⁺ calculated: 116.1434, found: 116.1445.

2.14.6 (*S*)-tetrahydrothiophene-3-amine²²

[α]_D -32.2, c 1, acetone (lit. ²³ -37.77), ¹H-NMR (400 MHz, CDCl₃) δ 1.51 (3 H, br s, NH₂ + H₂O), 1.81 – 1.90 (1 H, m, SCH₂CH_aH_b), 1.94 – 2.04 (1 H, m, SCH₂CH_aH_b), 2.59 (1 H, ddd (*J* 10.7, 4.4, 0.9 Hz), SCH_aH_bCHN), 2.84 – 2.96 (2 H, m, SCH₂CH₂), 2.98 (1 H, dd (*J* 10.6, 5.1 Hz), SCH_aH_bCHNH₂), 3.71 (1 H, p (*J* 4.9 Hz), CHNH₂); ¹³C-NMR (100 MHz, CDCl₃) δ 28.4 (SCH₂CH₂), 38.7 (SCH₂CH₂), 40.0 (SCH₂CHNH₂), 55.9 (CHNH₂); in agreement with lit.^{22,23} ESI-MS (*m/z*): [M+H]⁺: calc. 104.0528, found 104.0538.

2.14.7 (*S*)-2-phenoxyisopropylamine

Phenoxyacetone (274 μL, 2 mmol), IPA (5 mmol, from a pH adjusted 2 M stock in potassium phosphate buffer (50 mM, pH8)), PLP (2 mL of a 10 mM stock in potassium phosphate buffer (50 mM, pH8)), and DMSO (2 mL) were diluted with potassium phosphate buffer (50 mM, pH8) to a final volume of 20 mL. Lyophilized HEwT cfe (25.5 mg)¹⁰ was added and the mixture incubated with gentle agitation at 25-30 °C. After 24 h, an additional 25 mg of HEwT and after another 24 h a further 50 mg were added. The reaction was incubated for another 72 h after which the reaction was basified with 3 mL NaOH (5 M) and extracted with 3 × 20 mL EtOAc. The combined organic extracts were dried with MgSO₄, filtered, and concentrated *in vacuo* to ca 2 mL.

Methanolic HCl (80 μ L, 3M, prepared from acetyl chloride and methanol) was added and the sample concentrated *in vacuo*. The resulting oil was recrystallized from EtOAc to give the HCl salt of (*S*)-2-phenoxyisopropylamine as white needles (290.3 mg, 77% yield, >99.5% *ee*). $[\alpha]_D^{25}$ 33.8, *c* 2, methanol (lit.²⁴ -28.1 for the (*R*)-enantiomer), ¹H-NMR (400 MHz, DMSO-*d*₆) δ 1.29 (3 H, t (*J* 6.7 Hz), Me), 3.58 (1 H, pd (*J* 6.8, 4.0 Hz), CHNH₃), 3.99 (1 H, dd (*J* 10.2, 7.0 Hz), CH_aH_b), 4.12 (1 H, dd (*J* 10.2, 4.0 Hz), CH_aH_b), 6.95-7.02 (3 H, m, *o,p*-Ar-H), 7.32 (2 H, dd (*J* 8.8, 7.0 Hz), *m*-Ar-H), 8.22 (3 H, br s, NH₃); ¹³C-NMR (100 MHz, DMSO-*d*₆) δ 15.0 (Me), 46.0 (CHNH₃), 68.5 (CH₂), 114.6 (*o*-C), 121.2 (*p*-C), 129.5 (*m*-C), 157.8 (Ar-C-O); in agreement with lit.^{25,26} ESI-MS (*m/z*): [M+H]: calc. 152.1070, found 152.7078.

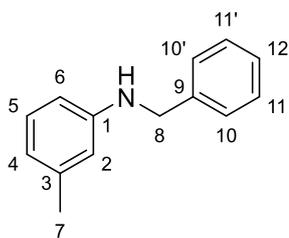
2.15 Buchwald-Hartwig Amination project

2.15.1 General method for the sequential cascade to produce *N*-aryl amines

Biotransformations were set up in a final volume of 5 mL, containing the ketone (or aldehyde) substrate (0.5 mmol), PLP (1 mM), DMSO (10% *v/v*), and KP₇-buffer (100 mM). Additionally, either isopropylamine (10 eq.) and HEwT (Iyo. cfe, 50 mg), or D-Ala (5 eq.), D-Glc (1.2 eq.), NAD⁺ (1 mM), TsRTA (Iyo. cfe, 25 mg), GDH (Iyo. cfe, 12.5 mg), and LDH-4 (Iyo. cfe, 12.5 mg) were added. Biotransformations were incubated at 30 °C with gentle shaking for 24h. Reaction mixtures were stored frozen at -20 °C or used directly in the BHA step.

Sodium *t*-butoxide (336 mg, 3.50 mmol, 7 eq. (D-Ala) or 577 mg, 6.00 mmol, 12 eq. (IPA)), *t*BuXPhos-Pd-G3 (19.9 mg, 25.0 μ mol, 5 mol% (BzH) or 39.7 mg, 50.0 μ mol, 10 mol% (all other substrates)) were placed in a flask which was flushed with argon, then the aryl halide (0.6 mmol, 1.2 eq.) was added while the flask was under argon. Next, toluene (1 mL) that had been thoroughly degassed (10 min sparging with argon under sonication) was added and the suspension sparged with argon for 5 min with vigorous stirring. Finally, the biotransformation mixture (which had been centrifuged (4800 g (4700 rpm), 5 min) to remove precipitated protein and degassed in the same way as the toluene) was added and the solution sparged for a further 2 min. Then, the solution was heated to 60 °C with vigorous stirring under argon for 24 h. After cooling to RT, samples of both the aqueous and organic phase were taken and analysed by RP-HPLC to assess conversions. Alternatively, reactions were extracted with ethyl acetate (3 \times 10 mL) and the combined organic phases analysed by GC-FID (following acetylation as described in **Section 2.16.6**). *Ees* were determined following extraction as above, either by chiral GC-FID or chiral RP-HPLC. See Section **2.16.8** for details on derivatization strategies.

2.15.2 3-benzylaminotoluene hydrochloride

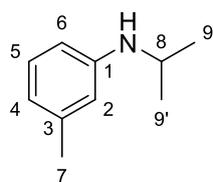


Benzylamine (98.3 μL , 96.4 mg, 0.9 mmol), 3-bromotoluene (121.2 μL , 171.03 mg, 1 mmol), [(cinnamyl)PdCl]₂ (5.7 mg, 11 μmol), 1.1 mol%), *t*BuXPhos (18.7, 44 μmol , 4.4 mol%), and TBAB (16.1 mg, 50 μmol , 5 mol%) were added to toluene (1 mL) and degassed by sparging with argon for 5 min. A degassed (15 min sparging with argon) aqueous solution of potassium hydroxide (2 M, 1 mL) was added, and the reaction heated to 50 °C under vigorous stirring (1200 rpm) for 19 h. Aqueous sodium hydroxide (5M, 200 μL) was added and the reaction mixture extracted with ethyl acetate (3 \times 3 mL). The combined organics were washed with brine (2 mL), dried over MgSO₄, and filtered through a silica plug (3 cm in Pasteur pipette). Methanolic HCl (approx. 3 M, prepared from methanol and acetyl chloride, 0.5 mL) was added and after cooling at -20 °C the resulting precipitate was obtained by vacuum filtration, washed with a mixture of ice cold hexane and ethyl acetate (6:4 (v:v), 5 mL), giving 3-benzylaminotoluene hydrochloride as white fine needles (157.7 mg, 0.675 mmol, 75% yield). ¹H-NMR (400 MHz, CDCl₃) δ 2.26 (3 H, s, C7H₃), 4.32 (2 H, s, C8H), 7.08–7.19 (4 H, m, C2H+C4H+C5H+C6H), 7.21–7.26 (3 H, m, C11H+C11'H+C12H), 7.37 (2 H, dd (*J* 6.6, 2.9 Hz), C10H+C10'H), 11.69 (2 H, br s, NH₂); ¹³C-NMR (101 MHz, CDCl₃) δ 21.2 (C7), 56.0 (C8), 120.9 (C6, 5, or 12), 124.2 (C2 or 4), 128.6 (C11+C11'), 129.3 (C6, 5, or 12), 129.4 (C6, 5, or 12), 129.5 (C9), 130.0 (C2 or 4), 131.1 (C10+C10'), 134.2 (C1), 140.0 (C3). *m/z* [M+H]⁺ calculated: 198.1277, found: 198.1279.

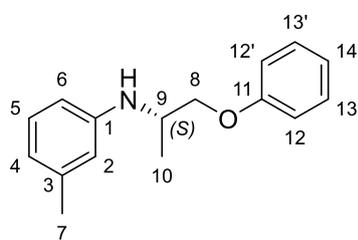
2.15.3 Synthesis of all other *N*-arylamine standards

Reactions containing amine (1.00 mmol), aryl halide (1.20 mmol), *t*BuXPhos-Pd-G3 (39.7 mg, 50.0 μmol , 5 mol%), and sodium *t*-butoxide (192.2 mg, 2.00 mmol) were set up as follows: solid reagents were placed in a flask which was flushed with argon, then liquid reagents were added while the flask was under argon. Next, toluene (1 mL) that had been thoroughly degassed (10 min sparging with argon under sonication) was added and the suspension sparged with argon for 2 min with vigorous stirring. Finally, aqueous buffer (1 mL; potassium phosphate (50 mM), DMSO 10 % (v/v), pH 8) was added and the solution sparged for a further 2 min. Then, the solution was heated to 60 °C with vigorous stirring under argon for 16–24 h. Reactions were allowed to cool and extracted with ethyl acetate (3 \times 3 mL). The organic phases were combined, and solvents removed *in vacuo*. Compounds were then purified by flash chromatography on silica gel.

2.15.3.1 3-isopropylaminotoluene

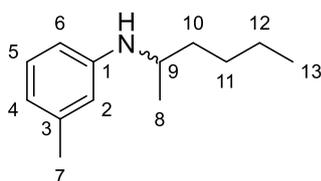


From isopropylamine (1 mL of a 1 M stock in the aqueous buffer) and 3-bromotoluene (145.6 μL , 205.2 mg, 1.20 mmol). Purified by flash chromatography on silica gel using hexane/ethyl acetate (10:0 to 9:1 v/v), giving 3-isopropylaminotoluene as a mandarin-coloured oil (117.0 mg, 0.784 mmol, 78% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 1.21 (6 H, d (J 6.3 Hz), $\text{C}_9\text{H}_3+\text{C}_9'\text{H}_3$), 2.28 (3 H, s, C_7H_3), 3.63 (1 H, hept (J 6.3 Hz), C_8H), 6.40–6.44 (2 H, m, $\text{C}_2\text{H}+\text{C}_6\text{H}$), 6.51 (1 H, d (J 7.5 Hz), C_4H), 7.06 (1 H, td (J 7.3, 1.5 Hz), C_5H); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 21.6 (C_7), 23.0 ($\text{C}_9+\text{C}_9'$), 44.3 (C_8), 110.5 (C_6), 114.1 (C_2), 118.0 (C_4), 129.1 (C_5), 139.0 (C_3), 147.4 (C_1). m/z $[\text{M}+\text{H}]^+$ calculated: 150.1277, found: 150.1277.

2.15.3.2 (*S*)-3-(2-phenoxyisopropylamino)toluene

From (*S*)-2-phenoxyisopropylamine hydrochloride (159.6 mg, 0.850 mmol) and 3-bromotoluene (123.8 μL , 174.5 mg, 1.02 mmol), using *t*BuXPhos-Pd-G3 (13.5 mg, 17.0 μmol , 2 mol%), and sodium *t*-butoxide (163.4 mg, 1.70 mmol). Purified by flash chromatography on silica gel using hexane/ethyl acetate (10:0 to 9:1 v/v), giving (*S*)-3-(2-phenoxyisopropylamino)toluene as a pale-yellow oil (125.0 mg, 0.518 mmol, 61% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 1.37 (3 H, d (J 6.3 Hz), C_{10}H_3), 2.28 (3 H, s, C_7H_3), 3.85–3.94 (2 H, m, $\text{C}_9\text{H}+\text{C}_8\text{H}_a$), 4.00–4.07 (1 H, m, C_8H_b), 6.47–6.52 (2 H, m, $\text{C}_2\text{H}+\text{C}_6\text{H}$), 6.56 (1 H, d (J 7.5 Hz), C_4H), 6.91 (2 H, dd (J 8.7, 1.2 Hz), $\text{C}_{12}\text{H}+\text{C}_{12}'\text{H}$), 6.97 (1 H, tt (J 7.4, 1.1 Hz), C_{14}H), 7.08 (1 H, td (J 7.4, 1.2 Hz), C_5H), 7.29 (2 H, dd (J 8.7, 7.3 Hz), $\text{C}_{13}\text{H}+\text{C}_{13}'\text{H}$); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 18.2 (C_{10}), 21.6 (C_7), 48.2 (C_9), 70.8 (C_8), 110.7 (C_6), 114.3 (C_2), 114.6 ($\text{C}_{12}+\text{C}_{12}'$), 118.7 (C_4), 120.9 (C_{14}), 129.2 (C_5), 129.5 ($\text{C}_{13}+\text{C}_{13}'$), 139.2 (C_3), 146.9 (C_1), 158.8 (C_{11}). m/z $[\text{M}+\text{H}]^+$ calculated: 242.1539, found: 242.1545.

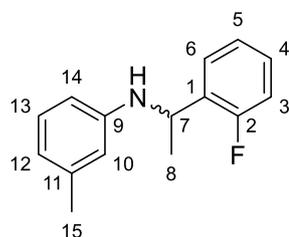
2.15.3.3 3-(hexan-2-amino)toluene



From 2-aminohexane (132.4 μL , 1.00 mmol) and 3-bromotoluene (145.6 μL , 205.2 mg, 1.20 mmol), using *t*BuXPhos-Pd-G3 (31.5 mg, 40.0 μmol , 4 mol%). Purified by flash chromatography on silica gel using hexane/ethyl acetate (100:0 to 98:2 v/v), giving 3-(hexan-2-amino)toluene as a slightly hazy yellow oil (156.2 mg, 0.816 mmol, 82% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 0.91 (3 H, t (J 6.8 Hz), C_{13}H_3), 1.17 (3 H, d (J 6.3 Hz), C_8H_3), 1.21–1.48 (5 H, m, $\text{C}_{11}\text{H}_2+\text{C}_{12}\text{H}_2+\text{C}_{10}\text{H}_a$), 1.51–1.63 (1 H, m, C_{10}H_b), 2.27 (3 H, s, C_7H), 3.44 (1 H, h (J 6.2 Hz), C_9H), 6.37–6.43 (2 H, m, $\text{C}_2\text{H}+\text{C}_6\text{H}$), 6.50 (1 H, d (J 7.4 Hz),

C4H), 7.05 (1 H, td (J 7.3, 1.3 Hz), C5H); ^{13}C -NMR (101 MHz, CDCl_3) δ 14.1 (C13), 20.8 (C8), 21.6 (C7), 22.8 (C12), 28.3 (C11), 36.9 (C10), 48.5 (C9), 110.3 (C6), 114.0 (C2), 117.8 (C4), 129.1 (C5), 139.0 (C3), 147.6 (C1). m/z $[\text{M}+\text{H}]^+$ calculated: 192.1747, found: 192.1748.

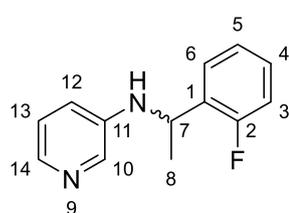
2.15.3.4 3-(*o*-fluoro- α -methylbenzylamino)toluene



From *o*-fluoro- α -methylbenzylamine (130.2 μL , 1.00 mmol) and 3-bromotoluene (145.6 μL , 205.2 mg, 1.20 mmol), using *t*BuXPhos-Pd-G3 (31.5 mg, 40.0 μmol , 4 mol%). Purified by flash chromatography on silica gel using hexane/ethyl acetate (100:0 to 98:2 v/v), giving 3-(*o*-fluoro- α -methylbenzylamino)toluene as a

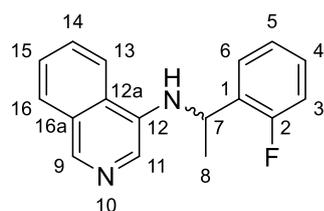
pale-yellow oil (192.4 mg, 0.839 mmol, 84% yield). ^1H -NMR (400 MHz, CDCl_3) δ 1.53 (3 H, d (J 6.7 Hz), C8H_3), 2.22 (3 H, s, C15H_3), 4.08 (1 H, br s, NH), 4.81 (1 H, q (J 6.7 Hz), C7H), 6.32 (1 H, dd (J 8.1, 2.4 Hz), C14H), 6.38 (1 H, t (J 2.0 Hz), C10H), 6.49 (1 H, d (J 7.4 Hz), C12H), 6.99 (1 H, t (J 7.7 Hz), C13H), 7.01–7.09 (2 H, m, $\text{C3H}+\text{C5H}$), 7.19 (1 H, tdd (J 7.5, 5.3, 1.8 Hz), C4H), 7.38 (1 H, td (J 7.8, 1.7 Hz), C6H); ^{13}C -NMR (101 MHz, CDCl_3) δ 21.6 (C15), 23.3 (C8), 47.5 (d ($J_{\text{C-F}}$ 2.9 Hz), C7), 110.3 (C14), 114.2 (C10), 115.4 (d ($J_{\text{C-F}}$ 21.7 Hz), C3), 118.6 (C12), 124.4 (d ($J_{\text{C-F}}$ 3.4 Hz), C5), 127.2 (d ($J_{\text{C-F}}$ 4.5 Hz), C6), 128.3 (d ($J_{\text{C-F}}$ 8.1 Hz), C4), 129.0 (C13), 131.7 (d ($J_{\text{C-F}}$ 13.2 Hz), C1), 138.9 (C11), 146.8 (C9), 160.5 (d ($J_{\text{C-F}}$ 244.4 Hz), C2); ^{19}F -NMR (376 MHz, CDCl_3) δ -120.4 (ddd ($J_{\text{F-H}}$ 10.7, 7.6, 5.1 Hz), C2F). m/z $[\text{M}+\text{H}]^+$ calculated: 230.1340, found: 230.1343.

2.15.3.5 3-(*o*-fluoro- α -methylbenzylamino)pyridine

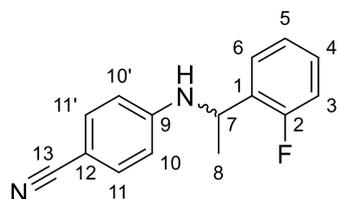


From *o*-fluoro- α -methylbenzylamine (130.2 μL , 1.00 mmol) and 3-bromopyridine (115.6 μL , 189.6 mg, 1.20 mmol). Purified by flash chromatography on silica gel using hexane/ethyl acetate (5:5 to 2:8 v/v), giving 3-(*o*-fluoro- α -methylbenzylamino)pyridine as an

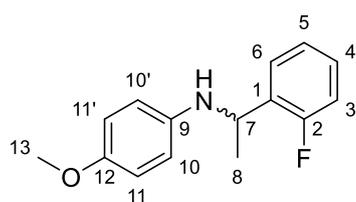
off-white powder (177.8 mg, 0.822 mmol, 82% yield). ^1H -NMR (400 MHz, CDCl_3) δ 1.56 (3 H, d (J 6.7 Hz), C8H_3), 4.24 (1 H, br s, NH), 4.75–4.84 (1 H, m, C7H), 6.73 (1 H, ddd (J 8.3, 2.9, 1.3 Hz), C12H), 6.98 (1 H, dd (J 8.3, 4.7 Hz), C13H), 7.01–7.09 (2 H, m, $\text{C3H}+\text{C5H}$), 7.16–7.24 (1 H, m, C4H), 7.32 (1 H, td (J 7.8, 1.9 Hz), C6H), 7.90 (1 H, dd (J 4.8, 1.3 Hz), C14H), 8.01 (1 H, d (J 3.0 Hz), C10H); ^{13}C -NMR (101 MHz, CDCl_3) δ 23.2 (C8), 47.4 (d ($J_{\text{C-F}}$ 2.9 Hz), C7), 115.7 (d ($J_{\text{C-F}}$ 21.8 Hz), C3), 119.0 (C12), 123.7 (C13), 124.5 (d ($J_{\text{C-F}}$ 3.6 Hz), C5), 127.0 (d ($J_{\text{C-F}}$ 4.4 Hz), C6), 128.7 (d ($J_{\text{C-F}}$ 8.1 Hz), C4), 130.7 (d ($J_{\text{C-F}}$ 13.1 Hz), C1), 136.4 (C10), 138.7 (C14), 142.9 (C11), 160.5 (d ($J_{\text{C-F}}$ 245.1 Hz), C2); ^{19}F -NMR (376 MHz, CDCl_3) δ -120.3 (ddd ($J_{\text{F-H}}$ 10.8, 7.6, 5.2 Hz), C2F). m/z $[\text{M}+\text{H}]^+$ calculated: 217.1136, found: 217.1143.

2.15.3.6 4-(*o*-fluoro- α -methylbenzylamino)isoquinoline

From *o*-fluoro- α -methylbenzylamine (130.2 μ L, 1.00 mmol) and 4-bromoisoquinoline (249.7 mg, 1.20 mmol). Purified by flash chromatography on silica gel using hexane/ethyl acetate (9:1 to 7:3 v/v), giving two fractions of 4-(*o*-fluoro- α -methylbenzylamino)isoquinoline, an impure off-white powder (97.7 mg) and a coral powder (88.2 mg, 0.331 mmol, 33% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 1.70 (3 H, d (J 6.7 Hz), C8H_3), 4.62 (1 H, br d (J 6.0 Hz), NH), 5.04 (1 H, p (J 6.4 Hz), C7H), 6.99–7.09 (2 H, m, $\text{C3H}+\text{C5H}$), 7.20 (1 H, dddd (J 8.8, 7.2, 5.3, 1.8 Hz), C4H), 7.36 (1 H, td (J 7.7, 1.8 Hz), C6H), 7.58 (1 H, ddd (J 8.0, 6.9, 1.1 Hz), C15H), 7.66 (1 H, s, C11H), 7.69 (1 H, ddd (J 8.4, 6.8, 1.4 Hz), C14H), 7.89 (1 H, d (J 8.1 Hz), C16H), 7.92 (1 H, d (J 8.6 Hz), C13H), 8.64 (1 H, s, C9H); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 23.4 (C8), 47.8 (d ($J_{\text{C-F}}$ 3.2 Hz), C7), 115.8 (d ($J_{\text{C-F}}$ 21.9 Hz), C3), 119.0 (C13), 124.5 (d ($J_{\text{C-F}}$ 3.3 Hz), C5), 124.5 (C11), 125.8 (C12a), 126.9 (d ($J_{\text{C-F}}$ 5.1 Hz), C6), 126.9 (C15), 128.2 (C16), 128.4 (C16a), 128.7 (d ($J_{\text{C-F}}$ 8.5 Hz), C4), 129.1 (C14), 130.6 (d ($J_{\text{C-F}}$ 13.0 Hz), C1), 136.1 (C12), 142.1 (C9), 160.6 (d ($J_{\text{C-F}}$ 245.6 Hz), C2); $^{19}\text{F-NMR}$ (376 MHz, CDCl_3) δ -120.0 (ddd ($J_{\text{F-H}}$ 10.8, 7.7, 5.3 Hz), C2F). m/z $[\text{M}+\text{H}]^+$ calculated: 267.1292, found: 267.1305.

2.15.3.7 4-(*o*-fluoro- α -methylbenzylamino)benzotrile

From *o*-fluoro- α -methylbenzylamine (130.2 μ L, 1.00 mmol) and 4-bromobenzotrile (218.4 mg, 1.20 mmol). Purified by flash chromatography on silica gel using hexane/ethyl acetate (100:0 to 85:15 to 80:20 v/v), giving an impure bright yellow powder (153.6 mg), which was recrystallized from ethyl acetate/hexane, giving 4-(*o*-fluoro- α -methylbenzylamino)benzotrile as a bright orange-yellow powder (127.2 mg, 0.529 mmol, 53% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 1.61 (3 H, d (J 6.7 Hz), C8H_3), 4.69 (1 H, br s, NH), 4.88 (1 H, q (J 6.7 Hz), C7H), 6.53 (2 H, d (J 8.4), $\text{C10H}+\text{C10'H}$), 7.05–7.16 (2 H, m, $\text{C3H}+\text{C5H}$), 7.26 (1 H, ddd (J 7.4, 5.3, 1.9 Hz), C4H), 7.32 (1 H, td (J 8.1, 2.0 Hz), C6H), 7.38 (2 H, d (J 8.6 Hz), $\text{C11H}+\text{C11'H}$); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 23.0 (C8), 47.3 (d ($J_{\text{C-F}}$ 3.0 Hz), C7), 99.3 (C12), 112.9 ($\text{C10}+\text{C10'}$), 115.8 (d ($J_{\text{C-F}}$ 21.7 Hz), C3), 120.3 (C13), 124.6 (d ($J_{\text{C-F}}$ 3.5 Hz), C5), 126.9 (d ($J_{\text{C-F}}$ 4.4 Hz), C6), 129.0 (d ($J_{\text{C-F}}$ 8.3 Hz), C4), 130.2 (d ($J_{\text{C-F}}$ 13.0 Hz), C1), 133.7 ($\text{C11}+\text{C11'}$), 149.9 (C9), 160.4 (d ($J_{\text{C-F}}$ 245.4 Hz), C2); $^{19}\text{F-NMR}$ (376 MHz, CDCl_3) δ -120.1 (ddd ($J_{\text{F-H}}$ 10.8, 7.5, 5.2 Hz), C2F). m/z $[\text{M}+\text{H}]^+$ calculated: 241.1136, found: 241.1136.

2.15.3.8 4-(*o*-fluoro- α -methylbenzylamino)anisole

From *o*-fluoro- α -methylbenzylamine (130.2 μ L, 1.00 mmol) and 4-bromoanisole (150.2 μ L, 224.4 mg, 1.20 mmol).

Purified by flash chromatography on silica gel using hexane/ethyl acetate (100:0 to 95:5 to 90:10 v/v), giving

4-(*o*-fluoro- α -methylbenzylamino)anisole as a bright green-yellow oil (188.1 mg, 0.767 mmol, 77% yield). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ 1.53 (3 H, d (J 6.7 Hz), C_8H_3), 3.70 (3 H, s, C_{13}H_3), 4.75 (1 H, q (J 6.7 Hz), C_7H), 6.46–6.53 (2 H, m, $\text{C}_{10}\text{H}+\text{C}_{10}'\text{H}$), 6.68–6.73 (2 H, m, $\text{C}_{11}\text{H}+\text{C}_{11}'\text{H}$), 7.00–7.09 (2 H, m, $\text{C}_3\text{H}+\text{C}_5\text{H}$), 7.19 (1 H, tdd (J 7.3, 5.2, 1.8 Hz), C_4H), 7.37 (1 H, td (J 7.6, 1.8 Hz), C_6H); $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ 23.4 (C_8), 48.3 (d ($J_{\text{C-F}}$ 2.7 Hz), C_7), 55.7 (C_{13}), 114.6 ($\text{C}_{10}+\text{C}_{10}'$), 114.8 ($\text{C}_{11}+\text{C}_{11}'$), 115.4 (d ($J_{\text{C-F}}$ 21.9 Hz), C_3), 124.4 (d ($J_{\text{C-F}}$ 3.5 Hz), C_5), 127.3 (d ($J_{\text{C-F}}$ 5.1 Hz), C_6), 128.2 (d ($J_{\text{C-F}}$ 8.5 Hz), C_4), 131.9 (d ($J_{\text{C-F}}$ 13.4 Hz), C_1), 141.0 (C_9), 152.1 (C_{12}), 160.5 (d ($J_{\text{C-F}}$ 244.5 Hz), C_2); $^{19}\text{F-NMR}$ (376 MHz, CDCl_3) δ -120.5 (ddd ($J_{\text{F-H}}$ 10.6, 7.7, 5.4 Hz), C_2F). m/z $[\text{M}+\text{H}]^+$ calculated: 246.1289, found: 246.1293.

2.16 Analytical techniques

2.16.1 Fmoc derivatization

Borate buffer (100 mM, pH 9; 200 μ L), the sample to be analysed (diluted to a total amine concentration not exceeding 25 mM while also not exceeding more than 10 mM of DMHamine; 100 μ L), and Fmoc-Cl (15 mM, in acetonitrile; 400 μ L) were combined and mixed by vortexing. 200 μ L of that mixture were added to 800 μ L of hydrochloric acid (0.2%) and acetonitrile (1:1 v/v) and analysed by reverse-phase HPLC.

2.16.2 DNPH derivatization

Sodium citrate buffer (400 mM, pH 3; 200 μ L), the sample to be analysed (diluted to a total ketone concentration not exceeding 10 mM; 100 μ L), and dinitrophenylhydrazine (DNPH; 3 mg/mL, in acetonitrile; 300 μ L) were combined, mixed by vortexing and incubated at 37 $^\circ\text{C}$ for 20 min. 200 μ L of that mixture were added to 800 μ L of hydrochloric acid (0.2%) and acetonitrile (1:1 v/v) and analysed by reverse-phase HPLC.

2.16.3 Reverse-phase HPLC analysis of conversions

Samples were analysed using a ThermoFisher Ultimate 3000 Reverse-phase HPLC (diode array detector) on a Waters XBridge C18 column (3.5 μm , 2.1 x 150 mm) at 45 $^\circ\text{C}$ with a flow rate of 0.8 mL/min. Injection volume 2-5 μ L. The following methods were used (A: 0.1% TFA in water, B: 0.1% TFA in acetonitrile):

Gradient 1: 0 min 95% A 5% B; 1 min 95% A 5% B; 5 min 5% A 95% B; 5.10 min 0% A 100% B; 6.60 min 0% A 100% B; 7 min 95% A 5% B; 10 min 95% A 5% B. UV chromatograms were obtained at 210, 240, 250, 265, and 290. Retention times in min: acetophenone (3.85), MBA (2.11), benzaldehyde (3.54), benzylamine (1.06), phenoxyacetone (3.86), 2-phenoxyisopropylamine (3.05), *o*-F-MBA (2.067), *o*-F-acetophenone (4.03), 2-Br-phenoxyacetone (4.47), 2-(2-Br-phenoxy)isopropylamine (3.52 (putative)), FMOC-2-aminohexane (5.71), 3-benzylaminotoluene (4.00), 3-isopropylaminotoluene (3.29), 3-(*o*-F-methylbenzylamino)toluene (4.48), 3-(hexan-2-amino)toluene (4.02), 3-(2-phenoxyisopropylamino)toluene (4.18), 3-(*o*-F-methylbenzylamino)pyridine (3.63), 4-(*o*-F-methylbenzylamino)isoquinoline (3.92), 4-(*o*-F-methylbenzylamino)benzotrile (4.97), 4-(*o*-F-methylbenzylamino)anisole (3.82), 3-bromotoluene (5.23), 5-bromoindole (4.77), 4-bromoanisole (4.90), 3-bromobenzoic acid (4.19), 3-bromopyridine (1.08), 4-bromoisoquinoline (3.53), 4-bromobenzotrile (4.46).

Gradient 2: 0 min 80% A 20% B; 1 min 80% A 20% B; 4 min 5% A 95% B; 4.10 min 0% A 100% B; 5.60 min 0% A 100% B; 6 min 80% A 20% B; 9 min 80% A 20% B. UV chromatograms were obtained at 265 and 370 nm. Retention times in min: FMOC-butan-2-amine (4.55), FMOC-DMBamine (4.83), FMOC-DMPamine (4.94), FMOC-DMHamine (5.09), FMOC-tetrahydrofuran-3-amine (3.97), FMOC-tetrahydrothiophene-3-amine (4.38), FMOC-hexan-2-amine (4.87), DPNH-butanone (4.41), DPNH-acetone (4.12).

Gradient 3: 0 min 100% A 0% B; 1 min 100% A 0% B; 5 min 5% A 95% B; 5.10 min 0% A 100% B; 6.60 min 0% A 100% B; 7 min 100% A 0% B; 10 min 100% A 2% B. UV chromatograms were obtained at 240 and 290 nm. Retention times in min: 2-acetylthiazole (3.32), 1-(thiazol-2-yl)ethylamine (1.20).

Conversions were calculated from a calibration curve of authentic standards (either commercially available or synthesised and characterized as described in **Sections 2.14** and **2.15**), following the production of product(s). Alternatively, conversions were calculated by comparing the areas of starting material and product, corrected by their response factor.

2.16.4 Reverse-phase LC-MS analysis of conversions

Samples were analysed using a Agilent 1260/Agilent 6120 Quadrupole LC-MS (diode array detector) on a Waters XBridge C18 column (3.5 μ m, 2.1 x 30 mm) at 40 °C, with a flow rate of 0.8 mL/min. Injection volume 2 μ L. The following method was used (A: 0.1% ammonium hydroxide (0.01%) in water, B: acetonitrile): Gradient: 0 min 95% A, 5% B; 3 min 5% A, 95%

B; 3.1 min 0% A 100% B; 3.5 min 0% A 100% B; 3.51 min 95% A, 5% B; 4.5 min 95% A, 5% B. UV chromatograms were obtained at 210. Conversions were calculated from a calibration curve of authentic standards, following the production of product(s). Retention times in min (FMOC derivatized samples): butan-2-amine (3.19), DMBAmine (3.45), DMHamine (3.63).

2.16.5 Reverse-phase UPLC analysis of conversions

Samples were analysed using an Agilent 1290 infinity II Reverse-phase UPLC (diode array detector) on an Agilent Eclipse plus C18 column (1.8 μ m, 2.1 x 50 mm) at 20 °C with a flow rate of 0.5 mL/min. Injection volume 2 μ L. The following method was used (A: 0.1% TFA in water, B: 0.1% TFA in acetonitrile): Gradient: 0 min 95% A 5% B; 0.5 min 95% A 5% B; 1 min 60% A 40% B; 3.5 min 10% A 90% B; 3.51 min 0% A 100% B; 4.5 min 0% A 100% B; 4.6 min 95% A 5% B; 5.1 min 95% A 5% B. UV chromatograms were obtained at 210, 250, and 265 nm. Retention times in min: MBA (1.28), acetophenone (1.81), FMOC-DMBAmine (3.43), FMOC-DMHamine (3.77). Conversions were calculated from a calibration curve of authentic standards, following the production of product(s).

2.16.6 Chiral GC-FID

Samples were basified by adding 1:10 sodium hydroxide (5 M), optionally saturated with sodium chloride for highly water-soluble amines, and extracted into 2 x 500 μ L ethyl acetate. Extracted samples were derivatized with 20 μ L each triethylamine and acetic anhydride (under these conditions *N*-arylamines are not acetylated) and analysed by GC-FID: Thermo Scientific™ Trace™ 1310 GC equipped with an Agilent CHIRASIL-DEX CB (25 m x 0.25 mm x 0.25 μ m) column:

Method 1: 0 min 40 °C, 2 min 40 °C, 7.5 min 150 °C, 12.5 min 150 °C, 14.167 min 200 °C, 18.167 min 200 °C. Injector temperature 230 °C, split ratios 1:10 to 1:100, continuous flow 1.7 mL/min, FID temperature 250 °C, injection volume 1 μ L. Helium was used as carrier gas. Retention times in min (primary amines acetylated): (*S*)-*o*-fluoro- α -methylbenzylamine (11.2), (*R*)-*o*-fluoro- α -methylbenzylamine (11.1), (*S*)-1-aminoindan (14.9), (*R*)-1-aminoindan (15.0), (*S*)-4-phenylbutan-2-amine (14.7), (*R*)-4-phenylbutan-2-amine (14.8), (*S*)-hexan-2-amine (8.0), (*R*)-hexan-2-amine (8.1), (*S*)-tetrahydrofuran-3-amine (9.6), (*R*)-tetrahydrofuran-3-amine (9.5), (*S*)-tetrahydrothiophene-3-amine (13.1), (*R*)-tetrahydrothiophene-3-amine (13.0), (*S*)- α -Ethylbenzylamine (13.3), (*R*)- α -Ethylbenzylamine (13.4), SMBA (12.3), RMBA (12.6), (*S*)-2-phenoxyisopropylamine (14.7), (*R*)-2-phenoxyisopropylamine (14.8); hexan-2-one (4.8), acetophenone (7.3), (*S*)-3-(hexan-2-amino)toluene (12.0), (*R*)-3-(hexan-2-amino)toluene (12.1), 3-isopropylaminotoluene (8.5).

Method 2: 0 min 40 °C, 1 min 40 °C, 4 min 100 °C, 5 min 100 °C, 15 min 110 °C, 16 min 110 °C, 17.8 min 200 °C, 22.8 min 200 °C. Injector temperature 230 °C, split ratios 1:3 to 1:100, continuous flow 1.7 mL/min, FID temperature 250 °C, injection volume 1 µL. Helium was used as carrier gas. Retention times in min (acetylated): (*S*)-2-aminobutane (8.5), (*R*)-2-aminobutane (8.7).

2.16.7 Chiral RP-HPLC (primary amines)

Alternatively, samples were derivatized with FMOC-Cl (100 µL sample, 200 µL borate buffer (100 mM, pH 9), 400 µL FMOC-Cl (15 mM in acetonitrile)), diluted 5-fold with acetonitrile and aq. HCl (0.2%) (1:1 v/v) and analysed by reverse-phase HPLC (diode array detector) on a Phenomenex Lux Cellulose-2 chiral column (5 µm, 44.6 x 250 mm); injection volume 2-20 µL, at ambient temperature with a flow rate of 1 mL/min, with the following isocratic methods: A: 0.1% TFA in water, B: 0.1% TFA in acetonitrile: Tetrahydrofuran-3-amine and butan-2-amine: 40% A 60% B; tetrahydrothiophene-3-amine 45%A 55%B; serine: 55% A 45% B. Retention times in min (FMOC derivatized, 265 nm): (*S*)-tetrahydrofuran-3-amine (11.3), (*R*)-tetrahydrofuran-3-amine (12.3), (*S*)-tetrahydrothiophene-3-amine (19.3), (*R*)-tetrahydrothiophene-3-amine (18.6), (*S*)-2-aminobutane (13.0), (*R*)-2-aminobutane (11.7), L-serine (6.0), D-serine (6.2; shoulder: acetophenone (6.4)).

Retention times of each enantiomer were identified by comparing to commercially available samples (either a racemate and one enantiomer, or both enantiomers), except for phenoxypropan-2-amine, where (*S*)-2-phenoxyisopropylamine synthesized using the *Halomonas elongata* transaminase¹⁰ was used (see below), tetrahydrothiophene-3-amine, where a commercial racemate and (*S*)-tetrahydrothiophene-3-amine synthesized from L-methioninol according to the procedure by Pan *et al.*²² was used, and serine, where commercial L-serine was used.

2.16.8 Chiral RP-HPLC (*N*-arylamines)

2.16.8.1 Anhydrous FMOC-derivatization

For purified compounds, approx. 1 mg or 1 µL were used. For EtOAc extracts of reactions, an aliquot containing approx. 1 mg was taken, the solvent removed *in vacuo*, and the remaining residue was used. The compounds, sodium bicarbonate (approx. 10 mg), and FMOC-Cl (approx. 10 mg) were added to 100 µL ACN and incubated at 30 °C, 900 rpm for 16-24 h. 20 µL of the supernatant were diluted with 980 µL ACN/HCl (0.2%) (1:1 v/v) and analysed by chiral RP-HPLC.

2.16.8.2 Solvent-less Acetylation or Trifluoroacetylation

For purified compounds, approx. 1 mg or 1 μL were used. For EtOAc extracts of reactions, an aliquot containing approx. 1 mg was taken, the solvent removed *in vacuo*, and the remaining residue was used. To this sample were added triethylamine (20 μL) and acetic anhydride (20 μL), or a mixture (60 μL) triethylamine and trifluoroacetic anhydride (2:3 *v/v*), prepared on ice (exotherm!). After incubation at 30 °C, 900 rpm for 16-24 h, 10 μL of the supernatant was diluted with 990 μL ACN/HCl (0.2%) (1:1 *v/v*) and analysed by chiral RP-HPLC.

2.16.8.3 Chiral RP-HPLC method

The samples, derivatized and prepared as described above, were analysed by reverse-phase HPLC (diode array detector) on a Phenomenex Lux Cellulose-2 chiral column (5 μm , 44.6 x 250 mm); injection volume 2-20 μL , at ambient temperature with a flow rate of 1 mL/min, with the following isocratic methods: A: 0.1% TFA in water, B: 0.1% TFA in acetonitrile (Retention times in min):

20% A, 80% B; Fmoc-derivatized: (*S*)-3-(hexan-2-amino)toluene (11.7), (*R*)-3-(hexan-2-amino)toluene (11.1);

50% A, 50% B; acetylated: (*S*)-4-(*o*-F-methylbenzylamino)anisole (14.2), (*R*)-4-(*o*-F-methylbenzylamino)anisole (15.2);

55% A, 45% B; acetylated: (*S*)-3-(*o*-F-methylbenzylamino)toluene (26.9), (*R*)-3-(*o*-F-methylbenzylamino)toluene (28.5), (*S*)-3-(2-phenoxyisopropylamino)toluene (35.3), (*R*)-3-(2-phenoxyisopropylamino)toluene (38.1);

60% A, 40% B; trifluoroacetylated: (*S*)-4-(*o*-F-methylbenzylamino)benzotrile (38.8), (*R*)-4-(*o*-F-methylbenzylamino)benzotrile (40.1), Rot1-(*S*)-4-(*o*-F-methylbenzylamino)isoquinoline (16.7), Rot1-(*R*)-4-(*o*-F-methylbenzylamino)isoquinoline (20.0), Rot2-(*S*)-4-(*o*-F-methylbenzylamino)isoquinoline (21.6), Rot2-(*R*)-4-(*o*-F-methylbenzylamino)isoquinoline (22.2) (Rot1, Rot2 = rotamers 1 & 2);

70% A, 30% B; trifluoroacetylated: (*S*)-3-(*o*-F-methylbenzylamino)pyridine (32.5), (*R*)-3-(*o*-F-methylbenzylamino)pyridine (37.5).

2.17 References

1. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
2. Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. Rational assignment of key motifs for function guides in silico enzyme identification. *Nat. Chem. Biol.* **6**, 807–813 (2010).
3. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
4. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
5. Trott, O. & Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **31**, 455–461 (2009).
6. DeLano, W. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **700**, (2002).
7. Planchestainer, M. *et al.* Continuous flow biocatalysis: production and in-line purification of amines by immobilised transaminase from *Halomonas elongata*. *Green Chem.* **19**, 372–375 (2017).
8. *Immobilization of Enzymes and Cells. Methods in Molecular Biology* vol. 1051 (Humana Press Inc., 2013).
9. Schätzle, S., Höhne, M., Redestad, E., Robins, K. & Bornscheuer, U. T. Rapid and sensitive kinetic assay for characterization of ω -transaminases. *Anal. Chem.* **81**, 8244–8248 (2009).
10. Cerioli, L., Planchestainer, M., Cassidy, J., Tessaro, D. & Paradisi, F. Characterization of a novel amine transaminase from *Halomonas elongata*. *J. Mol. Catal. B: Enzym.* **120**, 141–150 (2015).
11. Davies, M. T. A universal buffer solution for use in ultra-violet spectrophotometry. *Anal.* **84**, 248–251 (1959).
12. Copeland, R. A. *Enzymes. A Practical Introduction to Structure, Mechanism, and Data Analysis.* (Wiley-VCH, Inc., 2000). doi:10.1002/0471220639.
13. Evans, P. R. An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **67**, 282–292 (2011).
14. Powell, H. R., Battye, T. G. G., Kontogiannis, L., Johnson, O. & Leslie, A. G. W. Integrating macromolecular X-ray diffraction data with the graphical user interface iMosflm. *Nat. Protoc.* **12**, 1310–1325 (2017).
15. Łyskowski, A. *et al.* Crystal structure of an (R)-selective ω -transaminase from *Aspergillus terreus*. *PLoS one* **9**, e87350 (2014).
16. Vagin, A. & Teplyakov, A. MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
17. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **67**, 235–242 (2011).
18. Davis, I. W. *et al.* MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, 375–383 (2007).
19. Jaźwiński, J. Studies on adducts of rhodium(II) tetraacetate and rhodium(II) tetratetrafluoroacetate with some amines in CDCl₃ solution using ¹H, ¹³C and ¹⁵N NMR. *J. Mol. Struct.* **750**, 7–17 (2005).
20. Ipaktschi, J. Reduktion von Oximen mit Natriumborant in Gegenwart von Übergangsmetallverbindungen. *Chem. Ber.* **117**, 856–858 (1984).
21. Moss, Neil; Gauthier, Jean; Ferland, J.-M. The enantioselective Synthesis of neopentylamine derivatives. *Synlett* 142–144 (1995).
22. Pan, X. *et al.* An efficient synthesis of (R)-3-aminothiolane. *J. Chem. Res.* **35**, 729–730 (2011).

23. Dehmlow, E. V. & Westerheide, R. (S)-3-Aminothiolane: A New Chiral Building Block. *Synthesis* **10**, 947–949 (1992).
24. Koszelewski, D., Lavandera, I., Clay, D., Rozzell, D. & Kroutil, W. Asymmetric synthesis of optically pure pharmacologically relevant amines employing ω -transaminases. *Adv. Synth. Catal.* **350**, 2761–2766 (2008).
25. Franchini, C. *et al.* Optically Active Mexiletine Analogues as Stereoselective Blockers of Voltage-Gated Na⁺ Channels. *J. Med. Chem.* **46**, 5238–5248 (2003).
26. Knutsen, L. J. S. *et al.* N-substituted adenosines as novel neuroprotective A1 agonists with diminished hypotensive effects. *J. Med. Chem.* **42**, 3463–3477 (1999).

3 Transaminases for the synthesis of 2-aminobutane and 2,2-dimethylhexan-3-amine

3.1 2-aminobutane and 2,2-dimethylhexan-3-amine as target molecules

The two amines (2-aminobutane and 2,2-dimethylhexan-3-amine (DMHamine)) that are the focus of this chapter had been identified by Johnson Matthey (JM) as potentially interesting targets for enzymatic synthesis. The reason for this initial interest was the presence of both groups in drug candidates in clinical trials (**Figure 3-1**). 2-Aminobutane is a feature of the drug XL888, a HSP90 inhibitor currently in phase I clinical trials for use in conjunction with antibodies to treat cancers.¹⁻⁴ DMHamine is present in Veledimex that at the outset of this work was in phase I clinical trials for a gene therapy against glioma.⁵ It has since moved into phase 2 clinical trials.⁶ In addition, 2-aminobutane is a feature of other clinically relevant molecules (for example in PET tracers⁷ and featured in recent structure-activity relationship studies (SARs)^{8,9}). In each case it is usually one enantiomer that is of interest. It also appears in the herbicide Bromacil, which is commonly available as the racemate although the different enantiomers have been shown to have different effects.¹⁰ For both amines, the similar size of the R-groups (ethyl vs methyl, *t*-butyl vs *n*-propyl) might pose additional challenges to the enantioselectivity of the enzymes, further increasing the interest.

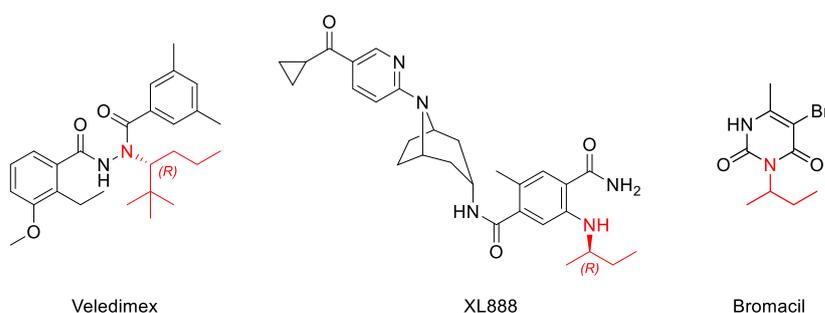


Figure 3-1: Structures of Veledimex and XL888, the chiral amine is highlighted in red.

While only two syntheses of DMHamine have been reported (in both cases via reduction of the corresponding chiral *N*-(α -methylbenzyl)imine followed by removal of the α -methylbenzyl group by hydrogenation),^{11,12} the synthesis of 2-aminobutane is more common, where the similar size of the R-groups (ethyl vs methyl) poses a challenge in achieving high enantioselectivity. Asymmetric syntheses of 2-aminobutane via reductive amination using α -methylbenzylamine (MBA)¹³ (71% yield, 74% *ee*) or using Ellman's sulfonamide¹⁴ (93% yield, 86% *de* of intermediate) as a chiral auxiliary have been reported.

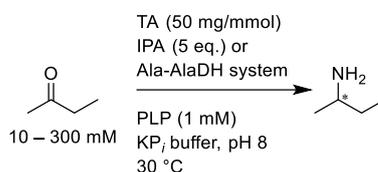
The requirement of several steps, chiral auxiliaries, and the use of “non-green” solvents¹⁵ (THF and DCM) makes these approaches less attractive today. A recently reported asymmetric synthesis employing a chiral ruthenium catalyst only achieved 22% *ee*, among the lowest enantioselectivities for the substrates tried by the authors.¹⁶ Kinetic resolutions of 2-aminobutane using CAL-B have also been investigated, requiring ethyl decanoate as the acyl donor to achieve high enantioselectivity (25% yield as the phosphate salt, 99.8% *ee* (*S*) on a 5 g scale).¹⁷ A dynamic kinetic resolution using the more conventional methyl 3-methoxypropionate resulted in 84% yield, 74% *ee* of the (*R*)-amide, on a 1 mmol scale.¹⁸ Kinetic resolutions employing several transaminases have also been investigated.^{19–21} Notably, using whole *E. coli* cells expressing the (*S*)-selective transaminase from *Vibrio fluvialis*, 53% conversion and 98% *ee* (*R*) were achieved at 400 mM concentration (under reduced pressure), although a large quantity of enzyme had to be used (100 U/ml).²² Deracemization on a 50 mM scale through sequential addition of two enantiocomplementary transaminases was demonstrated by Kroutil and co-workers, giving either enantiomer (>99% conversion, >99% *ee*) depending on the order the transaminases are added.²³

Syntheses of 2-aminobutane from butanone with transaminases has also been reported. Mutti *et al.* reported the use of whole cell systems screening four RTAs²⁴ and three STAs,²⁵ using alanine as the amine donor and either alanine dehydrogenase (AlaDH) or lactate dehydrogenase (LDH) for pyruvate removal. On a 50 mM scale, the highest conversion achieved with STAs (*V. fluvialis* TA, LDH system) was 31%, 98% *ee*. For the RTAs, the highest conversion was achieved with the *Aspergillus terreus* transaminase (*At*RTA) (AlaDH system), giving 69% conversion and >99% *ee*. Koszelewski *et al.* reported high conversions (98% and 97%) and high *ee* (>99% (*R*) and 98% (*S*)) with the Codexis enzymes ATA-117 and ATA-113, respectively (using enantiopure alanine and the Codexis LDH kit for pyruvate removal).²⁶ However, intensification beyond 50 mM and the use of the cheaper isopropylamine (IPA) as the amine donor have not been investigated. In addition, isolation of the product 2-aminobutane has not been attempted, presumably due to its high solubility in water and high volatility. Thus, the aim of the work in this chapter was the development of transaminase catalysed process giving 2-aminobutane and DMHamine, with a focus on scalability, use of IPA, and product isolation.

3.2 Asymmetric synthesis of 2-aminobutane from butanone

3.2.1 Screening of STAs and RTAs against butanone

Initial screening was performed with 12 transaminases. In addition to 8 transaminases from Johnson Matthey's portfolio, the STAs from *Halomonas elongata* (HEwT)²⁷ and *Chromobacterium violaceum* (CvSTA),²⁸ and the RTAs from *Thermomyces stellatus* (TsRTA, see **chapter 5**)²⁹ and *Aspergillus terreus* (AtRTA)³⁰ were also screened. All enzymes were used as lyophilized cell-free extracts (cfes) and tested with 3 substrate concentrations (10, 100 and 300 mM), using 5 equivalents of IPA while keeping a constant loading of enzyme relative to substrate of 50 mg/mmol (**Scheme 3-1**). By using a constant relative amount of enzyme, at each concentration the same level of conversion corresponds to the same turnover number (TON). In addition to TON, different concentrations of active enzyme in the crude enzyme mix (**Figure 3-2**) will also affect the conversion. While not correcting for active enzyme content may give a less precise indication of the specific performance of the catalyst, this was a deliberate choice as the crude lyophilizate is sold commercially by weight. Thus, higher expression becomes as valid a parameter to assess suitability of the enzymatic preparation as would be enzyme activity.



Scheme 3-1: Screening conditions for the transaminase catalysed synthesis of 2-aminobutane, starting from butanone. Ala-AlaDH system: L-Ala (1.2 eq.) or DL-Ala (2.4 eq.), AlaDH (25 mg/mmol), GDH (25 mg/mmol), D-glucose (1.2 eq.), NH₄Cl (2.4 eq.), and NAD⁺ (1 mM). Ala-AlaDH: KP_i (300 mM), IPA: KP_i (50 mM) Enzymes were employed as lyophilized cfes.

Almost all enzymes screened exhibited an increase in conversions at the 5 h mark when going from 10 mM to 100 mM concentration (**Figure 3-3**). This is likely due to the enzymes not having high affinity for butanone (although a similar effect with IPA is also possible). For some enzymes, such as STA-1 and -2, the conversions at 5 h decreased when scaling up to 300 mM, for others, such as *RTA-43, CvSTA, and, most notably, HEwT conversions further increased at the highest substrate concentration. While lower 5 h conversions always corresponded to lower final conversions, the contrary was not always true for the enzymes performing better at the 5 h mark with (as is the case for STA-14 and RTA-25).

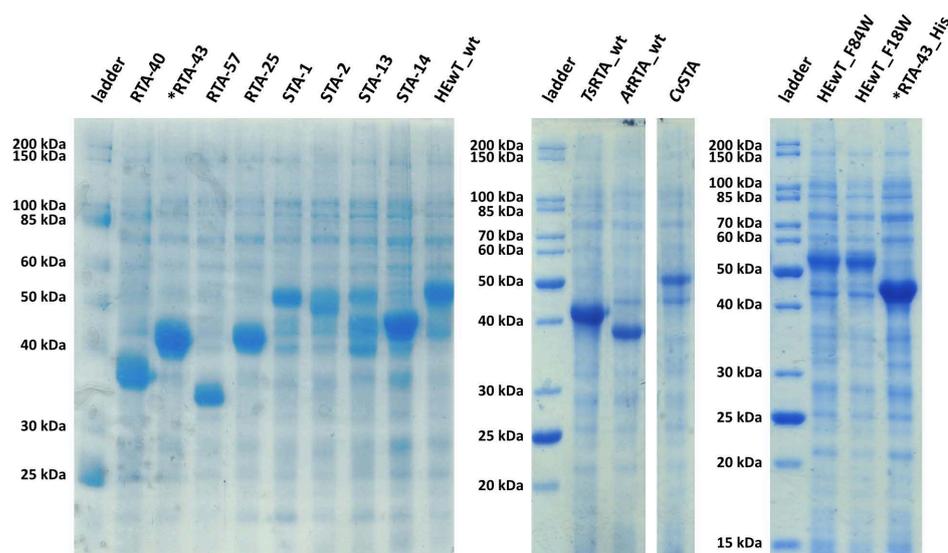


Figure 3-2: SDS-PAGE of all lyophilized cfes used. The same weight of lyophilized cfe was loaded in each well (1 μ L of a 25 mg/mL solution).

Of the STAs tested, HEwT exhibited the best performance in terms of conversions, reaching almost 70% at the 300 mM scale in 5 h, and achieving the highest final conversions (almost 90%). However, in terms of enantioselectivity it did poorly, yielding only 45% *ee* (*S*) (**Table 3-1**). CvSTA, which also showed a significant albeit smaller increase in conversion from 100 mM to 300 mM, also had an unsatisfactory 80% *ee*. For the RTAs, only *RTA-43 and RTA-25 showed acceptable levels of conversion, with *RTA-43 giving better scalability. For all RTAs, the enantioselectivities were excellent.

To evaluate whether the poor performing enzymes were due to the choice of IPA, the screening was repeated with alanine as the amine donor, employing an AlaDH provided by JM and a glucose dehydrogenase (GDH) from *Bacillus megaterium* for pyruvate recycling/removal. For the STAs, 1.2 equivalents of L-Ala were used as the donor, for the RTAs 2.4 equivalents of racemic alanine were used (effectively using 1.2 equivalents of D-Ala, **Scheme 3-1**). While all three enzymes showed improved performance using the new amine donor, in particular STA-13 and RTA-57 (**Figure 3-4**), they fell short of the best performing enzymes employing IPA. STA-13 showed a poor *ee* of 85%, and an increase in conversion from 100 mM to 300 mM as had been observed for CvSTA and HEwT.

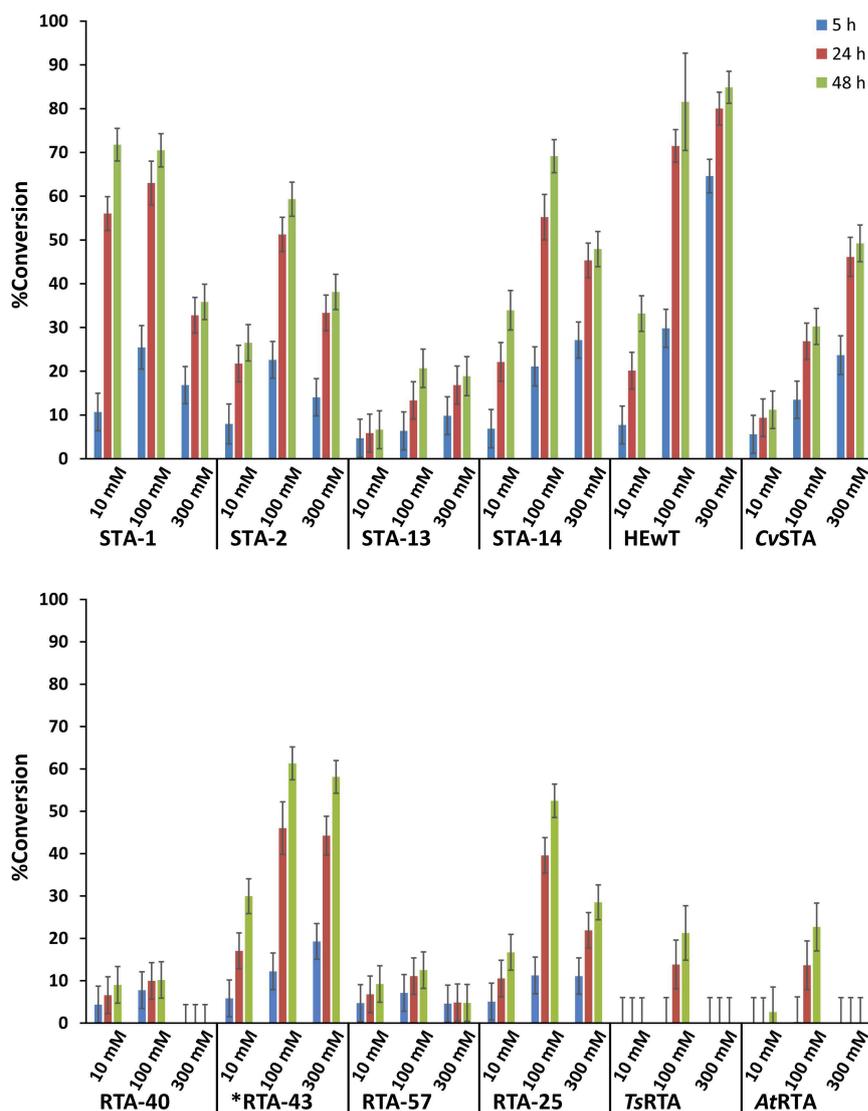
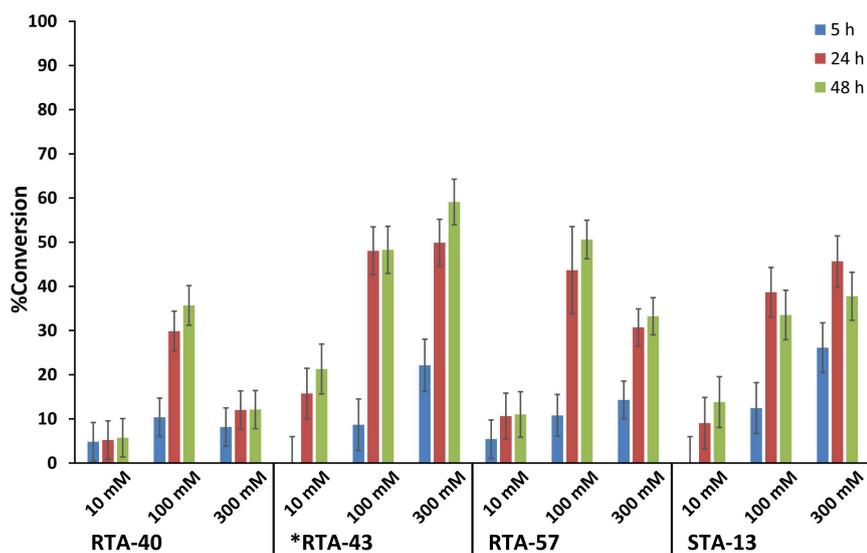


Figure 3-3: Transaminase screen: biotransformations of butanone, using IPA as the amine donor. Samples were taken after 5 h, 24 h, and 48 h, and conversions were determined following the production of 2-aminobutane by RP-HPLC, after FMOC derivatization, using a calibration curve. All experiments were carried out in duplicate, error bars represent the standard error (SE), and include the uncertainty associated with the calibration curve.^{31,32}

Table 3-1: Enantiomeric excess (*ee*) of the biotransformations employing IPA as the amine donor.

	10 mM	100 mM	300 mM
STA-1	96% (<i>S</i>)	96% (<i>S</i>)	99% (<i>S</i>)
STA-2	traces (<i>S</i>)	99.5% (<i>S</i>)	99.6% (<i>S</i>)
STA-13	traces (<i>S</i>)	85% (<i>S</i>)	84% (<i>S</i>)
STA-14	97% (<i>S</i>)	94% (<i>S</i>)	96% (<i>S</i>)
HEwT	68% (<i>S</i>)	45% (<i>S</i>)	46% (<i>S</i>)
CvSTA	traces (<i>S</i>)	79% (<i>S</i>)	81% (<i>S</i>)
RTA-40	traces (<i>R</i>)	traces (<i>R</i>)	traces (<i>R</i>)
*RTA-43	>99% (<i>R</i>) ^a	99.4% (<i>R</i>)	99.4% (<i>R</i>)
RTA-57	traces (<i>R</i>)	>99% (<i>R</i>) ^a	traces (<i>R</i>)
RTA-25	traces (<i>R</i>)	99.3% (<i>R</i>)	99.6% (<i>R</i>)
TsRTA	>99.5% (<i>R</i>) ^{ab}	>99.5% (<i>R</i>) ^{ab}	n/d ^b
AtRTA	>99.5% (<i>R</i>) ^{ab}	>99.5% (<i>R</i>) ^{ab}	n/d ^b

All *ees* represent the average of duplicates.
^a other enantiomer not detected
^b chiral RP-HPLC
n/d = not detected

**Figure 3-4:** Transaminase screen: biotransformations of butanone, using the alanine-AlaDH system. Samples were taken after 5 h, 24 h, and 48 h, and conversions were determined following the production of 2-aminobutane by RP-HPLC, after FMOC derivatization, using a calibration curve. All experiments were carried out in duplicate, error bars represent the standard error (SE), and include the uncertainty associated with the calibration curve.^{31,32}

All RTAs were outperformed by *RTA-43 in terms of conversion (all RTAs had excellent *ee*, **Table 3-2**), and it was chosen for the production of (*R*)-2-aminobutane. For this enzyme, employing either alanine or IPA as the amino donor made a negligible difference, and so the less complex IPA system was chosen. Among the STAs, STA-2 showed the highest *ee* of (*S*)-2-aminobutane; yet had slightly lower conversions at 100 mM and 300 mM compared to STA-14 and STA-14, respectively, both of which still had >90% *ee*. More importantly, all three enzymes showed a significant drop in conversion going from 100 mM to 300 mM. Engineering of HEwT to increase its *ee* while retaining the excellent performance was therefore attempted.

Table 3-2: Enantiomeric excess (*ee*) of the biotransformations employing alanine as the amine donor.

	10 mM	100 mM	300 mM
STA-13	n.d.	86% (<i>S</i>)	86% (<i>S</i>)
RTA-40	n.d.	>99.5% (<i>R</i>) ^a	>99.5% (<i>R</i>) ^a
*RTA-43	n.d.	99.8% (<i>R</i>)	>99.5% (<i>R</i>) ^a
RTA-57	traces (<i>R</i>)	>99.5% (<i>R</i>) ^a	>99.5% (<i>R</i>) ^a

All *ees* represent the average of duplicates.
^a other enantiomer not detected
n.d. = not detected

To this end, two phenylalanine residues in the small pocket (**Figure 3-5**; F18 and F84) were independently changed to the larger tryptophan (chosen to maintain the aromaticity, as the F18 position in particular is sensitive to changes³³). The residue F84 has been shown previously to be a hot-spot for enlarging the small pocket in other STAs,^{34,35} whereas no previous attempt to engineer an STA to enhance enantioselectivity for small substrates has been reported.

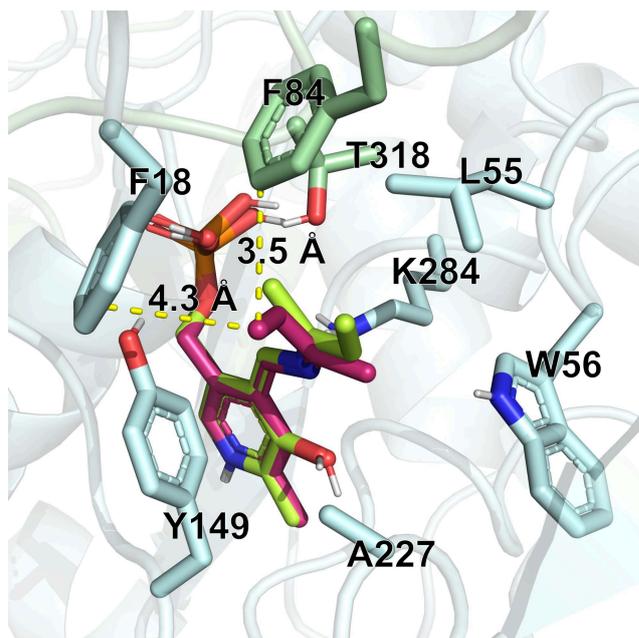


Figure 3-5: Docked quinonoid intermediate in the pro-(*S*) (lime) and pro-(*R*) (pink) orientation into the active site of HEwT (PDB: 6GWI).³⁶ Side-chains of residues surrounding the butane-amine moiety are shown as sticks; the two different subunits are coloured cyan and green. Distances of the ethyl-group in the pro-(*R*) orientation to F18 and F84 are shown. Docking was carried out using Autodock/vina,³⁷ the figure was generated using open source PyMOL 2.1.0.

While HEwT_F18W had significantly lower conversions (in particular at 300 mM) and unchanged enantioselectivity (50-55% *ee* (*S*)), HEwT_F84W exclusively produced (*S*)-2-aminobutane (*ee* >99.5%). However, speed and final conversion were reduced compared to wild-type HEwT, but remained higher than for the other STAs at 300 mM (compare **Figure 3-3** and **Figure 3-6**). Conversions at 10 mM and 100 mM were reduced significantly, while the increase in conversion from 10 mM all the way through to 300 mM was maintained. This suggests that the increased steric bulk in the small pocket may have reduced affinity for butanone, which would also explain the slightly lower final conversions at 300 mM. However, attempts at measuring the K_m or the specific activity using the acetophenone assay³⁸ were unsuccessful, as the activity was too low to be measured (no activity was detected at up to 1 mg/mL of enzyme and using 10-350 mM of butanone). The specific activities with SMBA and pyruvate were 0.01 U/mg and 0.2 U/mg for the lyophilized crude HEwT_F84W and wild-type, respectively.

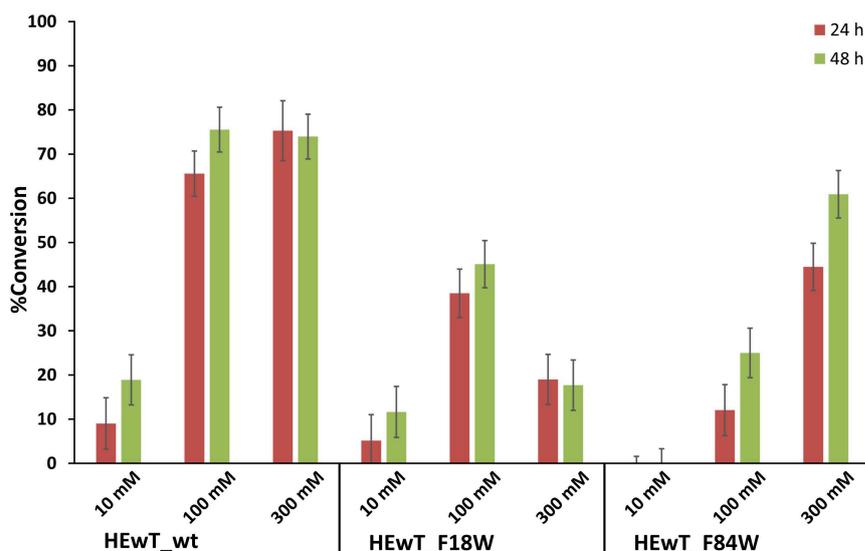


Figure 3-6: HEwT mutants: biotransformations of butanone, using IPA as the amine donor. Samples were taken after 24 h and 48 h, and conversions were determined following the production of 2-aminobutane by RP-HPLC, after FMOc derivatization, using a calibration curve. All experiments were carried out in duplicate, error bars represent the standard error (SE), and include the uncertainty associated with the calibration curve.^{31,32}

Additional substrate scope and effect on enantioselectivity for the two mutants are investigated in further in **Chapter 4**.

3.2.2 Determination of immobilization and operating conditions in flow

As the batch reactions required high enzyme loading (15 mg/mL at the 300 mM concentration), enzyme immobilization and use in continuous flow, in a packed-bed reactor, were investigated, to facilitate reusability of the enzyme. Additionally, moving to flow has previously resulted in increased performance.³⁹ Initial immobilizations were carried out on Co^{II}-derivatized Relizyme EP403/S beads, using a loading of 100 mg_{lyo}/g_{resin}. For both *RTA-43 and HEwT_F84W, complete immobilization was achieved within 3-4 h, giving recovered activities of 50% and 30%, respectively, where the recovered activity is defined as the fraction of activity contained in the resin compared to the total activity offered to the resin. Using these resins, operating conditions in flow were explored. Given the long reaction times despite the high enzyme loading employed in batch, a reaction in a single pass proved unfeasible with the equipment available. Thus, a recirculating strategy was adopted, where the output was fed back into the reaction mixture. Increasing the reaction temperature to 37 °C resulted in a loss of activity of the immobilized enzyme, whereas both enzymes were able to operate for >24 h without any detectable loss in activity at 30 °C. Conversions after recirculating four column volumes of reaction mixture for 24 h (6 h contact time; approx.

9×40 min residence time) were 50% and 30% for *RTA-43 and HEwT_F84W, respectively. Addition of DMSO or increasing the concentration of IPA did not improve performance.

Thus, the loading of enzyme on the resin was increased. Loadings of 170 mg_{lyo}/g_{resin} and 400 mg_{lyo}/g_{resin} were achieved for *RTA-43 and HEwT_F84W, respectively. At these loadings between 70-100% of the enzyme was immobilized (as determined from the residual activity of the supernatant; also see **Figure 3-7**). Further increasing the loading did not increase the amount of enzyme immobilized. Recovered activities reached up to 38% for *RTA-43 and up to 64% for HEwT_F84W. With these resins, conversions up to 75% were achieved in 4 h (approx. 12×20 min residence time), with no further increase observed with longer contact times.

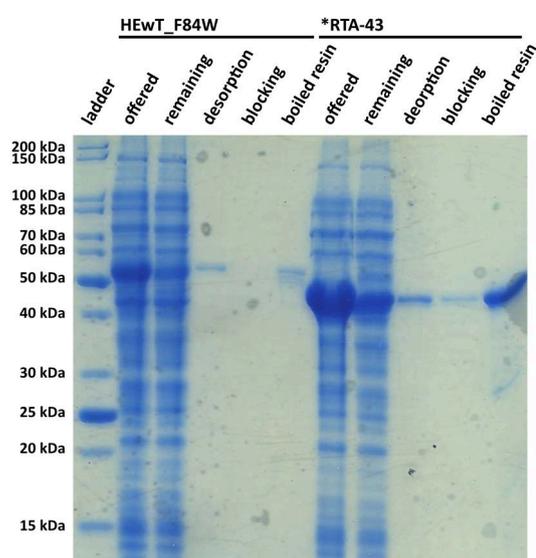


Figure 3-7: SDS-PAGE of the maximum loading immobilization, where “offered” refers to the protein solution prior to incubation with the resin and “remaining” refers to the same solution after incubation with the resin. “Desorption” and “blocking” refer to samples of the desorption and blocking buffer taken after incubation with the resin. “Boiled resin” refers to the buffer obtained from boiling resin in SDS-loading buffer for approx. 10 min. The amounts loaded in each well are directly comparable for each of the two enzymes.

3.2.3 Scaled-up synthesis of both enantiomers of 2-aminobutane in flow

Using these operating conditions, larger scale syntheses of both enantiomers of 2-aminobutane were carried out. By employing 6 g of resin, 46.4 mL of reaction mixture could be processed every 24h. For each of the enzymes, the reactor was left running continuously for one week, replacing the reaction mixture every 1-4 days, processing a total volume of 0.319 L, containing 6.90 g (95.7 mmol) of butanone. Conversions were determined each time the reaction mixture was changed (**Figure 3-8**). For HEwT_F84W, conversions were stable

over the duration of seven days (66±5%), and no loss in activity was detected in the resin afterwards. On the other hand, for *RTA-43 conversions dropped from 69±5% to 54±5% after seven days, and the specific activity of the resin decreased by 35%. Compared to the initial batch process, both catalyst productivity and Space-time yield have been improved, as envisaged (Table 3-3).

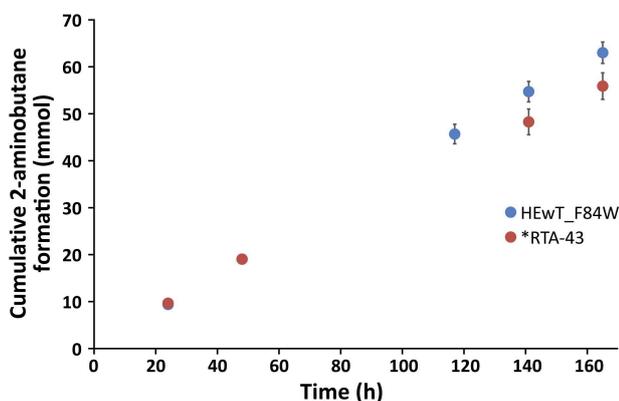


Figure 3-8: Cumulative production of 2-aminobutane during a 7-day run of the continuous flow process, employing either HEwT_F84W or *RTA-43. At each time point, the solution was replaced with a fresh reaction mixture and the conversion determined using RP-HPLC. Error bars represent standard error of the calibration curve. Data gaps between 48 h and 120 h or 144 h for HEwT_F84 and *RTA-43, respectively, are due to periods of lab-closure, where a proportionally larger volume of liquid was circulated.

Table 3-3: Comparison of key process parameters between the batch and flow process.

		Specific reaction rate ($\mu\text{mol mg}^{-1} \text{h}^{-1}$) ^a	Catalyst productivity ($\mu\text{mol mg}^{-1}$) ^b	Space-time yield ($\mu\text{mol mL}^{-1} \text{h}^{-1}$) ^c
Batch	HEwT_F84W	0.293	14.1	4.40
	*RTA-43	0.267	12.8	4.00
Flow	HEwT_F84W	0.159	26.3	49.4
	*RTA-43	0.332	54.8	43.8

^a Calculated according to $n_{BA}/m_{lyo-cfe} \times t_{total}$, at similar levels of conversions (achieved after 48 h in the case of batch reactions (Figure 3-3)), where n_{BA} corresponds to the overall amount of 2-aminobutane produced (as determined by HPLC), $m_{lyo-cfe}$ corresponds to the mass of lyophilized cfe used, and t_{total} corresponds to the total reaction time (48 h batch, 165 h flow). For the flow process, this has been calculated with respect to the amount of enzyme offered during the immobilization, neglecting any losses in activity.

^b Calculated according to $n_{BA}/m_{lyo-cfe}$.

^c Calculated as follows: batch: $n_{BA}/V_{rxn} \times t_{total}$, where V_{rxn} corresponds to the reaction volume.

Flow: $n_{BA}/V_{bed} \times t_{total}$, where V_{bed} corresponds to the reactor bed volume.

Given the highly volatile and highly water-soluble nature of 2-aminobutane, as well as the similar structure to IPA, purification by distillation appeared to be the best approach, avoiding both the need for water removal and the use of organic solvents. Alternative strategies, such as lyophilization for water removal or precipitation were also investigated.

However, preliminary tests had shown that when lyophilizing acidic solutions of 2-aminobutane and IPA in water, up to 25% of the initially added mass were lost, presumably due to the same apparent sublimation that is also observed for ammonium chloride. Precipitation of 2-aminobutane from the reaction mixture using tartaric acid⁴⁰ also proved to be ineffective (data not shown).

In a first vacuum distillation, the ketones were removed from an acidified reaction mixture. Subsequently, the remaining aqueous solution was alkalized, and the amines were distilled. However, separation of the amines during this initial distillation proved very challenging (**Table 3-4**). The compositions of the cuts obtained were determined by ¹H-NMR (**Figure 3-9**). Thus, a second fractional distillation (4-ball Snyder column) had to be employed on the amine mixture obtained from the second distillation (**Table 3-5**), giving 2.49 g of (*S*)-2-aminobutane and 1.97 g of (*R*)-2-aminobutane, corresponding to a yield of 35% and 28%, respectively (*ees* ≥99%). This represents a mass-loss of approx. 50%. While some of the apparently lost amine was contained in an impure cut (5% and 15% for (*S*)- and (*R*)-2-aminobutane, respectively), there was also loss through leaking joints and potentially during the initial vacuum distillation. However, this mass-loss should be reduced at a larger scale with a more effective (industrial) distillation set-up. Additionally, combining the second distillation and fractional distillation might further reduce this mass-loss.

Table 3-4: Masses and composition of the different cuts obtained during the first distillation.

	Cut 1		Cut 2		Cut 3	
	m (g)	2-AB ^a (mol%)	m (g)	2-AB ^a (mol%)	m (g)	2-AB ^a (mol%)
HEwT_F84W	17.2	6	7.44	18	1.81	35
*RTA-43	12.9	4	9.92	15	5.08	22

^a Composition determined by integrating the *CHN* peak in the ¹H-NMR.

Traces of acetone were also detected.

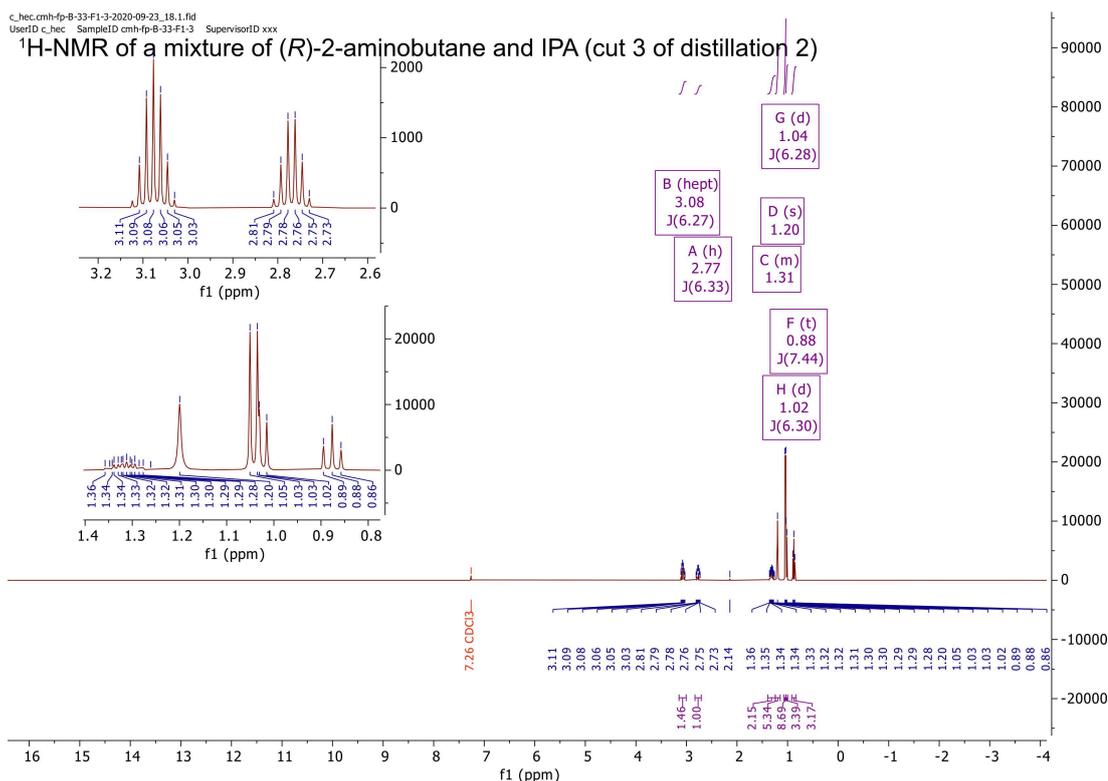


Figure 3-9: Representative ¹H-NMR spectrum obtained from a cut (cut 3 of distillation 2 for *RTA-43) during the purification of 2-aminobutane, showing the peaks of 2-aminobutane and IPA. For quantification, the proton attached to the α -carbon was used (multiplet B for IPA and multiplet A for 2-aminobutane).

Table 3-5: Masses and composition of the different cuts obtained during the second distillation.

	Cut 1		Cut 2		Cut 3		Cut 4	
	m (g)	2-AB ^a (mol%)						
HEwT_F84W	15.9	<1	5.17	1	1.33	16	2.49	≥99
*RTA-43	18.3	1	3.71	2	1.50	41	1.97	≥99

^a Composition determined by integrating the CHN peak in the ¹H-NMR.
Traces of acetone were detected in cuts 1-3.

The use of alternative amino donors, cadaverine and alanine, was also explored in flow. Employing 4 eq. of cadaverine at pH 9 under otherwise identical conditions, HEwT_F84W lost 60% and *RTA-43 lost 75% of activity over 24 h. Additionally, conversions for *RTA-43 were reduced to 25% (employing cadaverine in batch using 2 or 4 eq. at pH 8 or 9 resulted in no conversion with *RTA-43) while the conversion obtained with HEwT_F84W was unchanged (68%). The use of alanine in flow was hampered by the lack of compatible supports for the covalent immobilization of the transaminases, GDH, and AlaDH (or LDH), as attempts to immobilize the GDH on EP403/S resulted in low recovered activities (data not

shown). On the other hand, the transaminases were not operationally stable when immobilized on epoxy-agarose, which is the preferred support for GDH (data not shown).⁴¹

To assess the sustainability of this process, the atom economy and E-factor were calculated. The atom economy, defined as the molecular weight of the desired product divided by the molecular weight of all products, was 56%. The E-factor, which is defined as the mass of all waste (neglecting water) divided by the mass of product, was calculated with respect to the actual isolated yield obtained in this current work. The mass of the culture media and buffer salts used throughout, despite their benign nature (bio-renewable and biodegradable in the case of culture media) has been taken into account. Additionally, the supported enzyme catalyst used in the flow reaction is considered waste, despite being potentially usable for further cycles (in particular in the case of HEwT_F84W). Furthermore, the recovered IPA may be re-used and the unreacted butanone could in principle be separated from the waste acetone, but this has also been ignored in the calculation. Thus, this E-factor represents a “worst-case” estimate (see **Figure 3-10** for a breakdown); and has been calculated to be 48 for the HEwT_F84W process and 55 for the *RTA-43 process.

Chapter 3 — 2-Aminobutane and DMHamine

all masses in g (neglecting water)	For *RTA-43		For HEwT_F84W	
Mass of culture media /L	58.65		42.54	
	tryptone	12	Al ZYP mix	33.7
	yeast extract	24	MgSO ₄	0.24
	potassium phosphate monobasic	9.4	trace elements	1
	potassium phosphate dibasic	3.2	glycerol	5
	glycerol	5	glucose	0.5
	lactose	5	lactose	2
	antibiotic	0.05	antibiotic	0.1
Mass of cfe/Lculture	6.26		1.55	
Mass of culture media* /g cfe	9.37		27.53	
Mass of buffers+cofactor during cell lysis/g cfe	0.16		0.22	
	potassium phosphate	0.16	potassium phosphate	0.22
	PLP	0.00044	PLP	0.00062
Mass of buffers+cofactor+additives in preparation of resin/g resin	1.82		1.90	
	IDA	0.27		0.27
	sodium Borate	0.076		0.076
	NaCl	0.41		0.41
	Cobalt chloride	0.025		0.025
	glycine	0.90		0.90
	potassium phosphate	0.073		0.15
	PLP	0.00027		0.00049
	EDTA	0.067		0.067
Mass resin used	6.08		6.05	
Total Masswaste (buffers, media,...) during the production of immobilized catalyst	12.7		22.6	
Mass of buffers+cofactor in reaction (319 mL)	2.86		2.86	
	potassium phosphate	2.78	potassium phosphate	2.78
	PLP	0.079	PLP	0.079
Mass of Starting material (IPA and butanone) in reaction (319 mL)	52.6		52.6	
	IPA.HCl	45.7	IPA.HCl	45.7
	Butanone	6.9	Butanone	6.9
Mass of Starting materials – mass of isolated product	50.7		50.1	
Mass of HCl used during purification	2		2	
Mass of KOH used during purification	35		37	
Total amount of waste	109.3		120.7	
Mass of product	1.97		2.49	
E-factor	55		48	

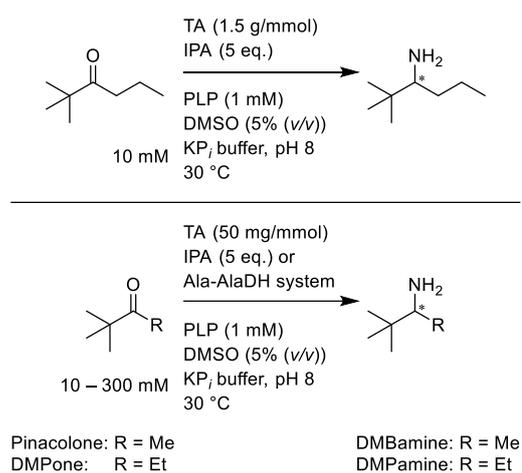
*by conservation of mass, this also accounts for cell debris and the non-buffer components of the lyophilized cfe

Figure 3-10: Spreadsheet for the calculation of the E-factors of the HEwT_F84W and *RTA-43 processes. Masses are colour coded for ease of legibility, masses in grey are a breakdown of the masses listed in black, which are then used to calculate the masses in red, which are then summed up to give final amount of waste, which is divided by the amount of product isolated (in red, boxed), to give the E-factor in green.

3.3 Screening and attempted engineering of transaminases toward the synthesis of 2,2-dimethylhexan-3-amine

3.3.1 Screening of STAs and RTAs against 2,2-dimethylhexan-3-one

The same panel of transaminases that was screened against butanone was screened against DMHone, however, screening was carried out on a 10 mM scale only to enable a higher loading of lyophilized enzyme (1500 g/mmol), with the addition of 5% DMSO to aid in solubility (**Scheme 3-2**). Initial screening with IPA (5 equivalents) showed no conversion for any of the enzymes, except for STA-14, where a small peak at the same retention time as DMHamine appeared in the LC-MS (as the FMOC derivative) but the identity of that peak could not be further confirmed and was of similar size to impurities. Thus, it was apparent that enzyme engineering was required. This was expected as it is well established that most TAs struggle with compounds where both groups are larger than a methyl group.



Scheme 3-2: Top: Screening conditions for the transaminase catalysed synthesis of DMHamine, starting from DMHone. Bottom: Screening conditions for the transaminase catalysed synthesis of the truncated analogues DMBamine and DMPamine, starting from pinacolone and DMPone, respectively. Ala-AlaDH system: L-Ala (1.2 eq.) or D-L-Ala (2.4 eq.), AlaDH (25 mg/mmol), GDH (25 mg/mmol), D-glucose (1.2 eq.), NH₄Cl (2.4 eq.), and NAD⁺ (1 mM). Ala-AlaDH: KP_i (300 mM), IPA: KP_i (50 mM) Enzymes were employed as lyophilized cfes.

To establish which of the TAs would be a suitable candidate for engineering, the panel was rescreened against pinacolone, an analogue of DMHone with the *n*-propyl group shortened to a methyl group. This screening was carried out under the conditions employed for butanone, with the addition of 5% DMSO (**Scheme 3-2**). TAs that had shown a preference for alanine with butanone (i.e. RTA-40, -57, and STA-13) were also screened with alanine, in addition to IPA, but showed either no or low conversion (ca. 5% conversion in the case of STA-13 with alanine). STA-1 and -2 showed low levels of conversion, STA-14, HEwT, *RTA-43 and RTA-25 showed moderate conversion (**Figure 3-11**). All other enzymes showed virtually

no conversion. In addition, STA-14 and RTA-25 showed a significant drop in conversion from 100 mM to 300 mM. The four most promising enzymes were next screened against 2,2-dimethylpentan-3-one. Screening was carried out under identical conditions, employing IPA as the amine donor. Traces of potential product were observed by HPLC for STA-14 and HEwT, which were deemed the most promising candidates for engineering, while no conversion was observed for either *RTA-43 nor RTA-57.

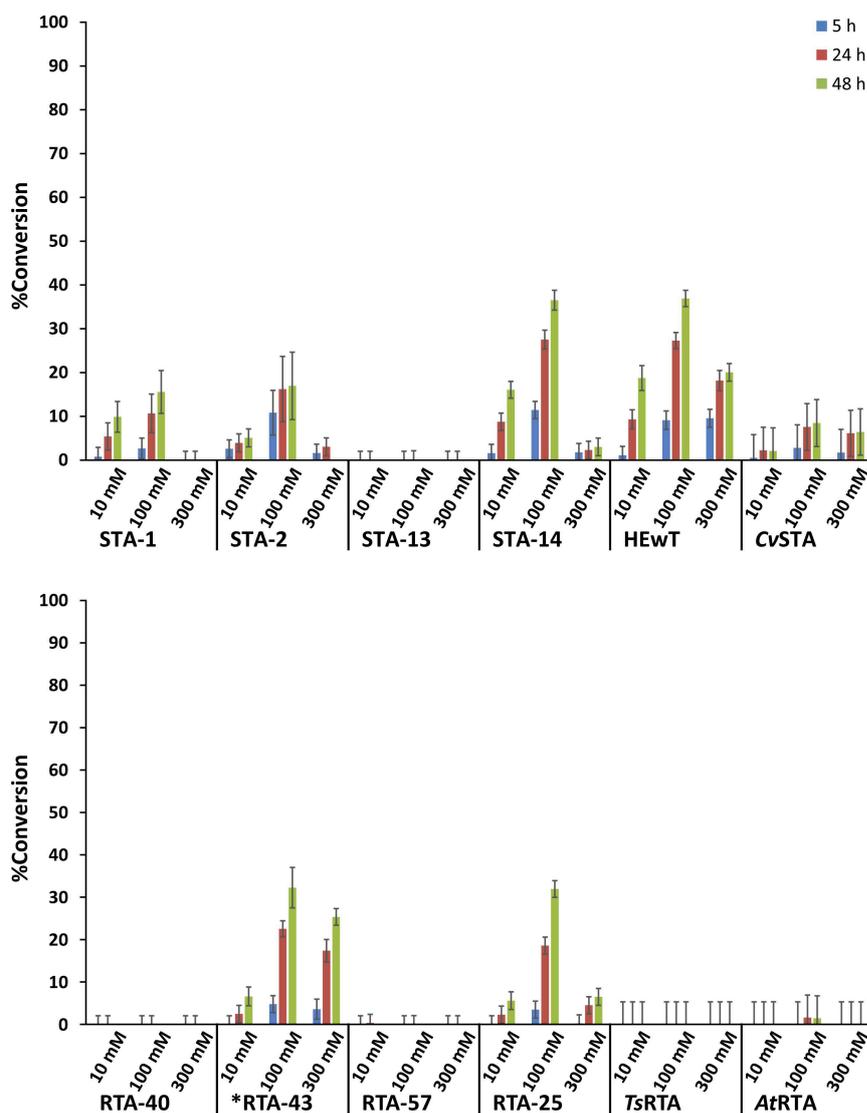


Figure 3-11: Transaminase screen: biotransformations of pinacolone, using IPA as the amine donor. Samples were taken after 5 h, 24 h, and 48 h, and conversions were determined following the production of DMBamine by RP-HPLC, after FMOC derivatization, using a calibration curve. All experiments were carried out in duplicate, error bars represent the standard error (SE), and include the uncertainty associated with the calibration curve.^{31,32}

3.3.2 Library design

As a crystal structure has been solved for HEwT (PDB entry: 6GWI),³⁶ and given the proprietary nature of STA-14, it was chosen as the initial template. Preliminary screening of

small (3000 variants) random libraries already present in the lab, employing a colony-based screen using *o*-xylylene diamine and background depletion with (*S*)MBA,³⁶ only returned false-positives (the screening methodology is explained in more detail in **Section 3.3.3**). Thus, a semi-rational approach focussing on the active site architecture was preferred. In particular, a CASTing approach⁴² was chosen over iterative site saturation mutagenesis in order to include potentially synergistic effects of mutations.

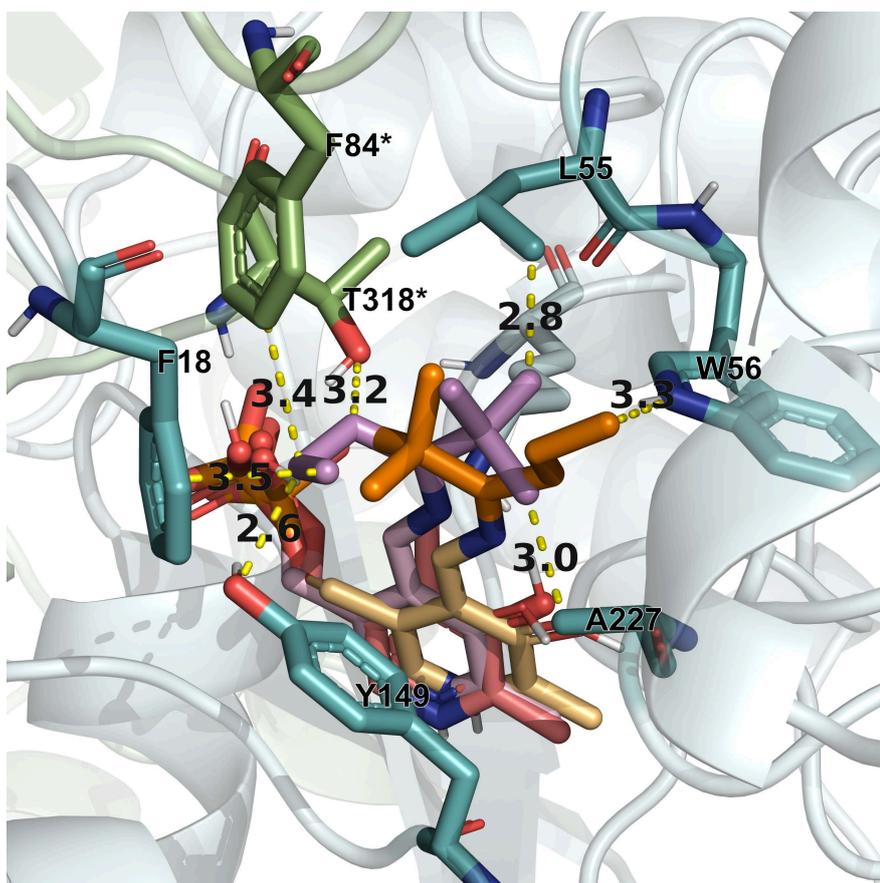


Figure 3-12: Residues within 4 Å of the DMHone moiety of the docked quinonoid intermediate (both a pro-(*S*) (periwinkle) and a pro-(*R*) (orange) conformation) are shown, as well as the crystal-structure PLP (strawberry)). The two chains forming the dimer are coloured teal and green, residues belonging to the other monomer are denoted with an asterisk.

To generate the CASTing libraries, residues surrounding the active site have to be identified, grouped into sets of 2 or 3 residues, and suitable alphabets for each position have to be determined. Thus, docking of the quinonoid intermediate of PLP and DMHone (chosen for its rigidity) in both a pro-(*S*) and a pro-(*R*) conformation into the active site of HEwT was performed using AutoDock Vina.³⁷ The position of the docked substrates was compared to the position of PLP. For the structures matching the PLP orientation in the active site, seven residues within 4 Å of the DMHone moiety were identified (**Figure 3-12**): F18, and F84* in the small pocket, W56 and A227 in the large pocket, and L55 in between the pockets.

Residues T318* and Y149 were not included due to their role in PLP binding (Hydrogen-bond to phosphate and π -stacking with the pyridine ring, respectively).

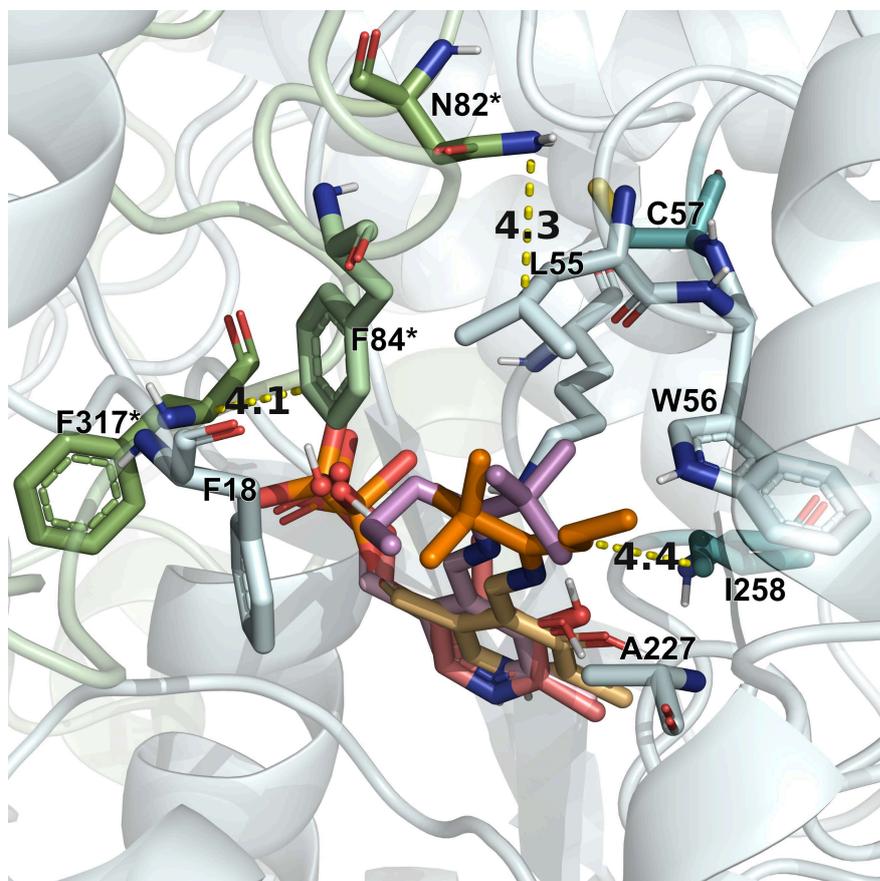


Figure 3-13: Residues included in the CAST libraries. The docked quinonoid intermediate (both a pro-(*S*) (periwinkle) and a pro-(*R*) (orange) conformation are shown, as well as the crystal-structure PLP (strawberry)). The two chains forming the dimer are coloured teal and green, residues belonging to the other monomer are denoted with an asterisk.

Co-conservation analysis was performed on Pfam PF00202 (STAs, 24998 sequences, aligned using Clustal Omega⁴³) using the MISTIC server.⁴⁴ Based on the strongest co-conservations observed, four additional residues were included in the list of target sites (**Figure 3-13**): F317*, N82*, C57, and I258. Using the co-conservation data, as well as physical proximity, residues were grouped as follows: F18+F84+F317, W56+C57+F317, L55+N82, and A227+I258. The alphabets at each position were chosen based on chemical similarity to the natural residue, chemical intuition (i.e. small/hydrophobic residues to carve out space and/or interact with the aliphatic chains of DMHone), and conservation at each position based on the co-conservation analysis. Lastly, additional amino acids were included if this enabled the use of fewer primers (**Table 3-6**). Since HEwT accepts pinacolone as a substrate, it was suspected that mutations in the small pocket in particular would be needed, thus a

larger alphabet was chosen for the “small pocket set” F18+F84+F317, with 12 residues in each position giving a library size of 1728 variants.

Libraries were generated using the QuikChange Lightning multi site directed mutagenesis kit, employing all primers in a single reaction per library. Primers were designed so melting points upstream and downstream of the mutation site matched as closely as possible, while ending the primers with at least one G or C at the 3'-terminus. Libraries were assessed by Sanger sequencing, using both the Q_{pool} parameter as well as Quick Quality Control pie charts, as described by Acevedo-Rocha, *et al.*⁴⁵ (Table 3-7). In some instances, the quality of the libraries generated was low; however, overall, the quality was deemed acceptable.

Table 3-6: The positions chosen for mutagenesis, the alphabet at each position, the set they have been assigned to for CASTing, and the degenerate codon(s) used for mutagenesis.

Position	Alphabet	Set	Codon
F18 & F84	AILVFIHSTDNP	1	NHC
F317	AILVFIHSTDNP	1 & 2	NHC
W56	GAWSC	2	KST, TGG
C57	GAILVFSTCP	2	NYC, KGC
L55	GAILVSC	3	KST+VTA
N82	AILVFIHSTDNP	3	NHC
I258	AILVFSTP	4	NYC
A227	GASC	4	KST

Library sizes: 1: 1728; 2: 600; 3: 84; 4: 32 variants.

Table 3-7: Evaluation of the libraries generated. The theoretical target composition at each randomized position in comparison to that determined from Sanger sequencing, and the calculated Q_{pool} , which is normalized on a scale of 0-1, where 1 means perfect randomization.

	G C	A T	Set 1			Set 2		Set 3		Set 4	
			F18	F84	F317	W56	C57	F317	L55	N82	A227
Theoretical											
	Q_{pool}	0.73	0.63	0.76	0.90	0.91	0.77	0.87	0.67	0.87	0.60

3.3.3 Library screening—colony-based screening

Screening the libraries at 3-fold oversampling would require screening of 7400 colonies, corresponding to 51 days of continuous HPLC time assuming a 10-minute method. This was not feasible and so it was decided to employ the colony-based screen as a pre-screen. Using the robotics available at JM, a larger sample of positive clones from the screen could be rescreened, and it was hoped that even though a large number of false positives were expected, real positives could also be found.

As already described in **Section 1.2.2.1**, the colony-based screening relies on the formation of a coloured product from the first half-reaction with the amine donor. However, even without the addition of a ketone substrate, colonies will turn coloured.^{36,46} This is presumably due to pyruvate in the cells acting as the ketone substrate in the reaction. This background can be greatly reduced by first incubating the cells with a non-colourogenic amine donor and allowing the reaction using any carbonyl acceptors in the cell to go to completion (**Figure 3-14**). This allows for colour formation due to activity toward the ketone substrate rather than endogenous pyruvate to be discriminated, and has previously allowed the identification of HEwT mutants with increased activity toward acetophenones and aliphatic heterocyclic ketones.^{36,47} However, while the background is reduced, it is not completely eliminated. As the cells remain alive, more pyruvate is produced from normal cell metabolism. In addition, due to the reversibility of transamination reactions, any ketone

formed from the sacrificial amine donor remaining in the cells can in principle also cause colour-formation.

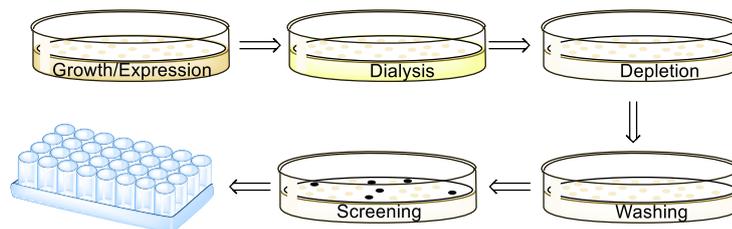


Figure 3-14: Colony based screening methodology: Following growth and expression on LB-agar (+IPTG) plates, colonies are transferred on a nitrocellulose membrane to dialysis plates (agarose (2% (w/v)), PLP (0.1 mM), Tris-HCl (10 mM), pH 8) and dialysed overnight at 4 °C, to arrest cell metabolism and wash away nutrients. Membranes are transferred to filter paper soaked in a non-colorogenic amine donor (e.g. MBA and/or IPA), incubated at RT for ca. 90 min, transferred to a filter paper soaked in buffer to wash away the amine donor and keto-product (ca. 90 min, RT or 4 °C), and finally incubated with the colorogenic amine donor and target ketone substrate. Coloured colonies are then picked and grown in 96-deep-well-plates (DWPs) for further assessment.

There are several parameters that can be varied throughout the screening. In particular, the nature of the sacrificial amine donor and the colourogenic amine donor can be changed. In addition, the incubation times and temperatures for each step can be varied. Initial conditions included depletion with (*S*)MBA (10 mM) followed by IPA (50 mM) for 45 min each, followed by washing with buffer for 90 min. Then, *o*-XDA (10 mM) and DMHone (50 mM) were added and any colonies showing colour formation were marked for about 1 h. Colonies were picked and grown in a 96-DWP (out of ca. 2000 colonies, 95 were picked). These were re-screened (after sonication and clarification) against DMHone (50 mM) and IPA (250 mM). Unexpectedly, using the UPLC at JM, traces of a peak matching DMHamine (FMOC derivatized) could be detected even for the wild-type control (however, apart from matching retention times, the identity of that peak could not be confirmed further). For most wells, this peak was decreased or unchanged (**Figure 3-15**); however, 35 wells showing an increase >20% were identified, out of these, two showed an increase >50% (peaks were still extremely small, and conversions could not be quantified, but were likely in the order of <0.1%). Re-screening was repeated using (*S*)MBA (50 mM) as the amine donor in case the variants picked had lost activity toward IPA. Overall, conversions were lower than with IPA, and followed a larger spread (29 wells showed a >2-fold improvement and two wells showed a >3-fold improvement). Correlation was observed between the (*S*)MBA and IPA experiments ($r = 0.83$, $p < 0.0001$), giving support to the hypothesis that variation across wells was not due to random noise (**Figure 3-16**).

IPA	1	2	3	4	5	6	7	8	9	10	11	12
A	0.11	0.12	0.10	0.09	0.13	0.08	0.09	0.09	0.11	0.11	0.09	0.11
B	0.09	0.11	0.10	0.10	0.08	0.09	0.08	0.08	0.08	0.07	0.06	0.08
C	0.07	0.10	0.09	0.10	0.10	0.08	0.11	0.07	0.08	0.10	0.10	0.09
D	0.11	0.10	0.11	0.10	0.07	0.10	0.08	0.09	0.08	0.09	0.10	0.09
E	0.09	0.09	0.08	0.09	0.08	0.08	0.08	0.07	0.09	0.09	0.07	0.06
F	0.11	0.09	0.08	0.09	0.10	0.09	0.08	0.06	0.07	0.07	0.07	0.09
G	0.06	0.09	0.08	0.10	0.07	0.06	0.07	0.10	0.06	0.09	0.12	0.06
H	0.08	0.08	0.08	0.07	0.10	0.06	0.07	0.10	0.12	0.10	0.11	0.08

(S)MBA	1	2	3	4	5	6	7	8	9	10	11	12
A	0.07	0.08	0.08	0.05	0.08	0.05	0.06	0.07	0.06	0.06	0.06	0.06
B	0.06	0.10	0.08	0.06	0.03	0.07	0.06	0.04	0.06	0.03	0.02	0.03
C	0.04	0.08	0.07	0.08	0.07	0.06	0.06	0.03	0.03	0.07	0.06	0.06
D	0.06	0.08	0.06	0.08	0.06	0.06	0.05	0.07	0.05	0.11	0.07	0.04
E	0.05	0.07	0.03	0.05	0.05	0.04	0.04	0.05	0.06	0.05	0.04	0.03
F	0.06	0.06	0.03	0.06	0.09	0.03	0.04	0.04	0.05	0.08	0.03	0.05
G	0.02	0.05	0.05	0.05	0.03	0.02	0.04	0.05	0.03	0.05	0.05	0.03
H	0.03	0.03	0.02	0.03	0.06	0.03	0.03	0.05	0.06	0.04	0.04	0.03

Figure 3-15: HEwT re-screening: DMHamine production (arb. U.) from DMHone (50 mM) across one 96-DWP after 24 h, using IPA (250 mM, top) or (S)MBA (50 mM, bottom) as the amine donor. Well H12: wild-type, all other wells: colonies picked from screening. Progressively greener shading indicates relative increase over wild-type. For clarity, the shading is only shown for wells with a >20% (IPA) or >100% ((S)MBA) increase. PLP (1 mM), DMSO (10% (v/v)), KP_i-buffer (50 mM), pH 8, 30 °C, using cfes. Samples were taken after 24 h and analysed by RP-HPLC, after FMOC derivatization, using a calibration curve (the uncertainty from the calibration curve was larger than the absolute value for DMHamine, and thus the numbers should not be seen as an accurate estimate of the actual conversion).

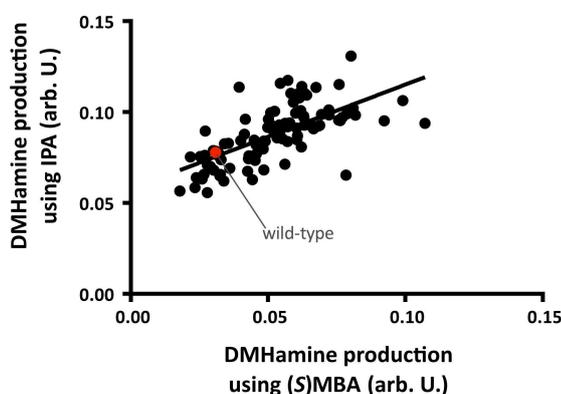


Figure 3-16: DMHamine production using IPA vs using (S)MBA as the amine donor in the 96-DWP shown in **Figure 3-15**, showing correlation between both donors ($r = 0.83$, $p < 0.0001$).

The screening conditions were then varied in order to try to reduce the high number of false positives. (S)MBA and IPA were combined in the background depletion step and washing was carried out at 4 °C vs room temperature to reduce metabolic activity. The screening was also carried out with *p*-nitrophenethylamine (*p*-NPEA), and DMSO concentrations were varied between 10% and 20%. Partial cell lysis by snap-freezing at -80 °C was also attempted. Noteworthy is that colour formation with *p*-NPEA did not occur

for several hours and colonies were picked the next day. Overall, fewer colonies had shown colour formation and the residual background observed in a wild-type control plate was lower (as has been reported before³⁶). To increase throughput, re-screening was occasionally carried out using whole cells, however it was soon realized that this had poor correlation with cfe data. Overall, a similar spread in conversions was observed in all plates irrespective of these changes. In total, ca. 5000 colonies were screened, 282 picked for re-screening and 35 candidates were identified as potentially having increased conversion over the wild-type. (An additional 8000 colonies were screened using LB+lac agar plates for growth instead of IPTG induction; however, this caused unacceptably high background in the wild-type and out of 96 colonies re-screened, none were taken further).

3.3.4 *Screening of additional transaminases*

Based on these results, libraries of STA-14 were also generated. For this, a proprietary definition of the active site of STA-14 was used (taking into account flexibility of the protein). The sequence was then aligned with a list of patented Codexis enzymes and additional hot spots were identified. Conservation analysis was carried out using consurf (400 representative sequences from UNIREF90, 35-95% homology, aligned using MAFFT-L-INS-I). Residues that were highly conserved in this analysis or in the Codexis set of sequences were discarded from the analysis. Finally, conservation and residues in the Codexis set were used to determine alphabets. In the end, the libraries were composed of very similar residues to those generated for HEwT, albeit with smaller alphabets at each position reducing the library size. Libraries were synthesised using a proprietary protocol at JM and analysed as before. The quality of libraries in general appeared lower than that of libraries generated for HEwT (data not shown). In particular, additional mutations were observed within the primer region and sequencing was not as clean. Regardless, libraries were screened using the same procedure outlined above. Ca. 1200 colonies were screened, and 147 colonies were picked for re-screening. Consistent with the previous results obtained in Nottingham, STA-14 itself showed a higher base-level of conversion over HEwT. Similar to HEwT, most colonies picked showed reduced or comparable conversion to the wild-type, and 13 wells with increased conversion were identified. As fewer than 192 colonies (two 96-DWPs) were picked, in one of the plates 35 wells were left empty and 14 wells were inoculated with the wild-type, as a more extensive control (**Figure 3-17**). The negative control wells were consistently among the lowest observed, and the conversions observed for the wild-type were also consistent, with one exception: the highest conversion observed in that plate corresponded to a wild-type in H12, and was nearly 2-fold higher than observed for the other wild-types.

24 h	1	2	3	4	5	6	7	8	9	10	11	12
A	0.13	0.16	0.14	0.17	0.18	0.15	0.14	0.14	0.14	0.16	0.18	0.15
B	0.12	0.07	0.06	0.07	0.08	0.12	0.08	0.12	0.11	0.10	0.12	0.13
C	0.14	0.06	0.07	0.06	0.09	0.11	0.11	0.11	0.11	0.12	0.11	0.13
D	0.10	0.15	0.12	0.12	0.05	0.06	0.05	0.06	0.06	0.14	0.14	0.13
E	0.07	0.07	0.05	0.07	0.06	0.05	0.05	0.06	0.04	0.08	0.06	0.08
F	0.13	0.12	0.13	0.14	0.13	0.13	0.13	0.12	0.11	0.11	0.14	0.14
G	0.07	0.06	0.07	0.07	0.07	0.08	0.08	0.07	0.07	0.07	0.09	0.10
H	0.07	0.06	0.07	0.08	0.07	0.06	0.06	0.07	0.08	0.08	0.09	0.23

Figure 3-17: STA-14 re-screening: DMHamine production (arb. U.) from DMHone (50 mM) across one 96-DWP after 24 h, using IPA (250 mM, top) or (S)MBA (50 mM, bottom) as the amine donor. Wells A1, F1-12, and H12: wild-type STA-14, wells E1-12, G1-12, and H1-11: blank, all other wells: colonies picked from screening. Progressively greener shading indicates relative increase over wild-type, progressively redder shading indicates relative decrease over wild-type. PLP (1 mM), DMSO (10% (v/v)), KP_i -buffer (50 mM), pH 8, 30 °C, using whole cells. Samples were taken after 24 h and analysed by RP-HPLC, after FMOc derivatization, using a calibration curve (the uncertainty from the calibration curve was larger than the absolute value for DMHamine, and thus the numbers should not be seen as an accurate estimate of the actual conversion).

Given that no mutants with appreciable conversion were identified, additional transaminases were screened: the Codexis transaminase kit, and 3 patented Codexis enzyme sequences. In all cases, conversions were extremely low with the highest conversions only being ca. 5-fold larger than for wild-type STA-14. Interestingly, the enzymes in this set performing best with pinacolone did not correspond to the ones performing best with DMHone. In particular, ATA117-Rd11⁴⁸ showed very poor conversion, even though it is able to accept very bulky substrates. Thus, it became apparent that the mutagenesis and screening effort that were likely required to induce significant activity toward DMHone into a transaminase was not feasible with the resources available.

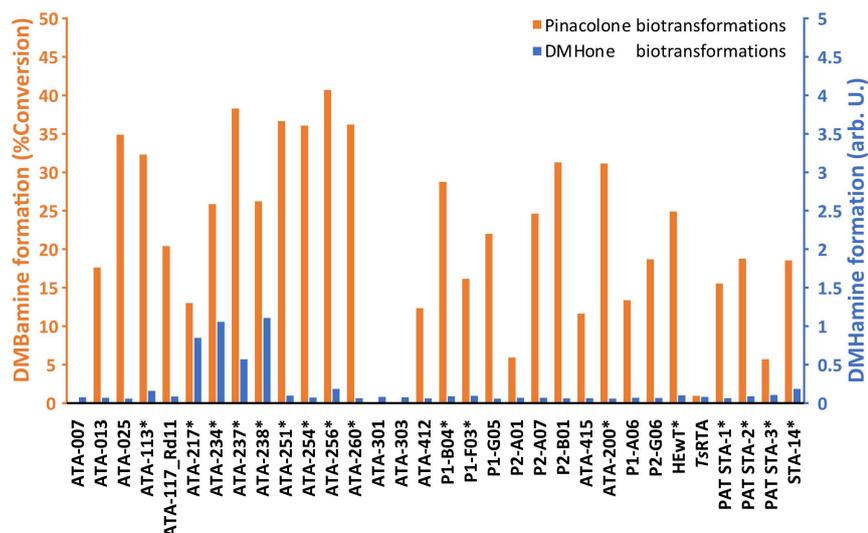


Figure 3-18: Additional screening of transaminases. Codexis kit enzymes, three patented Codexis sequences, HEWT, TsRTA, and STA-14 were screened against both pinacolone and DMHone (50 mM), using IPA (250 mM) as the amine donor, PLP (1 mM), DMSO (10% (v/v)), KP_i-buffer (50 mM), pH 8, 30 °C, using lyophilized cfs (10 mg/mL). Samples were taken after 24 h and analysed by RP-UPLC, after FMOC derivatization, using a calibration curve (the uncertainty from the calibration curve was larger than the absolute value for DMHamine, and thus the numbers should not be seen as an accurate estimate of the actual conversion). No replicates. Stars indicate (S)-selective transaminases.

3.3.5 Limitations of the screening

Further analysing the variants picked during the screen, a subset of colonies was sequenced. Several sequences that were picked due to their promising “conversions” with DMHone in fact contained frame-shift mutations and were truncated and likely not functional. Those sequences should not have shown colour formation in the colony-based screen (despite its shortcomings it is very robust regarding the need for an overexpressed transaminase that can accept the colourogenic donor). Two explanations appear plausible: two colonies, one with an active transaminase and one with the truncated sequence were inadvertently picked and only the non-functional sequence remained following the subsequent growth in DWPs until plasmid extraction for sequencing. Alternatively, the frameshift mutation may have occurred during the growth in the DWPs (an expression strain being used). Several wells included multiple sequences at the targeted positions and some wells included random mutations away from the targeted positions, making both explanations plausible. It should be noted that colonies had to be picked by hand and the colony-picker at JM was not suitable for use with the membranes used for the screening. Additionally, plasmids had to be propagated in BL21(DE3) due to the T7-promoter rather than in a DNA propagation strain, up until plasmid isolation prior to sequencing.

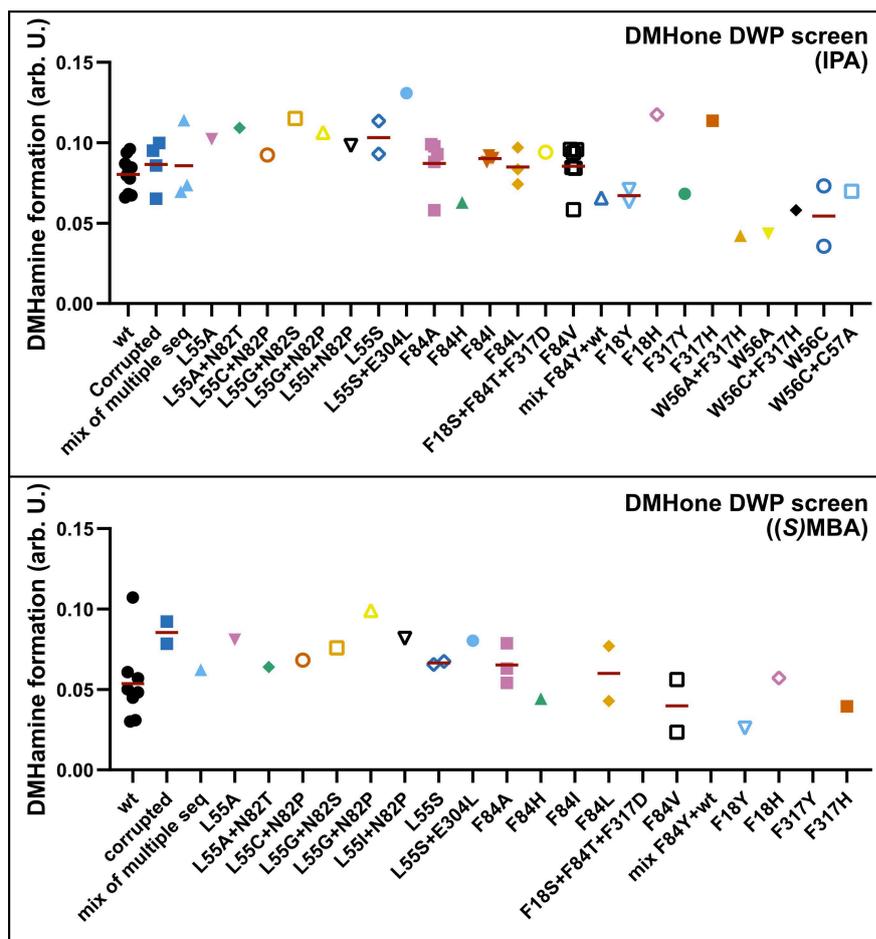


Figure 3-19: HEwT variants: DMHamine formation in the DMHone DWP screen (IPA and (S)MBA donor). Not all DWPs were screened with (S)MBA.

Furthermore, variation was observed between wells containing identical variants. Plotting the conversions with DMHone against all HEwT sequences identified, shows that no meaningful difference exists between all variants (**Figure 3-19**). Furthermore, plotting the conversions obtained with acetophenone (or the specific activities with SMBA), a lot of variation is also observed in most variants, but especially in the wild-type (**Figure 3-20**). Clearly, understanding this variation is crucial to enable successful screening of similar libraries in the future. Given that conversion data between multiple copies of the same plate appeared to be correlated (**Figure 3-16**), it seems likely that variability in growth and/or expression may be the underlying factor (as opposed to simply random variation). This could be caused by varying viabilities of the glycerol stocks in each well of the master plate. Indeed, reduced pellet size and even complete lack of growth were observed in some wells.

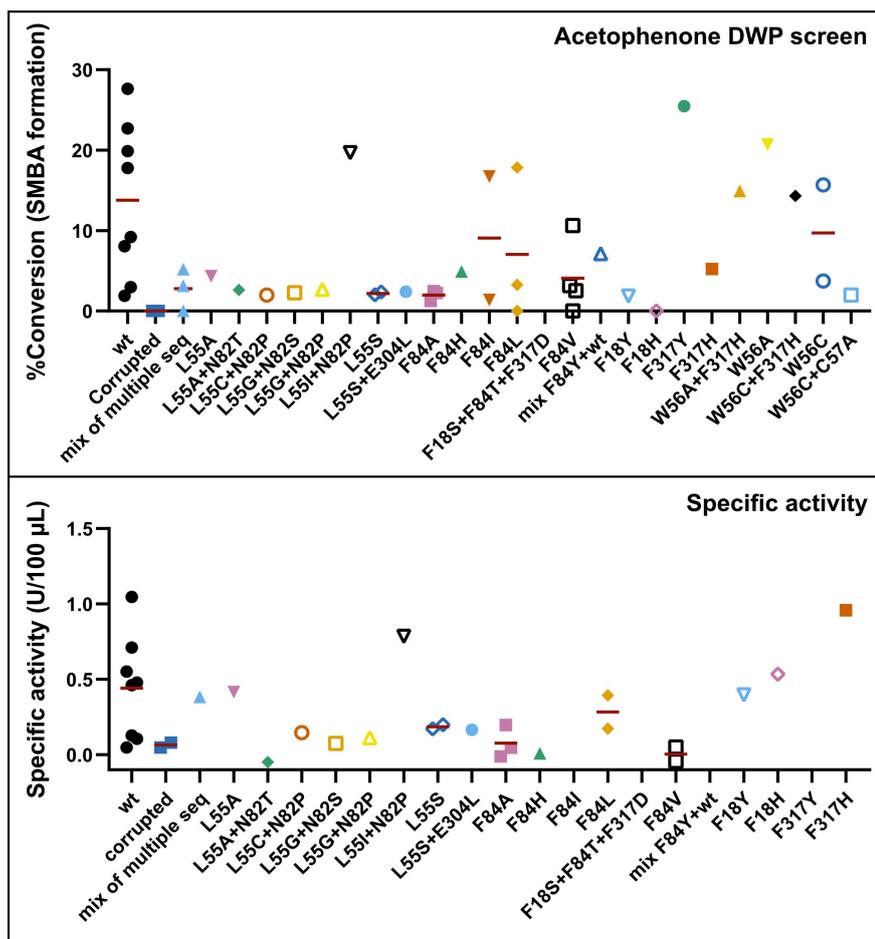


Figure 3-20: HEWT variants: Conversions with acetophenone in DWPs, and specific activities of the cfes. Activity was not determined in all DWPs.

For STA-14, a similar analysis of all sequences variants showed improvement for a few variants and more consistent results overall (**Figure 3-21**). This may be due to an undisclosed promotor that permitted use of the DNA propagation strain NEB 10-beta, resulting in a more stable plasmid.

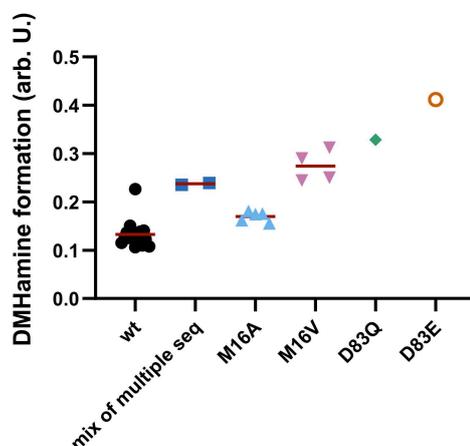


Figure 3-21: STA-14 variants: DMHamine formation in the DMHOne DWP screen (IPA donor).

3.4 Conclusions and future work

A panel of transaminases was screened for the synthesis of 2-aminobutane and DMHamine. While suitable candidates were identified for the synthesis of the former and successfully applied in a multi-gram synthesis in continuous flow, only very low (<1%) conversions were observed for the synthesis of the later. Thus, the directed evolution of two candidates that showed promising conversion on a truncated analogue, HEwT and STA-14, was attempted. However, challenges in the screening, background in a colony-based screen and irreproducibility (likely due to non-uniform growth or expression) in DWP screen did not allow for suitable variants to be identified. This was particularly the case for HEwT. Thus, it appears that in order to engineer a suitable transaminase catalyst for the synthesis of DMHamine, a more reliable screening protocol is needed. Additionally, it appears likely that multiple rounds of evolution on a wider set of residues, if not the entire protein, would be needed. For this, higher throughput screening and/or machine learning would likely be needed. Lastly, STA-14 appears to be the more promising candidate for such future work.

3.5 References

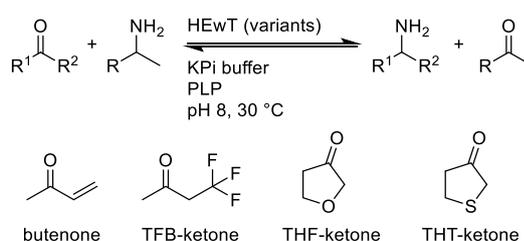
1. Azimi, A. *et al.* Targeting CDK2 overcomes melanoma resistance against BRAF and Hsp90 inhibitors. *Mol. Syst. Biol.* **14**, e7858 (2018).
2. Pembrolizumab and Hsp90 inhibitor XL888 in Treating Patients with Advanced Gastrointestinal Cancer. <https://www.cancer.gov/about-cancer/treatment/clinical-trials/search/v?id=NCI-2016-01594&r=1>.
3. XL888 + Vemurafenib + Cobimetinib for Unresectable BRAF Mutated Stage III/IV Melanoma. <https://clinicaltrials.gov/ct2/show/NCT02721459>.
4. Study of XL888 With Vemurafenib for Patients With Unresectable BRAF Mutated Stage III/IV Melanoma. <https://clinicaltrials.gov/ct2/show/NCT01657591>.
5. Chiocca, E. A. *et al.* Regulatable interleukin-12 gene therapy in patients with recurrent high-grade glioma: Results of a phase 1 trial. *Sci. Transl. Med.* **11**, (2019).
6. Study of Ad-RTS-hIL-12 + Veledimex in Combination With Cemiplimab in Subjects With Recurrent or Progressive Glioblastoma. <https://clinicaltrials.gov/ct2/show/NCT04006119>.
7. Berroterán-Infante, N. *et al.* (R)-[18F]NEBIFQUINIDE: A promising new PET tracer for TSPO imaging. *Eur. J. Med. Chem.* **176**, 410–418 (2019).
8. Vandyck, K. *et al.* Synthesis and Evaluation of N-Phenyl-3-sulfamoyl-benzamide Derivatives as Capsid Assembly Modulators Inhibiting Hepatitis B Virus (HBV). *J. Med. Chem.* **61**, 6247–6260 (2018).
9. Osborne, J. *et al.* Discovery of potent and selective 5-azaindazole inhibitors of leucine-rich repeat kinase 2 (LRRK2) – Part 1. *Bioorg. Med. Chem. Lett.* **29**, 668–673 (2019).
10. Chen, Z., Zou, Y., Wang, J., Li, M. & Wen, Y. Phytotoxicity of chiral herbicide bromacil: Enantioselectivity of photosynthesis in *Arabidopsis thaliana*. *Sci. Total Environ.* **548–549**, 139–147 (2016).
11. Moss, Neil; Gauthier, Jean; Ferland, J.-M. The enantioselective Synthesis of neopentylamine derivatives. *Synlett* 142–144 (1995).
12. Hayes, J. F., Shipman, M., Slawin, A. M. Z. & Twin, H. Multi-component reactions of 2-isopropylidene-aziridines: Application to the synthesis of enantiopure neopentylamines. *Heterocycles* **58**, 243–250

- (2002).
13. Nugent, T. C., Ghosh, A. K., Wakchaure, V. N. & Mohanty, R. R. Asymmetric reductive amination: Convenient access to enantioenriched alkyl-alkyl or aryl-alkyl substituted α -chiral primary amines. *Adv. Synth. Catal.* **348**, 1289–1299 (2006).
 14. Cogan, D. A., Liu, G. & Ellman, J. Asymmetric synthesis of chiral amines by highly diastereoselective 1,2-additions of organometallic reagents to N-tert-butanesulfinyl imines. *Tetrahedron* **55**, 8883–8904 (1999).
 15. Alder, C. M. *et al.* Updating and further expanding GSK's solvent sustainability guide. *Green Chem.* **18**, 3879–3890 (2016).
 16. Ghosh, T., Ernst, M., Hashmi, A. S. K. & Schaub, T. Ruthenium Catalyzed Direct Asymmetric Reductive Amination of Simple Aliphatic Ketones Using Ammonium Iodide and Hydrogen. *Eur. J. Org. Chem.* **2020**, 4796–4800 (2020).
 17. Goswami, A., Guo, Z., Parker, W. L. & Patel, R. N. Enzymatic resolution of sec-butylamine. *Tetrahedron: Asymmetry* **16**, 1715–1719 (2005).
 18. Poulhès, F., Vanthuyne, N., Bertrand, M. P., Gastaldi, S. & Gil, G. Chemoenzymatic dynamic kinetic resolution of primary amines catalyzed by CAL-B at 38–40°C. *J. Org. Chem.* **76**, 7281–7286 (2011).
 19. Hanson, R. L. *et al.* Preparation of (R)-amines from racemic amines with an (S)-amine transaminase from *Bacillus megaterium*. *Adv. Synth. Catal.* **350**, 1367–1375 (2008).
 20. Malik, M. S., Park, E.-S. & Shin, J.-S. ω -Transaminase-catalyzed kinetic resolution of chiral amines using l-threonine as an amino acceptor precursor. *Green Chem.* **14**, 2137 (2012).
 21. Yoon, S. *et al.* Deracemization of Racemic Amines to Enantiopure (R)- and (S)-amines by Biocatalytic Cascade Employing ω -Transaminase and Amine Dehydrogenase. *ChemCatChem* **11**, 1898–1902 (2019).
 22. Yun, H., Cho, B. K. & Kim, B. G. Kinetic resolution of (R,S)-sec-butylamine using omega-transaminase from *Vibrio fluvialis* JS17 under reduced pressure. *Biotechnol. Bioeng.* **87**, 772–778 (2004).
 23. Koszelewski, D., Clay, D., Rozzell, D. & Kroutil, W. Deracemisation of α -chiral primary amines by a one-pot, two-step cascade reaction catalysed by ω -transaminases. *Eur. J. Org. Chem.* 2289–2292 (2009) doi:10.1002/ejoc.200801265.
 24. Mutti, F. G., Fuchs, C. S., Pressnitz, D., Sattler, J. H. & Kroutil, W. Stereoselectivity of four (R)-selective transaminases for the asymmetric amination of ketones. *Adv. Synth. Catal.* **353**, 3227–3233 (2011).
 25. Mutti, F. G. *et al.* Amination of ketones by employing two new (S)-selective ω -transaminases and the his-tagged ω -TA from *Vibrio fluvialis*. *Eur. J. Org. Chem.* 1003–1007 (2012) doi:10.1002/ejoc.201101476.
 26. Koszelewski, D., Lavandera, I., Clay, D., Rozzell, D. & Kroutil, W. Asymmetric synthesis of optically pure pharmacologically relevant amines employing ω -transaminases. *Adv. Synth. Catal.* **350**, 2761–2766 (2008).
 27. Cerioli, L., Planchestainer, M., Cassidy, J., Tessaro, D. & Paradisi, F. Characterization of a novel amine transaminase from *Halomonas elongata*. *J. Mol. Catal. B: Enzym.* **120**, 141–150 (2015).
 28. 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.* **41**, 628–637 (2007).
 29. Heckmann, C. M., Gourlay, L. J., Dominguez, B. & Paradisi, F. An (R)-Selective Transaminase From *Thermomyces stellatus*: Stabilizing the Tetrameric Form. *Front. Bioeng. Biotechnol.* **8**, 707 (2020).
 30. Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. Rational assignment of key motifs for function guides in silico enzyme identification. *Nat. Chem. Biol.* **6**, 807–813 (2010).
 31. Theodorou, D., Zannikou, Y. & Zannikos, F. Estimation of the standard uncertainty of a calibration curve: Application to sulfur mass concentration determination in fuels. *Accredit. Qual. Assur.* **17**, 275–281 (2012).

32. *Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement*. (2012). doi:0 948926 15 5.
33. Guidi, B. *et al.* Strategic single point mutation yields a solvent- and salt-stable transaminase from *Virgibacillus* sp. in soluble form. *Sci. Rep.* **8**, 16441 (2018).
34. Pavlidis, I. V. *et al.* Identification of (S)-selective transaminases for the asymmetric synthesis of bulky chiral amines. *Nat. Chem.* **8**, 1076–1082 (2016).
35. Guo, F. & Berglund, P. Transaminase biocatalysis: optimization and application. *Green Chem.* **19**, 333–360 (2017).
36. Planchestainer, M., Hegarty, E., Heckmann, C. M., Gourlay, L. J. & Paradisi, F. Widely applicable background depletion step enables transaminase evolution through solid-phase screening. *Chem. Sci.* **10**, 5952–5958 (2019).
37. Trott, O. & Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **31**, 455–461 (2009).
38. Schätzle, S., Höhne, M., Redestad, E., Robins, K. & Bornscheuer, U. T. Rapid and sensitive kinetic assay for characterization of ω -transaminases. *Anal. Chem.* **81**, 8244–8248 (2009).
39. Planchestainer, M. *et al.* Continuous flow biocatalysis: production and in-line purification of amines by immobilised transaminase from *Halomonas elongata*. *Green Chem.* **19**, 372–375 (2017).
40. Publicover, E. A., Kolwich, J., Stack, D. L., Doué, A. J. & Ylijoki, K. E. O. Crystal structure of (S)-sec-butylammonium l-tartrate monohydrate. *Acta Crystallogr. Sect. E Crystallogr. Commun.* **73**, 716–719 (2017).
41. Dall’Oglio, F. *et al.* Flow-based stereoselective reduction of ketones using an immobilized ketoreductase/glucose dehydrogenase mixed bed system. *Catal. Commun.* **93**, 29–32 (2017).
42. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angew. Chemie Int. Ed.* **44**, 4192–4196 (2005).
43. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**, 135–145 (2018).
44. Simonetti, F. L., Teppa, E., Chernomoretz, A., Nielsen, M. & Marino Buslje, C. MISTIC: Mutual information server to infer coevolution. *Nucleic Acids Res.* **41**, 8–14 (2013).
45. Acevedo-Rocha, C. G., Reetz, M. T. & Nov, Y. Economical analysis of saturation mutagenesis experiments. *Sci. Rep.* **5**, 1–12 (2015).
46. 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.* **53**, 10714–10717 (2014).
47. Planchestainer, M. H. *elongata* aminotransferase: a resourceful enzyme for modern biocatalysis. (University of Nottingham, 2017).
48. Savile, C. K. *et al.* Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **329**, 305–310 (2010).

4 Further investigations of HEwT wild-type, F18W, and F84W

Given the successful implementation of HEwT_F84W in the synthesis of (*S*)-2-aminobutane, in particular regarding the enhancement of enantioselectivity from 45% *ee* to >99.5%, the effect of the small-pocket mutations on other small chiral amines was investigated. To this end, the pro-chiral ketones butenone, 4,4,4-trifluorobutanone (TFB-ketone), tetrahydrofuran-3-one (THF-ketone), and tetrahydrothiophen-3-one (THT-ketone) were chosen (**Scheme 4-1**). The first two were chosen because of their similarity to butanone; while the latter two were chosen as wt HEwT had shown modest and variable *ee*'s during the formation of the corresponding amine in previous work.¹⁻³ Curiously, while wild-type HEwT had been shown to preferentially produce the (*S*)-enantiomer of tetrahydrofuran-3-amine (THF-amine), it preferentially produced (*R*)-tetrahydrothiophen-3-amine (THT-amine).⁴



Scheme 4-1: Synthesis of small chiral amines using variants of HEwT (wt, F18W, F84W) from butenone, 4,4,4-trifluorobutanone (TFB-ketone), tetrahydrofuran-3-one (THF-ketone), or tetrahydrothiophen-3-one (THT-ketone).

4.1 Comparison of HEwT wild-type, F18W, and F84W variants

Reactions were set up containing 91 mM of the ketone substrate and one equivalent of SMBA, using lyophilized cfes of the different variants. Unfortunately, <1% conversion was observed with butenone or TFB-ketone (comparing peak areas of SMBA and acetophenone by GC-FID). Additionally, unidentified peaks appeared (three and one, respectively). As these peaks did not depend on the addition of enzyme, and coincided with a consumption of SMBA, these are likely the imine (in the case of TFB-ketone), and/or Michael addition product and oligomers (in the case of butenone). On the other hand, this was not observed in the biotransformations with THF- and THT-ketone which instead produced the expected amines, as can be seen in **Table 4-1**.

Table 4-1: Synthesis of 3-amino-THF and 3-amino-THT using variants of HEwT. Reactions on a 91 mM scale, containing HEwT (lyo. cfe, 2.7 mg/mL), SMBA (1 eq.), PLP (0.9 mM), KP_r-buffer (100 mM), and DMSO (1% (THF-ketone), or 11% (THT-ketone)); pH 8. Reactions were incubated at 30 °C for 24 h. THT-ketone was not fully soluble under these conditions.

Ketone	HEwT_wt		HEwT_F18W		HEwT_F84W	
	%Conversion ^a	ee (%)	%Conversion ^a	ee (%)	%Conversion ^a	ee (%)
THF-ketone ^b	88	14 (<i>S</i>)	52±1	34 (<i>S</i>)	14	48 (<i>S</i>)
THT-ketone ^c	53±2	37 (<i>R</i>)	18	43 (<i>R</i>)	4	45 (<i>S</i>)

^a Conversion determined by GC-FID using the relative peak areas of SMBA and acetophenone. Average of two duplicates, unless otherwise indicated SE < 1.

^b ee determined by chiral RP-HPLC following FMOC-derivatization.

^c ee determined by chiral GC-FID following acetylation.

The two mutant variants showed reduced conversions, in particular the F84W variant, but increased ee. In the case of THT-ketone, both the wild-type and the F18W variant are (*R*)-selective, however the F84W variant is (*S*)-selective. This is likely due to the increased steric bulk in the small pocket clashing with the large sulfur atom. Notably, the ee obtained for the wild-type was lower than previously obtained results in the lab (30-70 %).² Next, the intensification potential of all three variants was assessed, using 5 eq. of IPA as the amine donor and carrying out reactions on a 10, 100, and 300 mM scale. The results of this intensification study are shown in **Table 4-2**.

Table 4-2: Intensification of biotransformations of 3-amino-THF and 3-amino-THT using variants of HEwT. Reactions on a 10, 100 or 300 mM scale, containing HEwT (lyo. cfe, 50 mg/mmol), IPA (5 eq.), PLP (1 mM), KP_r-buffer (50 mM), and DMSO (10%); pH 8. Reactions were incubated at 30 °C for 48 h. THT-ketone was not fully soluble at the 100 and 300 mM scales.

Ketone	Scale (mM)	HEwT_wt		HEwT_F18W		HEwT_F84W	
		%Conversion	ee (%)	%Conversion	ee (%)	%Conversion	ee (%)
THF-ketone ^{ab}	10	82±3	11 (<i>S</i>)	81±1	32 (<i>S</i>)	5±2	n.d.
	100	76±2	4 (<i>R</i>)	61±2	20 (<i>S</i>)	20±2	50 (<i>S</i>)
	300	54±2	17 (<i>R</i>)	12±3	3 (<i>R</i>)	11±2	43 (<i>S</i>)
THT-ketone ^{bc}	10 ^d	58±3	23 (<i>R</i>)	50±3	28 (<i>R</i>)	6±3	31 (<i>S</i>)
	100	76±4	17 (<i>R</i>)	48±3	29 (<i>R</i>)	10±3	45 (<i>S</i>)
	300	78±3	20 (<i>R</i>)	27±3	32 (<i>R</i>)	11±3	50 (<i>S</i>)

^a Conversion determined by RP-HPLC, following the production of THF-amine (after FMOC derivatization), using a calibration curve.

^b ee determined by chiral GC-FID following acetylation.

^c Conversion determined by chiral RP-HPLC, following the production of THT-amine (after FMOC-derivatization), using a calibration curve.

^d ee determined by chiral RP-HPLC following FMOC-derivatization.

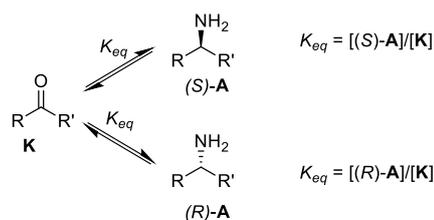
For the wild-type, conversions decreased at increasing concentrations of THF-ketone, but increased with increasing concentrations of THT-ketone. For the F18W variant, conversions decreased with increasing concentrations of either substrate, and for the F84W

variant conversions increased from 10 mM to 100 mM with either substrate, but were significantly lower than for the other variants, presumably due to steric clash with the methylene (C5) or heteroatom (O/S1), consistent with the enhanced discrimination of methyl vs ethyl in butanone as described in the previous chapter.

Unexpectedly, the wild-type switched from being (*S*)-selective to being (*R*)-selective at the higher concentrations for THF-ketone. This switching behaviour was also observed for the F18W variant, but the switch occurred at a higher substrate concentration (approx. 300 mM vs 100 mM). At each concentration, it was more (*S*)-selective (or less (*R*)-selective) than the wild-type. The F84W variant maintained its enantioselectivity throughout and was more (*S*)-selective than either of the other variants, yielding (*S*)-enriched THF-amine at all tested substrate concentrations. However, a slight decrease in (*S*)-selectivity at increasing substrate concentrations was also observed. For THF-ketone, no such concentration dependence was observed for the wild-type or F18W variant, with the latter showing slightly larger (*R*)-selectivity. However, for the F84W variant, the product was increasingly enriched with the (*S*)-enantiomer at increasing substrate concentrations.

4.2 Enantioselectivity THF-ketone—Kinetic and thermodynamic effects

Given that previous results obtained in the lab with wild-type HEwT and THF-ketone produced *ees* ranging from 15% (*S*)³ up to 70% (*S*),^{1,2} attempts were made to better understand this variability and switching behaviour, focussing on the wild-type. An initial hypothesis was postulated that different conditions in the experiments (catalyst loading, reaction temperature and time, immobilized or soluble enzyme, and substrate loading) could explain the differences in *ee* through thermodynamic vs kinetic control; in particular, longer reaction times or higher catalyst loadings (or a more stable catalyst) would favour a thermodynamic outcome, *i.e.* the racemate. This behaviour has been previously reported with chiral organocatalysts.⁵ While it may seem counterintuitive why the racemate should be favoured thermodynamically if each enantiomer has identical thermodynamic properties, it becomes apparent when considering the equilibrium between the ketone and each enantiomer (**Scheme 4-2**). Given the identical equilibrium constants for each of the two equilibria, the only way for both reactions to be at equilibrium is at equal concentrations of each enantiomer. Alternatively, since the difference in Gibbs free energy between the two enantiomers is zero, the equilibrium constant describing the equilibrium between the two is one.



Scheme 4-2: Simplified equilibrium for a generic reductive amination. Since both enantiomers have identical thermodynamic properties (standard Gibbs free energy), the equilibrium constants between the prochiral ketone (**K**) and either enantiomer of the amine (**A**) are identical. Since both equations have the same concentration of ketone in the denominator, the concentration of each enantiomer has to be identical for both equilibria to be satisfied.

To observe the thermodynamic effects, the *ee* of a reaction was monitored over time. Additionally, the *ee* of a series of reactions with increasing enzyme concentrations was also observed. Highest *ees* are expected at the shortest reaction times and with the lowest enzyme concentration and should decrease as either time or enzyme concentration increases. As can be seen in **Figure 4-1A-C**, this behaviour is indeed observed, however the maximum *ees* obtained under those conditions were approx. 16% (*S*), and thus it seemed unlikely that this effect was strong enough to explain the range of *ees* previously observed. Indeed, this effect would also not explain the concentration-dependent switching behaviour that is also observed with the purified enzyme (**Figure 4-1D**), and not only with the crude.

However, the presence of thermodynamic effects means that difference in *ee* between the variants as described in **Table 4-1** and **Table 4-2** should be taken with a degree of caution; in particular in the case of THT-ketone and the wild-type and F18W variants, where rather small differences in *ee* may in fact be due to increased thermodynamic effects with the wild-type (which reaches higher conversions).

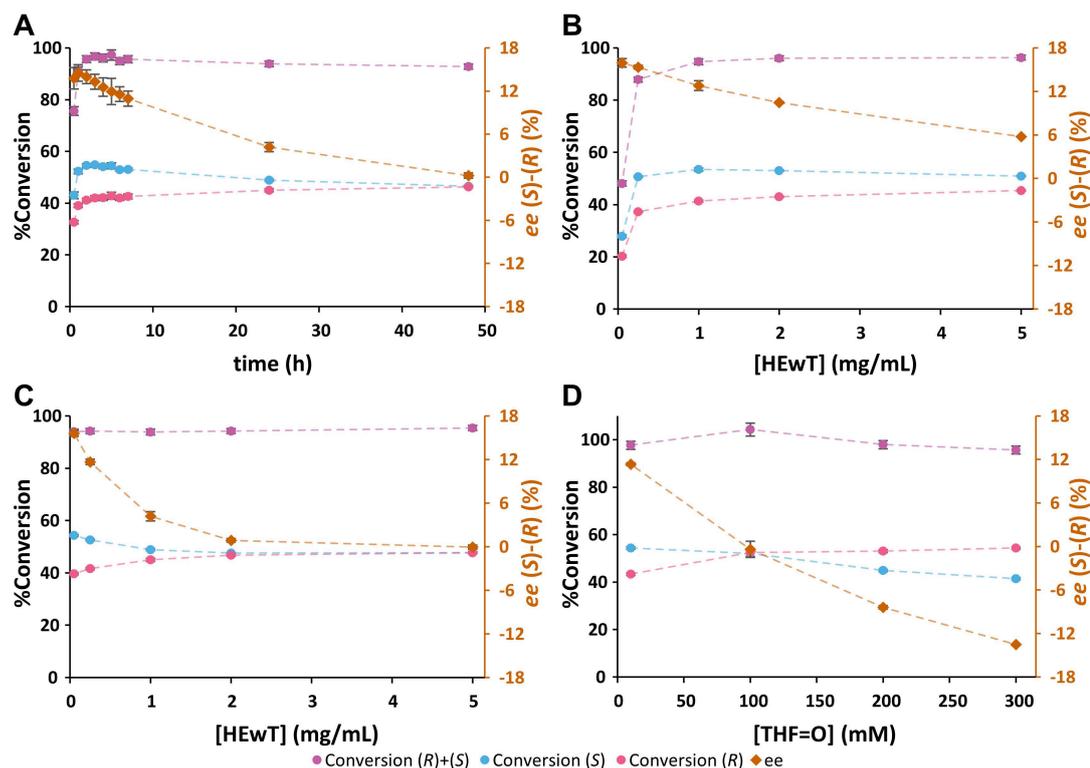


Figure 4-1: **A:** The production of (*S*)- and (*R*)-THF-amine from THF-ketone over time, in a biotransformation containing THF-ketone (10 mM), SMBA (1 eq.) and purified HEWT (1.0 mg/mL). **B:** The production of (*S*)- and (*R*)-THF-amine from THF-ketone with increasing concentrations of purified HEWT, in biotransformations containing THF-ketone (10 mM) and SMBA (1 eq.). Samples taken after 3 h and **C:** samples taken after 24 h. **D:** The production of (*S*)- and (*R*)-THF-amine from THF-ketone (THF=O) at increasing substrate concentrations, in biotransformations containing IPA (5 eq.), and purified HEWT (0.72 mg/mL). Samples taken after 3 h. All reactions contained PLP (1 mM), KP₁-buffer (100 mM); pH 8, 30 °C. Conversions determined by chiral RP-HPLC, after FMOC-derivatization. Error bars represent standard errors (n=2) and include the error associated with the calibration curve.^{6,7} Signed ees; +ve (*S*)-enantiomer, -ve (*R*)-enantiomer. Connecting lines added for clarity.

4.3 Enantioselectivity THF-ketone—Concentration effects

Attempting to gain some insight into this behaviour, biotransformations were set up containing fixed concentrations of either THF-ketone or IPA, to verify whether the behaviour was caused by an increase in either substrate, or both substrates. As can be seen in **Figure 4-2A**, increasing the concentration of THF-ketone from 10 to 300 mM while fixing the concentration of IPA to 50 mM, the enzyme remains (*S*)-selective, with slight variation probably due to thermodynamic effects (higher quantities of product correlating with smaller ees). However, when THF-ketone was fixed to 100 mM and IPA was varied from 50 to 1500 mM, the same switching behaviour was observed as when both were varied (compare **Figure 4-1D** and **Figure 4-2B**).

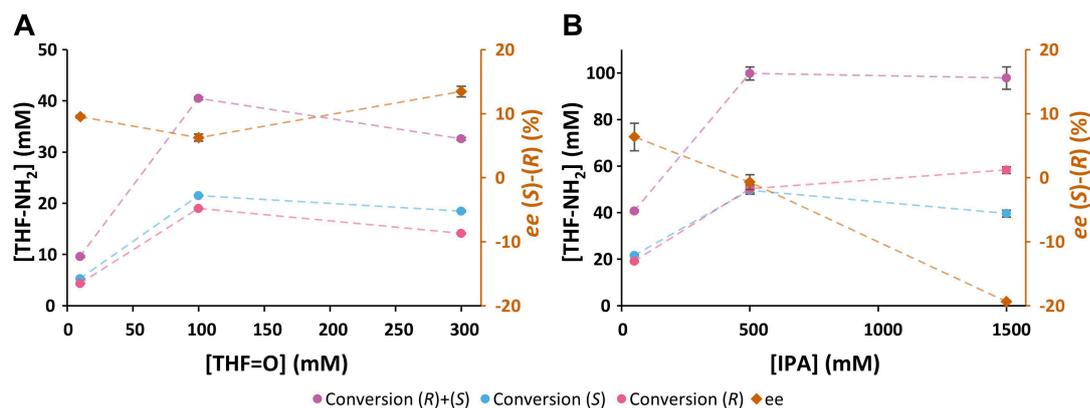


Figure 4-2: **A:** The production of (*S*)- and (*R*)-THF-amine from THF-ketone (THF=O) at increasing concentration of THF-ketone, in biotransformations containing a fixed concentration of IPA (50 mM), and purified HEwT (1 mg/mL). Samples taken after 3 h. **B:** The production of (*S*)- and (*R*)-THF-amine from THF-ketone (THF=O) at increasing concentration of IPA, in biotransformations containing a fixed concentration of THF-ketone (100 mM), and purified HEwT (1 mg/mL). Samples taken after 3 h. All reactions contained PLP (1 mM), KP_i-buffer (100 mM); pH 8, 30 °C. Conversions determined by chiral RP-HPLC, after FMOC-derivatization. Error bars represent standard errors (n=2) and include the error associated with the calibration curve.^{6,7} Signed *ees*; +ve (*S*)-enantiomer, -ve (*R*)-enantiomer. Connecting lines added for clarity.

Previous docking studies in the group² suggested that the enantioselectivity of HEwT could be explained via hydrogen bonding of the ring oxygen to W56 as the substrate enters the active site (**Figure 4-3**). From this docking, it appears that the neighbouring methylene (C5) is facing the entrance of the active site and could possibly interact with the solvent. Thus, an initial hypothesis was formed that competing hydrogen bonding of THF-ketone to the reaction medium (*i.e.* with IPA, which at pH 8 is protonated (pK_a 10.63)⁸) might favour the substrate entering the active site “C5-first” rather than “oxygen-first,” favouring the production of (*R*)-THF-amine. However, other effects on the enzyme are also possible, *e.g.* the increasing organic content (decreasing water content) or ionic strength might result in structural changes of the enzyme that favour the production of the (*R*)-enantiomer.

To separate the different effects of increasing IPA concentrations (*i.e.* hydrogen bonding, organic content, and ionic strength), experiments were set up with varying concentrations of the following additives: isopropyl alcohol (a weaker hydrogen bond donor than (protonated) IPA, with similar size and structure), sodium chloride (increasing the ionic strength without altering the concentration of hydrogen bond donors or the organic content), and ammonium chloride (which, despite its lower pK_a of 9.21, is also protonated at pH 8 and should provide a similar hydrogen bond donor capacity to IPA, whilst not increasing the organic content).

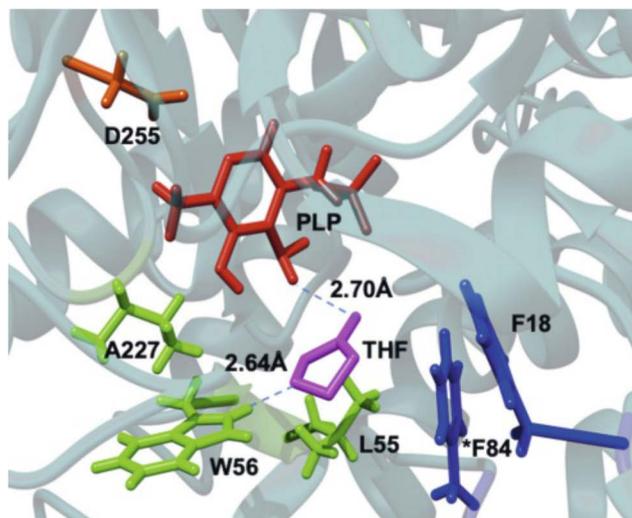


Figure 4-3: Docking of THF-ketone into the entrance of the active site of wild-type HEwT, showing the hydrogen bond to W56. Figure reproduced from ref.² (CC BY 4.0 license).

Thus, if the switch in enantioselectivity is indeed due to competing hydrogen bonding to the reaction medium, a similar effect should be observed for ammonium chloride and to a lesser extent isopropyl alcohol. However, if it is due to increasing organic content changing the structure of the enzyme, it should be observed only with isopropyl alcohol; and if it is due to increasing ionic strength it should be observed only with sodium chloride and ammonium chloride. As can be seen in **Figure 4-4A-C**, increasing the isopropyl alcohol concentration does indeed affect the *ee*, but in the opposite direction to what was expected. On the other hand, both increasing concentrations of sodium chloride and ammonium chloride switch the enantioselectivity from (*S*) to (*R*).

This suggests that the switch in enantioselectivity is caused by structural changes of the enzyme, with a more (*S*)-selective structure dominating in hydrophobic/organic reaction media and a more (*R*)-selective structure in hydrophilic/ionic reaction media. Increasing the pH (**Figure 4-4D**) also enhances the (*S*)-selectivity of HEwT, which may be due to lower amounts of protonated amines, rendering the reaction medium more hydrophobic/lowering the ionic strength. However, in the absence of direct evidence of such structural changes (*e.g.* cryo-electron microscopy or circular dichroism spectroscopy), this is not a definitive conclusion and it might indeed be the case that a more hydrophobic environment favours THF-ketone entering the active site “oxygen-first” and a more hydrophilic environment favours it entering “C5-first.” Increasing ionic strength might also disrupt the hydrogen bonding to W56 or alter the complexation of THF-ketone in solution.

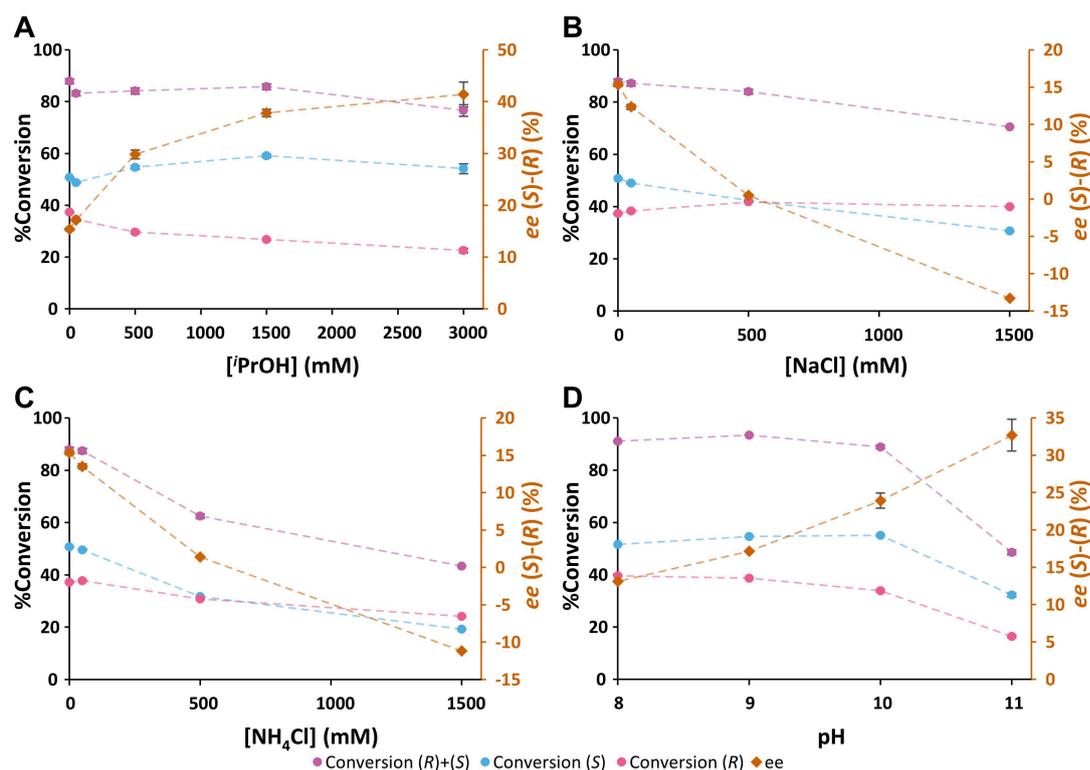


Figure 4-4: A-C: The production of (S)- and (R)-THF-amine from THF-ketone (THF=O) in biotransformations containing THF-ketone (10 mM), SMBA (1 eq.), and purified HEwT (0.25 mg/mL), with varying concentrations of additives. Samples taken after 3 h. **A:** *i*PrOH (50-3000 mM), **B:** NaCl (50-1500 mM), **C:** NH₄Cl (50-1500 mM). In the case of **C**, the decrease of conversion at increasing concentrations of ammonium chloride might be due to interference of ammonium in the Fmoc derivatization, *i.e.* an artefact of the analysis rather than a real reduction in conversion. This is supported by the complete consumption of Fmoc-Cl, the appearance of an unidentified peak (presumably Fmoc-NH₃), and a reduction in size of the Fmoc-derivatized SMBA peak. **D:** The production of (S)- and (R)-THF-amine from THF-ketone (THF=O) in biotransformations containing THF-ketone (50 mM), SMBA (1 eq.), and purified HEwT (0.25 mg/mL), with varying pH. Samples taken after 3 h. All reactions contained PLP (1 mM), KP₇-buffer (100 mM); pH 8, 30 °C. Conversions determined by chiral RP-HPLC, after Fmoc-derivatization. Error bars represent standard errors (n=2) and include the error associated with the calibration curve.^{6,7} Signed ees; +ve (S)-enantiomer, -ve (R)-enantiomer. Connecting lines added for clarity.

4.4 Conclusions

Attempting to expand the substrate scope of the small-pocket mutants from the previous chapter, a curious switching in enantioselectivity was observed for THF-ketone. Initially observed at increasing substrate loading, this was then shown to be caused by an increase in ionic strength, with increasing organic content having the opposite effect. This is most likely due to structural changes in the enzyme under these varying conditions. However, if the enantioselectivity of HEwT is indeed determined by the orientation in which the substrate enters the active site, a direct effect of the reaction medium on that orientation could not be fully ruled out. Clearly, further studies are needed, and may include obtaining structures of the enzyme under varying ionic strengths/organic content, circular dichroism spectroscopy, and point mutations of the W56 residue to confirm its role in the enantioselectivity of HEwT for that substrate.

4.5 References

1. Planchestainer, M., Hegarty, E., Heckmann, C. M., Gourlay, L. J. & Paradisi, F. Widely applicable background depletion step enables transaminase evolution through solid-phase screening. *Chem. Sci.* **10**, 5952–5958 (2019).
2. Hegarty, E. & Paradisi, F. Implementation of biocatalysis in continuous flow for the synthesis of small cyclic amines. *Chimia* **74**, 890–894 (2020).
3. Keil, A. A sequential chemo-enzymatic transamination—Buchwald-Hartwig cross coupling cascade. (University of Nottingham, 2019).
4. Heckmann, C. M., Gourlay, L. J., Dominguez, B. & Paradisi, F. An (R)-Selective Transaminase From *Thermomyces stellatus*: Stabilizing the Tetrameric Form. *Front. Bioeng. Biotechnol.* **8**, 707 (2020).
5. Rulli, G. *et al.* Direction of kinetically versus thermodynamically controlled organocatalysis and its application in chemoenzymatic synthesis. *Angew. Chemie Int. Ed.* **50**, 7944–7947 (2011).
6. Theodorou, D., Zannikou, Y. & Zannikos, F. Estimation of the standard uncertainty of a calibration curve: Application to sulfur mass concentration determination in fuels. *Accredit. Qual. Assur.* **17**, 275–281 (2012).
7. *Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement.* (2012). doi:0 948926 15 5.
8. Hall, H. K. Correlation of the Base Strengths of Amines. *J. Am. Chem. Soc.* **86**, 5709 (1964).

5 An (*R*)-Selective Transaminase from *Thermomyces stellatus*: Stabilizing the Tetrameric Form

5.1 Background

As described in **Section 1.2.2.1**, ATAs belong to two distinct fold-types of PLP-dependent enzymes: STAs belong to fold type I, and RTAs belong to fold-type IV. RTAs share their fold type with branched chain amino transferases (BCATs) and D-amino acid transaminases (DATAs), which are significantly more abundant in nature and well characterized, while RTAs remained elusive prior to Höhne *et al.*¹ describing a consensus motif, distinguishing them from those other members. Since then, several more RTAs have been reported, with just a few crystal structures solved. However, the overall number of known RTAs (and their solved structures) is still significantly smaller compared to STAs.^{2,3} Thus, the recruitment of additional RTAs was deemed of interest to aid in the projects described in **Chapter 3**.

5.2 Search for RTA sequences

The application of TAs (as well as other classes of enzymes) in chemical synthesis is often hindered by their insufficient stability.⁴ Recruitment of biocatalysts from extremophilic organisms⁵ has proved to be a valuable strategy to overcome this limitation, and thus only extremophilic organisms were considered as a source in this project. Two sequences predicted to be RTAs were identified from a protein BLAST search: one from the halophilic bacterium *Halofilum ochraceum*⁶ (*HoRTA*) and one from the thermotolerant fungus *Thermomyces stellatus* (*TsRTA*), having 43% and 81% identity with the RTA from *Aspergillus terreus* (*AtRTA*),⁷ respectively (**Figure 5-1**). Initial cloning of *HoRTA* into the expression vector pMP89b (modified pRSET-B)⁸ resulted in no overexpressed protein being observed under any expression conditions trialed (LB, TB, and ZYP-AI at 37 °C). Cloning into the expression vector pCH93b (modified pET22b),⁸ with a stop codon preceding the C-terminal His-tag encoded by the vector, overexpressed protein was observed, yet insoluble. Lowering the expression temperature to 16 °C, soluble protein was obtained but no activity was observed (**Figure 5-2a**). The function, if any, of the crude enzyme could not be confirmed. Mutagenesis to remove the stop codon and include the His-tag in the expressed protein was also attempted, but again the protein was only obtained in insoluble form. Thus, *HoRTA* was not pursued any further. Following these results, *TsRTA* was directly cloned into pCH93b,

including the His-tag. During initial expression trials at 37 °C, active, soluble protein could be obtained. However, a large fraction of inclusion bodies was also observed.

```
#
#
# Percent Identity Matrix - created by Clustal2.1
#
#
```

1: GEO1900	100.00	78.62	29.31	28.52	30.21	30.93	28.18	28.62	30.34	27.93
2: AF0933	78.62	100.00	32.29	31.49	33.45	31.83	29.76	30.21	30.90	29.86
3: ATA-117	29.31	32.29	100.00	91.82	37.92	40.00	43.12	43.75	42.01	42.99
4: ATA-117-Rd11	28.52	31.49	91.82	100.00	38.13	37.38	39.25	41.25	40.00	40.37
5: HoRTA	30.21	33.45	37.92	38.13	100.00	40.86	43.51	44.97	45.78	42.86
6: Aspergillus-fumigatus	30.93	31.83	40.00	37.38	40.86	100.00	68.42	69.06	72.98	72.36
7: Exophiala-xenobiotica	28.18	29.76	43.12	39.25	43.51	68.42	100.00	73.44	70.06	73.68
8: Nectria-haematococca	28.62	30.21	43.75	41.25	44.97	69.06	73.44	100.00	78.06	77.43
9: TsRTA	30.34	30.90	42.01	40.00	45.78	72.98	70.06	78.06	100.00	81.42
10: AtRTA	27.93	29.86	42.99	40.37	42.86	72.36	73.68	77.43	81.42	100.00

Figure 5-1: Identity matrix of TsRTA and HoRTA with a variety of previously reported RTAs and two thermostable BCATs (GEO1900 and AF0933), created using MUSCLE.⁹

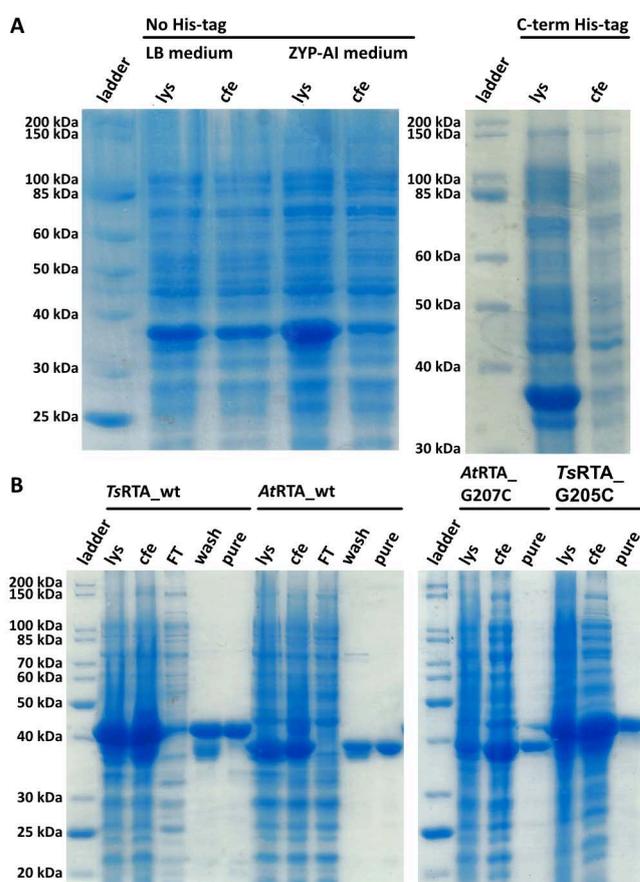


Figure 5-2: SDS-PAGE gels of **A)** HoRTA, expressed at 16 °C, both without a His-tag and with a C-terminal poly-His-tag. **B)** TsRTA, AtRTA, AtRTA_G207C, TsRTA_G205C. lys: lysate, cfe: cell-free extract, FT: flow-through (IMAC purification), wash: eluted with 10% elution buffer, pure: eluted with 100% elution buffer.

5.3 Expression and initial characterization of TsRTA

TsRTA was fully soluble when expressed at 25 °C, and, following Ni-IMAC, the enzyme was obtained in high yields (800 mg_{enzyme}/L_{culture}) and was judged to be pure by SDS-PAGE (**Figure 5-2b**). The specific activity with pyruvate and RMBA was determined to be 2.5 U/mg (AtRTA: 3 U/mg). The catalytic efficiency (k_{cat}/K_m) was comparable to AtRTA for RMBA and

ca. 2-fold lower for pyruvate (**Table 5-1**). Substrate inhibition was observed in both *Ts*RTA and *At*RTA for both pyruvate and RMBA, with higher inhibition from the latter (**Figure 5-3**).

Table 5-1: Kinetic parameters for wild-type (wt) and mutant *Ts*RTA and *At*RTA.

		K_m (mM)	k_{cat} (s ⁻¹)	K_i (mM)	k_{cat} / K_m (s ⁻¹ mM ⁻¹)
<i>Ts</i>RTA_wt	RMBA	0.13 ± 0.01	1.82 ± 0.03	15 ± 1	13.8 ± 0.6
	Pyruvate	0.57 ± 0.03	2.21 ± 0.04	39 ± 5	3.9 ± 0.1
<i>Ts</i>RTA_G205C	RMBA	0.12 ± 0.01	1.35 ± 0.03	20 ± 2	11.0 ± 0.6
	Pyruvate	0.42 ± 0.01	1.33 ± 0.02	130 ± 30	3.2 ± 0.1
<i>At</i>RTA_wt	RMBA	0.17 ± 0.01	2.24 ± 0.04	19 ± 1	13.3 ± 0.5
	Pyruvate	0.23 ± 0.01	2.25 ± 0.02	52 ± 4	10.0 ± 0.2
<i>At</i>RTA_G207C	RMBA	0.30 ± 0.02	1.96 ± 0.05	19 ± 2	6.4 ± 0.3
	Pyruvate	0.19 ± 0.01	1.71 ± 0.03	60 ± 10	9.2 ± 0.5

Parameters were obtained by fitting a substrate inhibition curve (using GraphPad Prism) to the data obtained when either RMBA or pyruvate concentrations were varied. Reaction velocities at each concentration were measured in triplicate. Standard errors (SE) are quoted, accounting for covariance.

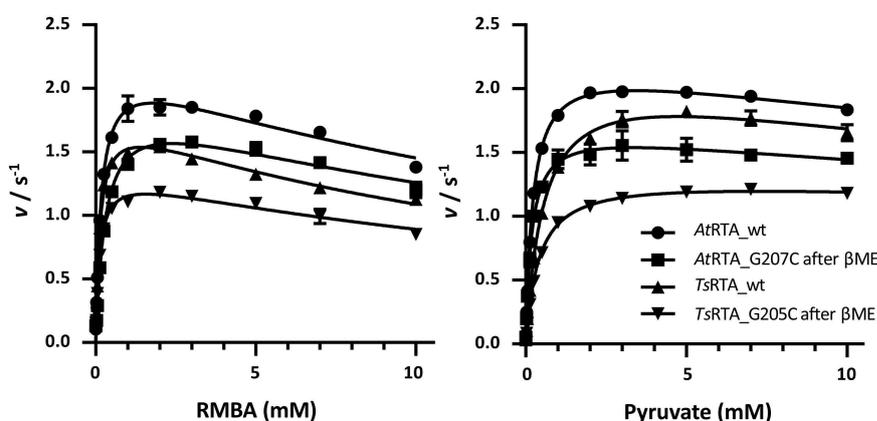


Figure 5-3: Substrate inhibition curves fitted to the reaction velocities obtained when RMBA and pyruvate were varied, respectively. Reaction velocities at each concentration were measured in triplicate.

The resting stability and activity of *Ts*RTA were then investigated under varying conditions. *Ts*RTA was stable in universal buffer between pH 5-9 for at least 14 days at 4 °C (**Figure 5-4a**), and most active between pH 8-9, as is commonly observed for RTAs¹⁰ (**Figure 5-4b**; it should be noted the overall specific activity in universal buffer¹¹ was lowered by ca. 20-fold). *Ts*RTA was stable in the presence of co-solvents such as 20% (v/v) methanol, ethanol, and DMSO, which are commonly used co-solvents in biotransformations, with no loss of activity after 1 week (**Figure 5-4d**). Stability decreased with increasing chain length of

alcohols, as well as THF, acetonitrile and to a lesser extent DMF. The activity in the presence of co-solvents followed the same trends as the stability (Figure 5-4c).

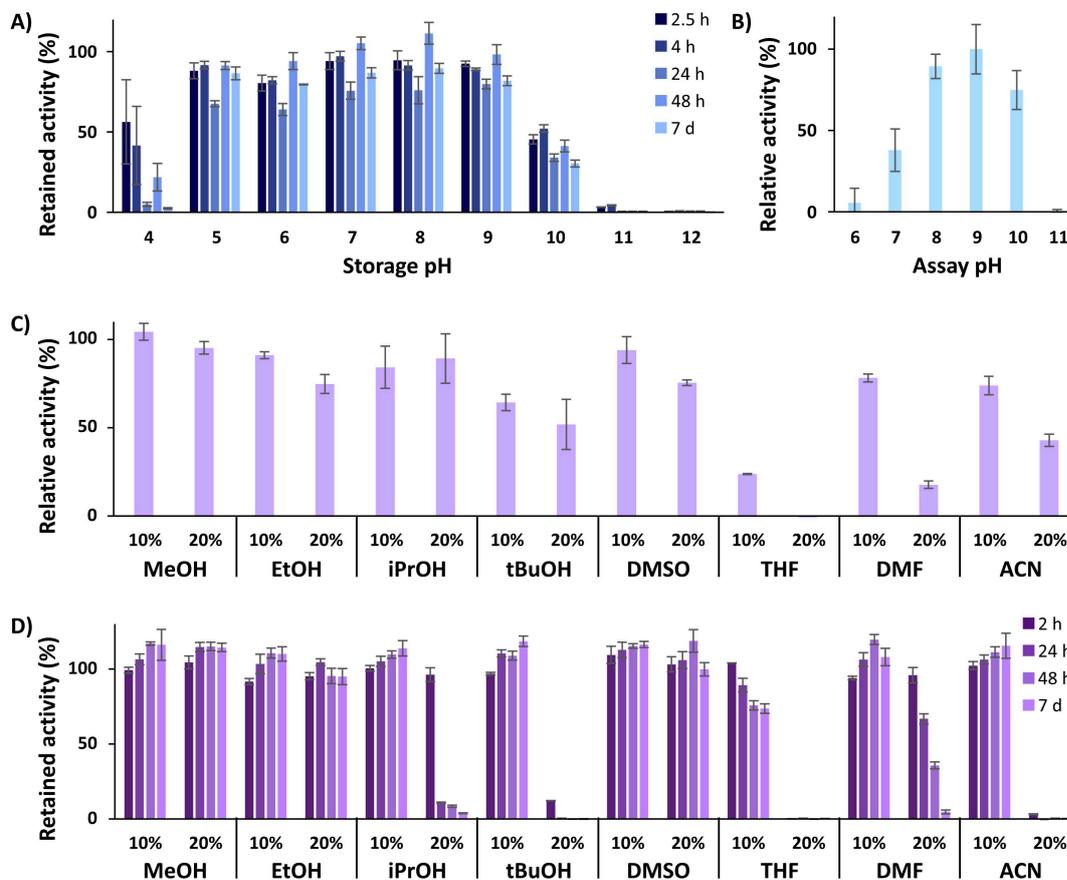


Figure 5-4: **A)** pH stability profile of *Ts*RTA: retained activity after incubation at pH 4 – 12 for 2.5 h – 7 d (4 °C), relative to freshly purified enzyme. **B)** pH–activity relationship of *Ts*RTA: relative activity at pH 6 – 11, relative to maximum activity at pH 9. **C)** Co-solvent–activity relationship of *Ts*RTA: relative activity in the presence of 10 or 20 % (*v/v*) co-solvents, relative to activity without co-solvent. **D)** Co-solvent stability profile of *Ts*RTA: retained activity after incubation with 10 or 20 % (*v/v*) co-solvents for 2 h – 7 d (25 °C, pH 8), relative to freshly purified enzyme. Error bars represent standard errors (n=3).

To assess the extent to which the thermotolerant origin of *Ts*RTA impacts its stability, thermostability assays were carried out in parallel with both *Ts*RTA and *At*RTA. *Ts*RTA retained 40% activity when incubated at 40 °C for 7 days, whereas *At*RTA lost almost 90% activity within 24 h. However, at 45 °C, *Ts*RTA almost completely lost its activity within 2h, while *At*RTA retained ca. 25% activity after 4h (Figure 5-5). Indeed, during temperature activity assays *At*RTA showed a higher optimum temperature compared to *Ts*RTA (Figure 5-6), which showed loss of activity over the 10 min activity assay from 50 °C onward.

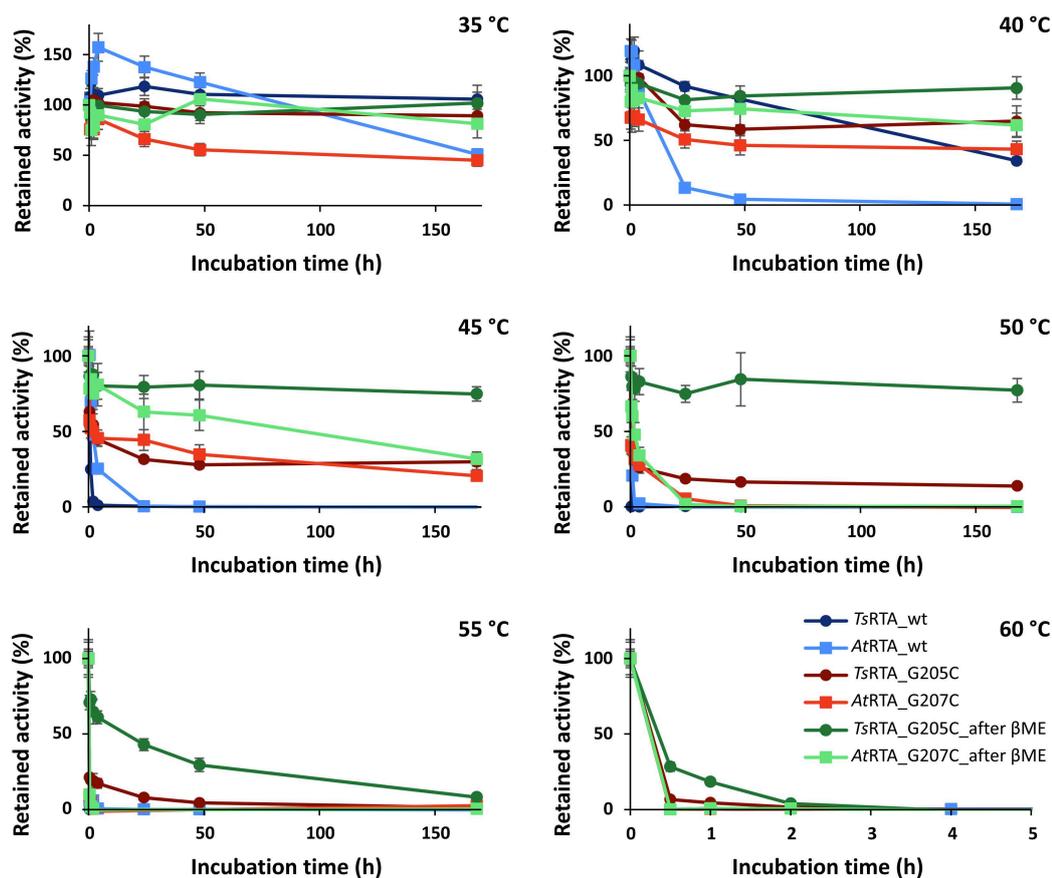


Figure 5-5: Temperature stability profiles of wild-type *Ts*RTA and *At*RTA, as well as mutants both before and after equilibration in the presence of β -mercaptoethanol. Retained activity after incubation at 35 – 60 °C for 30 min – 7 d (pH 8). Activity expressed relative to the activity of each enzyme at $t=0$. Error bars represent standard errors ($n=3$).

SEC unexpectedly revealed that 90% of *Ts*RTA in solution exists as a tetramer, with only 10% adopting the dimeric form (**Figure 5-7**). Other highly similar RTAs, including *At*RTA (4ce5, 81% identity),⁷ and RTAs from *Nectria haematococca* (4cmd, 78% identity),¹² *Exophiala xenobiotica* (6fte, 70% identity),¹³ and *Aspergillus fumigatus* (4uug, 73% identity)¹⁴ are reported as dimers based on their crystal packing. Two thermostable BCATs, belonging to the same fold type IV as RTAs, with hexameric structures have recently been reported by Isupov *et al.*¹⁵ In addition, increased operational stability of tetrameric STAs compared to dimeric STAs has been reported.¹⁶ Thus, the improved thermostability of *Ts*RTA was initially attributed to its tetrameric nature.⁵

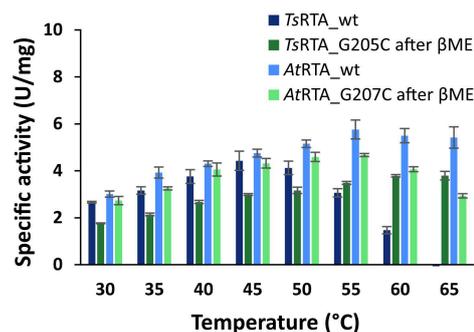


Figure 5-6: Temperature–activity relationship of wild-type and mutant TsRTA and AtRTA: specific activity at 30 – 65 °C. Error bars represent standard errors (n=3).

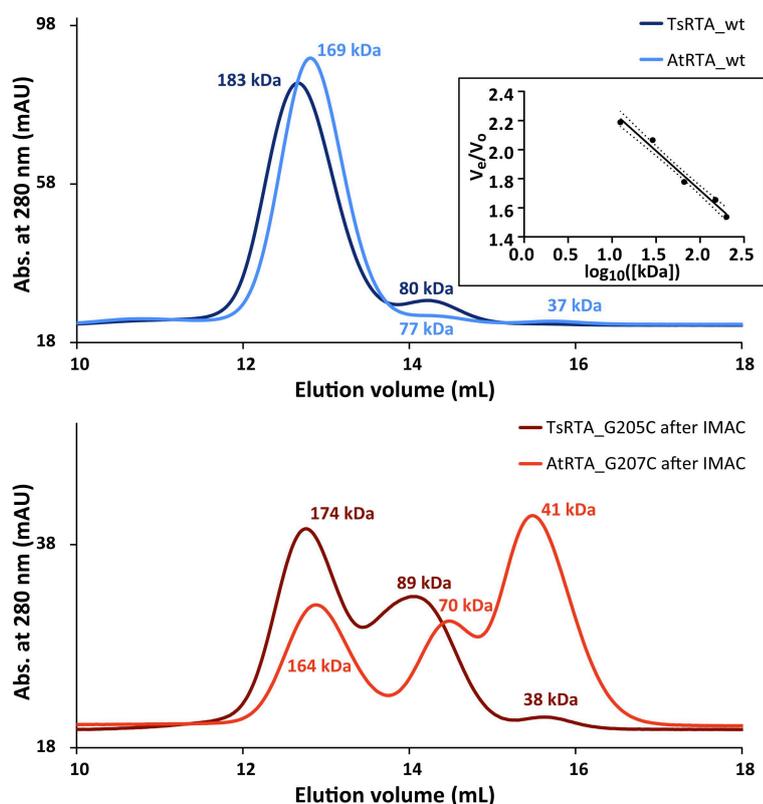


Figure 5-7: Gel filtration of TsRTA_G205C and AtRTA_G207C immediately after IMAC purification in contrast to the wild-type enzymes, showing the initially disrupted quaternary structure of the mutants. Insert: calibration curve using the Sigma Aldrich Gel Filtration Markers Kit for Protein Molecular Weights 12,000-200,000 Da (MWGF200).

5.4 3D Structural analysis and relation to thermostability

TsRTA was successfully crystallized by our collaborator Dr. Louise Gourley (University of Milan). Electron density was well-defined for residues 1 to 319 in both polypeptide chains (A and B) present in the asymmetric unit, but it was absent for TsRTA residues 320 to 334, in addition to the C-terminal His-tag, due to flexibility in this region. Both polypeptide chains exhibit high structural similarity, with an RMSD of 0.24 Å for 319 aligned C α atoms, as calculated using the CCP4i program SUPERPOSE.¹⁷ Regions of poor electron density were

observed for several side chains from residues 126 to 132 that are located in a loop region and solvent exposed. There was an area of positive density that appeared to be continuous with the sidechain of D65 (Chain A), observed in the difference map (mFo-DFc), that could not be identified or modelled.

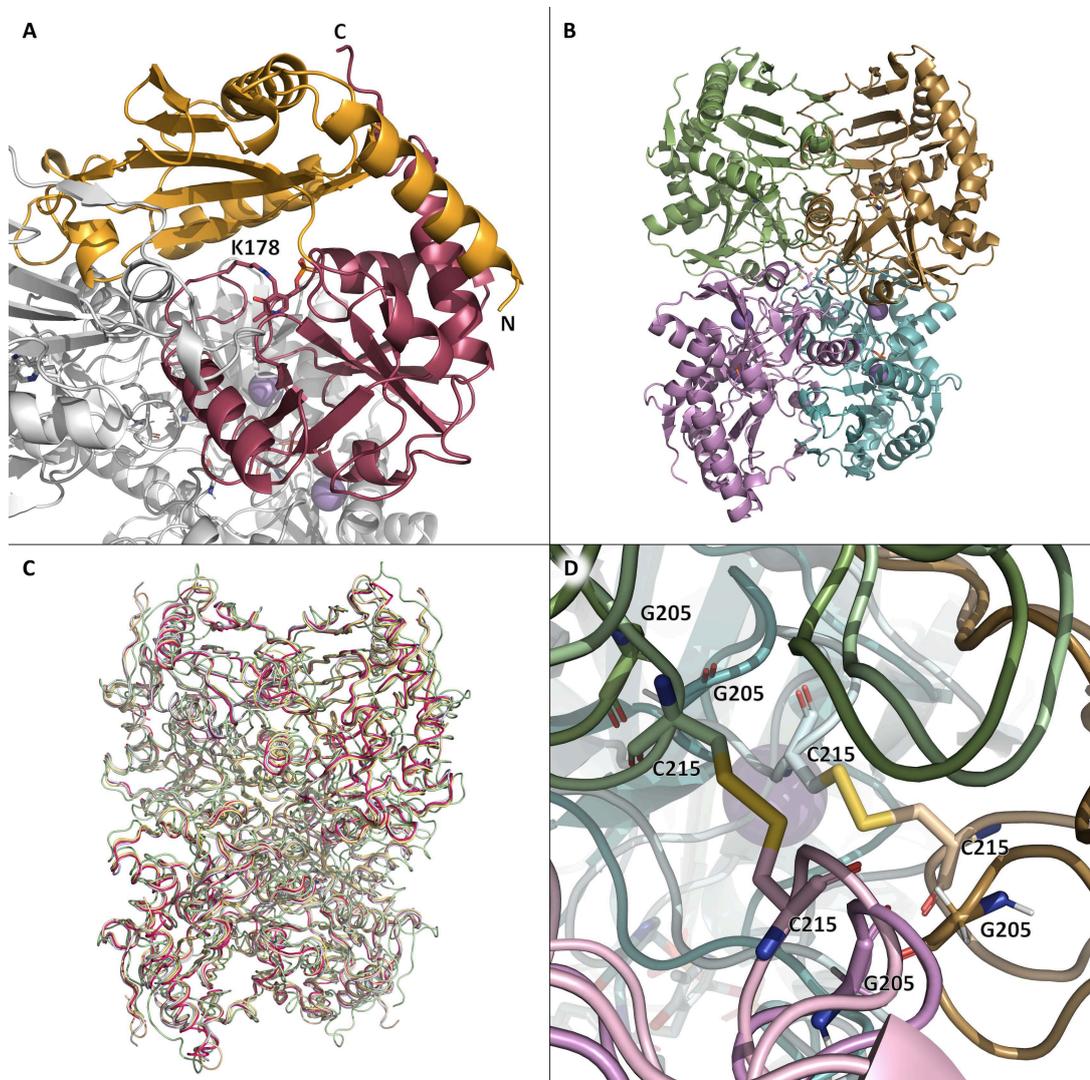


Figure 5-8: **A)** The monomer (chain B) of TsRTA in the context of the overall quaternary structure. The two sub-domains are coloured orange and pink. The co-factor PLP (attached to K178) is shown as sticks. **B)** the tetramer of TsRTA obtained by applying the symmetry operation $(-x, y, -z)$ to the asymmetric unit dimer. **C)** superimposition of the crystallographic tetramers of TsRTA (pink), AtRTA (4ce5, wheat), ATA-117-Rd11 (pale green), *N. haematococca* RTA (4cmd, blue-white), and *A. fumigatus* RTA (4uug, pale yellow). **D)** zoom view of the two disulfide bridges (C215) that extend across the dimer-dimer interface of 5fr9 (lighter colours) and the equivalent glycine residues (G205) of TsRTA (darker colours). This figure was generated with open source PyMOL 2.1.0.

The TsRTA monomer adopts the classical aminotransferase class IV fold (InterPro: IPR001544; Pfam PF01063) with a significant content of both α -helices and β -strands (3 β -sheets) and can be observed to be sub-divided into two smaller sub-domains (see **Figure 5-8a**). A Profunc search (<http://www.ebi.ac.uk/thornton->

srv/databases/profunc/),¹⁸ as expected, identified *At*RTA as the top structural homolog (RMSD of 0.67 Å over 318 C-alpha matched pairs).

The active site pocket houses a single PLP molecule, covalently bound to each *Ts*RTA chain via an imino bond with the conserved active site lysine residue (K178) and additional hydrogen bonds (2.6 – 3.0 Å) formed between the phosphate group of PLP and side chain atoms from residues R77, E211, T237 and T273 (conserved in *At*RTA); the phosphate group is further stabilized by hydrogen bonds with R77 and N219 via two conserved water molecules (**Figure 5-9**).⁷ E211 forms a hydrogen bond (2.8 Å) with the pyridine nitrogen; for other class IV members, it is suggested that the role of this residue is to maintain the pyridine ring in its protonated form, thus stabilizing the carbanion reaction intermediate. The pyridine ring is further stabilized by hydrophobic residues with a conserved leucine residue (L233) and the backbone of a conserved phenylalanine (F215).

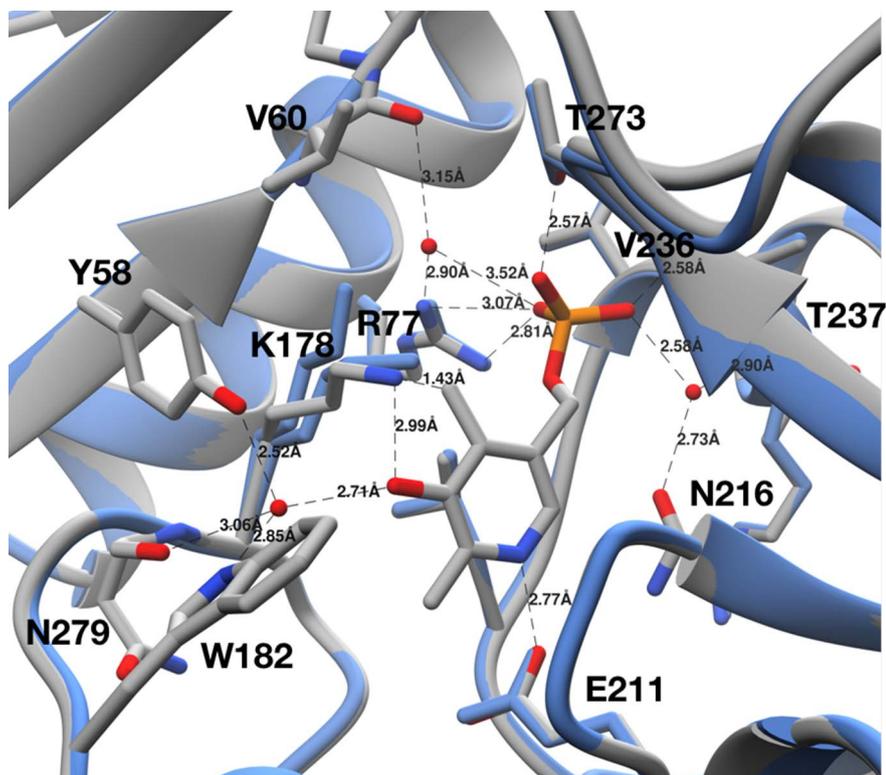


Figure 5-9: Detailed view of the interactions between PLP and *Ts*RTA. The stabilizing hydrogen bonds formed between active site residues and water molecules and PLP are highlighted. To illustrate the high structural identity between *Ts*RTA and *At*RTA, the structure of the latter (PDB entry 4ce5; blue ribbons and sticks) is superposed with Chain A of *Ts*RTA (grey ribbons and sticks).⁷ All panels were generated using Chimera.¹⁹

An analysis of the dimer interface formed with the PDBePISA server (<https://www.ebi.ac.uk/pdbe/pisa/>) revealed an interaction surface area of 2143.5 Å² mediated by 57 interacting residues (Complexation Significance Score (CSS) of 1.0). Contact

surface interactions comprise 30 hydrogen bonds and 10 salt bridges, with an estimated solvation energy gain upon interface formation of $\Delta G^i = -22.6 \text{ kcal mol}^{-1}$.

```

GEO1900 -----MSELLVYMNGEFVPEAQKVSFVDHGF
AF0933 -----MLYVYMDGEFVPEAEKVSIFDHF
ATA-117 MAFSADTSEIVYTHDTHDGLDYITYSDYELDPANPLAGGAAWIEGAFVPPSEARISIFDQGY
ATA-117-Rd11 MAFSADTPEIVYTHDTHDGLDYITYSDYELDPANPLAGGAAWIEGAFVPPSEARISIFDQGF
HoRTA -----DP-GAGVALIEDEIVVPAEARLPLDWF
Aspergillus-fumigatus MASMDKVFSGYARQKLLERSD-----NPFSGKIAYVEGKLVLPSDARIPLLDEGF
Exophiala-xenobiotica MATMEKVFAGYAEARQKVEAST-----NPPAKGVAVWEGKLVVPVNEARIPLMQDGF
Nectria-haematococca MATMDKVFAGYAEARQAVLEASK-----NPLAKGVANIQGLVLPHEARIPLLDQGF
TsRTA MATMDKVFAGYAAARQKAMEAAG-----NPLSEGIAWVEGEMVPLHEARIPMLDEGF
AtRTA MASMDKVFAGYAAARQATLESTETT-----NPPAKGIAWVEGELVPLAEARIPLLDQGF
      . . . : * : * . . . : * :

GEO1900 LYGDGVFEGIRAYNGKVFKLYEHIDRLYDCARVIDLKIPLSKEEFAEAIETLRRNNLRD
AF0933 LYGDGVFEGIRAYNGRVFRLKEHIDRLYDSAKAIDLEIPITKEEFMEIILETLRKNLRD
ATA-117 LHSDVYTYTVFHWNGNAFRLDDHIERLFSNAESMRIIPPLTQDEVKEIALELVAKTELRE
ATA-117-Rd11 YTSDATYTTTFHWNGNAFRLGDHIERLFSNAESIRLIPPLTQDEVKEIALELVAKTELRE
HoRTA LHSDATYDVVAHVWQGRFFRLEEHDRFFEGMADALRMIYPHDREAVSNRLHDVAASGLRD
Aspergillus-fumigatus MHSDLTYDVISVWDGRFFRLDDHLQRILESCDKMRKFPALSSVKNI LAEMVAKSGIRD
Exophiala-xenobiotica LHSDLTYDVPVSWDGRFFRLDDHLDRFELSCSKMRFPKMLPQEVKRI LVDVMAKSGIKD
Nectria-haematococca MHSDLTYDVPVSWDGRFFRLDDHLNRLEASCCKMRLRMLPREEVIKTLVDMVAKSGIRD
TsRTA MRSDLTYDVPVSWDGRFFRLDDHLRSLEASCALRLRLKPLPPEVKKILVEMVAKSGIRD
AtRTA MHSDLTYDVPVSWDGRFFRLDDHITRELEASCCKLRLRLPLPRDQVKQILVEMVAKSGIRD
      . * . : . . : * . * : * : * : * . . . : * . . . : * . . . :

GEO1900 AYIRPIVTRGA-GDLGLDPRKCPSPNVIITKP--WGKLYGDLYEKGLKAITVAIRRNAI
AF0933 AYIRPIVTRGI-GDLGLDPRKQNPVIITKP--WGKLYGDLYEKGLAITVAIRRNRF
ATA-117 AFVSVSITRGYSSTPGERDITKHRPQVYMYAVPYQWIVPFDRIRDGVHAMVAQSVRRTPR
ATA-117-Rd11 AMVTVTITRGYSSTPFERDITKHRPQVYMSACPYPQWIVPFDRIRDGVHLMVAQSVRRTPR
HoRTA AYVEMICTRGQFRAGSRDPRCT-NRFLFAVFPVVIADPAKQETGLHLTISR-MQRIPPP
Aspergillus-fumigatus AFVEIVTRGLTGVRGSKPEDLYNNIYLLVLPYIWMVAPENQHLGGEAII TRTVRRTPP
Exophiala-xenobiotica AFVEIIVTRGLKGVRLKAGESLNNLYMVIQPIYVWVPEMQRTGGSAI IARTVKTSTP
Nectria-haematococca AFVELIVTRGLTGVRGAKPEELLNLLNLYMFIQPIYVWVMDPDVQYTGGRAIVARTVRRVPP
TsRTA AFVEIIVTRGLKGVRSRPEEIV-NRLYMLVQPIYVWVMEPEVQPVGGDAVIARTVRRVPP
AtRTA AFVELIVTRGLKGVGRTRPEDIV-NNLYMVFQPIYVWVMEPDMQVRVGGSAVARTVRRVPP
* : *** . . . : * . . . : * . . . :

GEO1900 DSLPPIKLSLNYLNNILAKIEANAKGGDEAIFLDHNGYISEGSGDNI FIVKNGTITTPPT
AF0933 DALPPIKLSLNYLNNILAKIEANAKGGDEAIFLDHNGYVSESGDNI FVVKNGAIITTPPT
ATA-117 SSIDPQVKNFQWGLIRAVQETHDRGFELPLLLDGDGLLAEGSGFNVVVKDGVVRSFGR
ATA-117-Rd11 SSIDPQVKNFQWGLIRAVQETHDRGFELPLLLDGNLLAEGSGFNVVVKDGVVRSFGR
HoRTA ASVDPTVKNYHWLDMVQALFEAYDRGAETAITVDAEDNVVEGPGFNLFVQGGDLATPAT
Aspergillus-fumigatus GAFDPTTKNLQWGLDTRKGLFEAMDRGATYPFLTDGDTNLTGSGGFNI VLKNGI IYTPDR
Exophiala-xenobiotica GSMDPVTKNLQWGLDTRGMLAQRDGRADYPFLTDGDNLTGSGGFNI VFKDGLVYTPDR
Nectria-haematococca GSIDPTIKNLQWGLDVRGLFEANDRGATYPFLTDGDANLTGSGGFNI VLIKDGVLVTPDR
TsRTA GSIDPTVKNLQWGLFVRGLFEASDRGATYPFLTDGDANLTGSGGFNI VLKDGVLVYTPDR
AtRTA GAIDPTVKNLQWGLDVRGMFEAADRGATYPFLTDGDAHLTEGSGGFNI VLKDGVLVYTPDR
      . . * : * . . . : * . . . : * . . . : * . . . : * . . . :

GEO1900 LNNLKGITRQVVI ELINELEIPFREANIGLFDLYSADEIFVTGTAAEIAPVTVIDGRITV
AF0933 INNLRGITREAVIEIINRLGIPFKETNIGLYDLYTAEVFTGTAAEIAPVVIDGRKIG
ATA-117 -AALPGITRKTVLEIAESLGHEAILADITLAELLDADEVLCCTAGGVWPFVSV DGNPIS
ATA-117-Rd11 -AALPGITRKTVLEIAESLGHEAILADITPAELYDADEVLCSTGGGVWPFVSV DGNPIS
HoRTA -GVLPGVTRRTVIEELGDAHGRGVQAQSVRADTVRGADVEFITSTAGGVMEFVTRVDGALIG
Aspergillus-fumigatus -GVLRGITRKSVIDARANSIDIRLEVVEQYVHSDEIFMCTTAGGIMPITLLDGGQFV
Exophiala-xenobiotica -GVLKGVTRKSVADAANKANGIEMRIEFVPTEMAYQCCECFMCTTAGGVMEITRSDMGQFV
Nectria-haematococca -GVLQGITRKSVIDAARS CGYEIRVEHVP I EATYQADEIILMCTTAGGIMPITLLD DPKV
TsRTA -GVLQGVTRKSVI DVANAKGF E VRVEYVPEAA YH ADEIFMCTTAGGIMPITRSLD GPKV
AtRTA -GVLQGVTRKSVINAAEAFGEIVRVEFVPELAYRCEIFMCTTAGGIMPITLLDGMV
      * : * * * : * : * . . . : * . . . : * . . . :

GEO1900 NGKPGKVTKMLMEKFRERTENEGVEI----YR-----
AF0933 DGKPGBITRKLMEEFKLTSEGVPI----YE-----
ATA-117 DGVPGPVTQSIIRRYWELNVESSLLTPVQY-----
ATA-117-Rd11 DGVPGPVTQSIIRRYWELNVESSLLTPVQYA-----
HoRTA DGKPGPVTALRDAYWRLHEEPWYTO-AVSYERA----AGVRHGD
Aspergillus-fumigatus DGQVGPITTKIWDGYWEMHYNPAYSF-PVDYSG-----
Exophiala-xenobiotica DGKVGVPVTKIWDGYWAMHYDDKYSF-KIDYEESKANGINGVNGVH
Nectria-haematococca DGKVGVPITKAIWDGYWAMHWEDEFSF-KINY-----
TsRTA DGKVGVPITKAIWDGYWEMHYDPAYSF-EIKYQVAEGKPLAGYRFQE
AtRTA GGQIGPITTKIWDGYWAMHYDAAYSF-EIDYNERN-----
      . * * : * . . . : * . . . : * . . . :

```

Figure 5-10: CLUSTAL multiple sequence alignment by MUSCLE (3.8).⁹ Bold red: residues differing from AtRTA. Highlighted in blue: Cysteine forming a disulfide bridge in the tetrameric structure of ATA-117-RD11, and the corresponding glycine in ATA-117.

A tetrameric structure of ATA-117-Rd11 (40% sequence identity) has been deposited in the PDB as 5fr9.²⁰ This structure contains two disulfide bridges at the tetrameric interface (**Figure 5-8c,d**), while glycine residues are found in all of the above mentioned RTAs, as well as wild type ATA-117 (**Figure 5-8d**, **Figure 5-10**). Despite the fact that a tetramer interface was not predicted by PDBePISA, yet in agreement with our experimental findings, a TsRTA

tetramer was generated by applying the symmetry operation (-x, y, -z) to the asymmetric dimer, which superimposed well with the tetramer of ATA-117-Rd11 (**Figure 5-8b**). Additionally, for all of the above mentioned structures of RTAs, except for the RTA from *E. xenobiotica*, a similar crystallographic-tetramer can be obtained (**Figure 5-8c**), yet, interface analysis with PDBePISA, only predicted AtRTA to form a tetramer. Gel-filtration chromatography of AtRTA then revealed a composition of 97% tetramer, 2% dimer and 1% monomer (**Figure 5-7**). Analysis of known tetrameric STAs¹⁶ (PDB entries 4b9b, 4atp, 3n5m, 3a8u, and 5lh9) showed PDBePISA only predicted the first three to be stable tetramers in solution. Thus, the analysis of the “dimer of dimers” structure by PDBePISA was deemed unreliable. Attempts to disrupt the tetrameric structure by simple variation of the ionic strength of the buffer (0-1200 mM NaCl) were made but no effect was observed.

TsRTA_G205C and AtRTA_G207C variants, produced to mimic the equivalent tetramer-bridging G215C mutation of ATA-117-Rd11, exhibited a disrupted quaternary structure immediately following purification: a composition of ca. 55% tetramer, 42% dimer, and 2% monomer and 28% tetramer, 19% dimer, and 52% monomer were observed for TsRTA_G205C and AtRTA_G207C, respectively (**Figure 5-7**), as judged by SEC. This can be attributed to an enhanced energy barrier for tetramer formation due to increased steric hinderance from the cysteines. Indeed, incubation at room temperature with gentle agitation showed slow reconstitution of the tetrameric form, but it tapered off over time (**Figure 5-11**), possibly due to incorrect disulfide bond formation within the dimer. Incubation at 4 °C in the presence of β -mercaptoethanol (10 mM) for two days, followed by dialysis to remove the reducing agent and an overnight incubation with aeration (to form the correct disulfide bond), on the other hand, resulted in 89% tetramer and 11% dimer, and 89% tetramer, 5% dimer, and 6% monomer for TsRTA_G205C and AtRTA_G207C, respectively (**Figure 5-11**). The presence of a disulfide bond bridging two subunits was confirmed by MALDI-TOF MS (**Figure 5-12**). For AtRTA_G207C, a 65% increase in activity was observed following β -mercaptoethanol pre-treatment. This matches the increase in the relative amount of dimer and tetramer from 57% to 94%. Thus, the increase in activity is due to the inactive monomer being converted into the active dimeric and tetrameric forms. For TsRTA_G205C, no such increase was observed implying that both the dimeric and tetrameric forms have similar activities. Both enzymes, displayed only marginally lower specific activities over their respective wild type forms (1.8 U/mg and 2.7 U/mg vs 2.5 U/mg and 3 U/mg for TsRTA_G205C and AtRTA_G207C, respectively) due to a decreased k_{cat} . Inhibition from pyruvate was no longer detectable. The equivalent G207C mutation in AtRTA largely had the

same effect, but, in addition, it increased the K_m of RMBA, lowering the catalytic efficiency further. Any effect on substrate inhibition was less pronounced and statistically insignificant (Table 5-1).

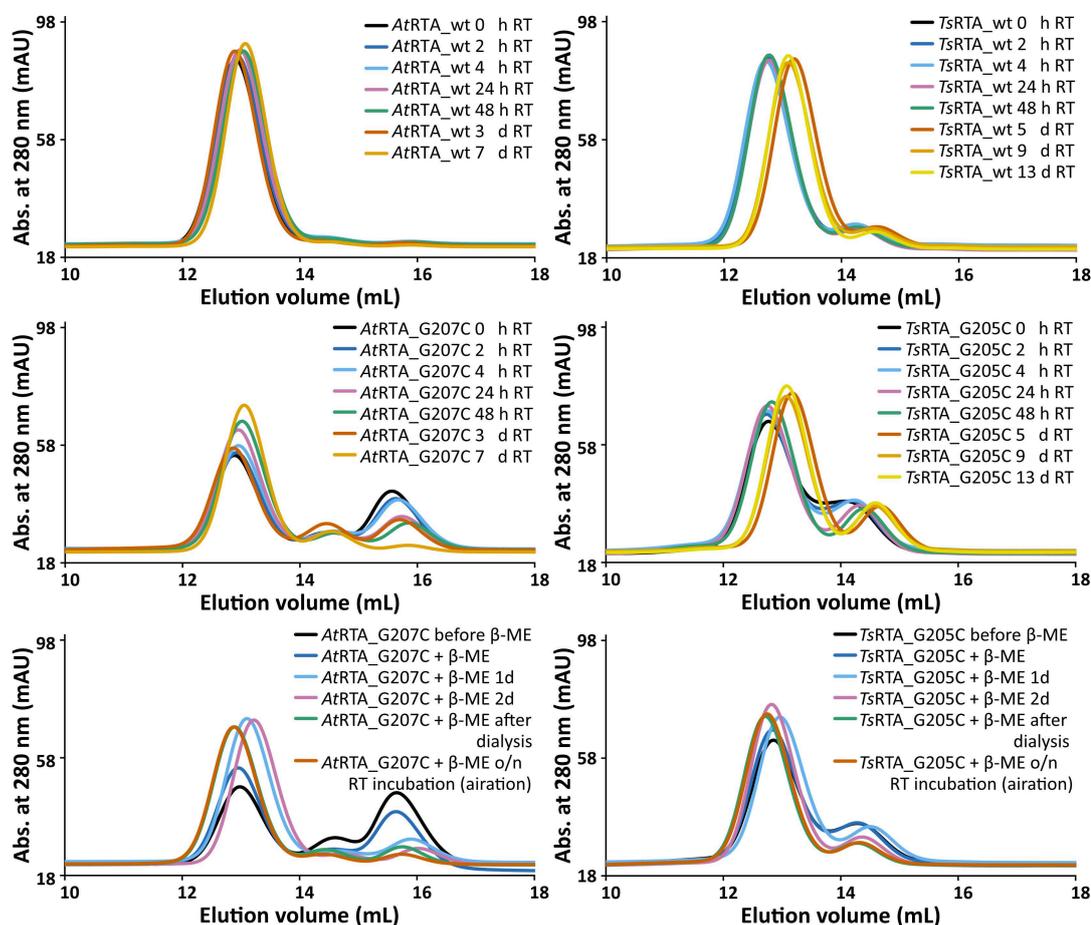


Figure 5-11: Gel-filtration chromatograms. Top: following the incubation of wild-type *At*RTA and *Ts*RTA at ambient temperature with gentle agitation (aeration); Middle: following the incubation of mutant *At*RTA and *Ts*RTA at ambient temperature with gentle agitation (aeration); following the β -mercaptoethanol treatment of *Ts*RTA_G205C and *At*RTA_G207C.

The slow formation of the quaternary structure provided an opportunity to probe the thermostability of *Ts*RTA_G205C and *At*RTA_G207C with different proportions of the tetrameric form. Both mutants showed increased thermostability compared to the corresponding wild type. However, the enzyme samples with a lower proportion of tetramer (In this case, *Ts*RTA_G205C: 75% tetramer, *At*RTA_G207C: 55% tetramer) showed a drop in activity within 30 min, followed by a more stable profile. Samples with 90% tetramer showed a much less pronounced activity loss (Figure 5-5). All samples showed a very similar profile, offset by the initial drop, thus supporting that the tetramer, stabilized by disulfide bonds, has improved thermostability over the dimer. *Ts*RTA_G205C showed higher thermostability than *At*RTA_G207C, and the switch in thermostability observed for the wild type enzymes

between 40 °C and 45 °C was absent in case of the mutants. The temperature optimum was shifted for *TsRTA_G205C* showing considerable activity at 60 °C and 65 °C but not for *AtRTA_G207C* which followed a similar profile to the wild type (**Figure 5-6**).

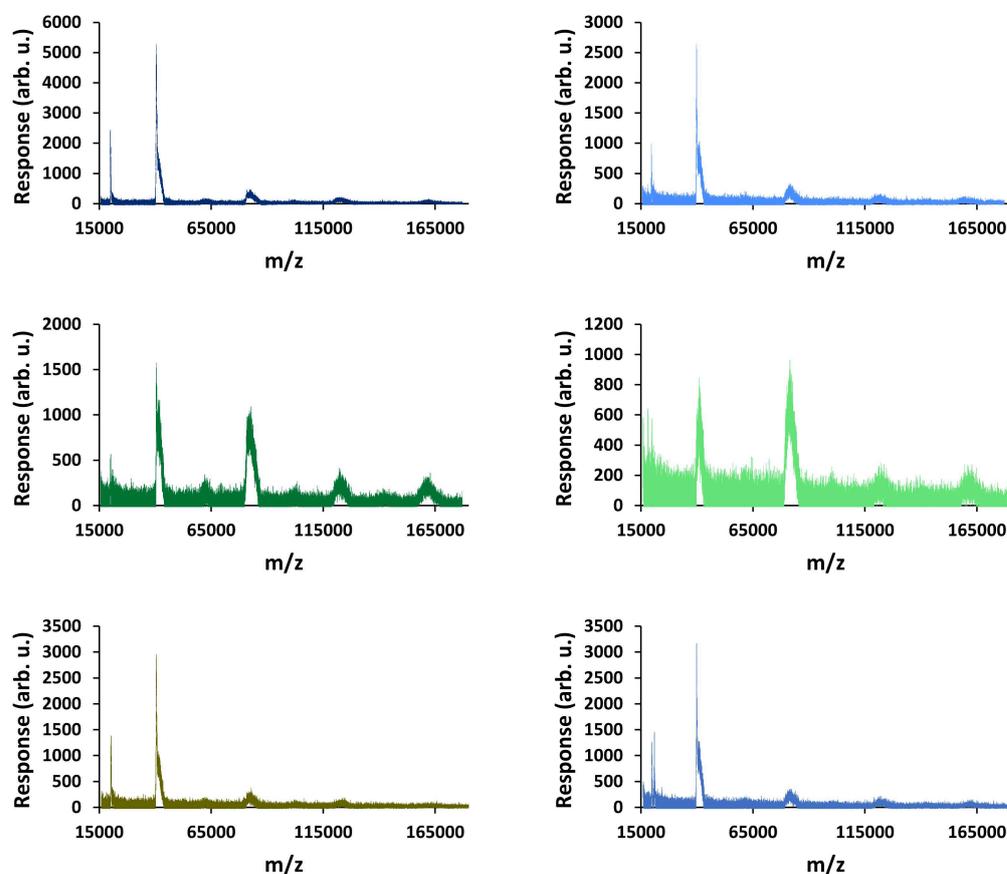


Figure 5-12: MALDI-TOF MS traces. Left: *TsRTA*, right: *AtRTA*. Top: wt. Middle: mutant. Bottom: mutant after incubation with TCEP.

These data suggest two different mechanisms by which these enzymes may lose their activity at elevated temperatures. The first, proceeding via dissociation of the tetramer followed by unfolding of the dimer, and the second being the direct unfolding of the tetramer. The first mechanism presumably is eliminated by the disulfide bond, and it implies that *TsRTA* is inherently more stable toward the second mechanism. This could explain why, at 40 °C, wild type *TsRTA* is more stable than *AtRTA*. Yet, at 45 °C, *TsRTA* might more rapidly dissociate into the dimeric form than *AtRTA*, thus being less stable at higher temperatures (a weaker dimer-dimer interaction is consistent with the higher proportion of dimer observed for *TsRTA* at equilibrium).

5.5 Substrate scope

To further characterize the newly identified TsRTA, the scope toward carbonyl substrates was examined (**Table 5-2**). Aldehydes such as benzaldehyde, cinnamaldehyde, and vanillin were accepted, but only traces of conversion were observed with phenylacetaldehyde. Pyruvate was an excellent substrate. α -Ketoglutarate, the natural ketone acceptor for branched-chain aminotransferases,¹ was not a substrate, as well as the bulky 2,2-dimethylhexan-3-one. However, traces of conversion were observed when the propyl group was shortened to a methyl group. Butanone gave low levels of conversion. Extending the chain-length to hexan-2-one increased conversion. In both cases only the (*R*)-enantiomer was detected, showing excellent discrimination between the methyl and ethyl group in butanone in particular. Cyclohexanone on the other hand gave only low conversions. The heterocyclic ketones tetrahydrofuran-3-one and tetrahydrothiophene-3-one gave higher conversions, yet poor enantioselectivity. While the expected (*R*)-3-aminotetrahydrothiophene was produced, (*S*)-3-aminotetrahydrofuran was preferred, in both cases with ca. 20% *ee*. This is most likely due to different electronic interactions of the negatively polarized oxygen compared to the more neutral and bulkier sulfur atom. Coincidentally, this revealed that the enantiomer produced by the *Halomonas elongata* STA had been mis-assigned as (*S*),²¹ and it is the unexpected (*R*)-enantiomer that is produced preferentially in that case. β -Hydroxy-pyruvate reached its final conversion of ca. 50% after just 30 min, probably due to thermodynamic limitations. Acetophenone (with 5 eq. of isopropylamine) gave only traces of conversion, similarly to 1-indanone and propiophenone. *o*-Fluoroacetophenone and phenoxyacetone gave good conversions. Substituting the oxygen of the latter with a methylene (4-phenylbutanone) resulted in moderate conversion. In all cases only the (*R*)-enantiomer was detected. To verify that TsRTA_G205C has a similar substrate scope, biotransformations with 5 representative substrates which showed lower conversions with the wild type, were carried out with the mutant variant: vanillin, hexan-2-one, *o*-fluoroacetophenone, 4-phenylbutanone, and phenoxyacetone. In all cases higher conversions were achieved after 24h, while conversions after 30 min remained similar to the wild type. This indicates that the higher resting stability of TsRTA_G205C also translates into improved operational stability, increasing the number of turnovers of the mutant. As expected, the mutant maintained excellent enantioselectivity.

Table 5-2: Carbonyl substrate scope (10 mM scale) of wild-type TsRTA, unless stated otherwise.^a

		$ \begin{array}{c} \text{R}^1-\text{C}(=\text{O})-\text{R}^2 + \text{R}^3-\text{CH}_2-\text{NH}_2 \xrightarrow[\text{5 or 10% (v/v) DMSO}]{\text{TsRTA (0.5 mg/mL)}} \text{R}^1-\text{CH}(\text{NH}_2)-\text{R}^2 + \text{R}^3-\text{C}(=\text{O})-\text{R}^4 \\ \text{KP; buffer (50 mM)} \\ \text{0.1 mM PLP} \\ \text{pH 8, 37 }^\circ\text{C} \end{array} $		
Substrate		Conversion (%)		Product ee (%)
		after 30 min	after 24 h	
Benzaldehyde		79 ± 1	89 ± 2	
Phenylacetaldehyde		1.5 ± 0.8	2.5 ± 0.8	
Cinnamaldehyde		38 ± 1	60 ± 1	
Vanillin		16 ± 1 12 ± 1 ^d	72 ± 3 93 ± 3 ^d	
Butanone		<0.5	12 ± 1	>99.5 (R)
Hexan-2-one ^b		30 ± 1 30 ± 1 ^d	57 ± 1 70 ± 2 ^d	>99.5 (R)
Pinacolone		<0.5	1.8 ± 0.8	n.d.
2,2-Dimethylhexan-3-one		<0.5	<0.5	n.d.
Cyclohexanone		<0.5	7.7 ± 0.7	
Tetrahydrothiophene-3-one ^b		5.5 ± 0.7	32 ± 1	21 ± 1 (R)
Tetrahydrofuran-3-one ^b		1.0 ± 0.8	25 ± 1	22 ± 1 (S)
Pyruvate		75 ± 1	95 ± 1	n.d.
β-Hydroxy-pyruvate		52 ± 1	46 ± 1	>99.5 (R)
α-Ketoglutarate		<0.5	<0.5	n.d.
Acetophenone ^{bc}		<0.5	4 ± 3	traces (R)
<i>o</i> -Fluoroacetophenone ^b		34 ± 1 35 ± 2 ^d	70 ± 1 80 ± 3 ^d	>99.5 (R)

... Table continued on next page

Substrate		Conversion (%)		Product ee (%)
		after 30 min	after 24 h	
1-Indanone ^b		1.8 ± 0.8	3.6 ± 0.8	traces (R)
Propiophenone ^b		<0.5	1.2 ± 0.8	traces (R)
Phenoxyacetone ^b		65 ± 1 50 ± 1 ^{de}	81 ± 4 95 ± 1 ^{de}	>99.5 (R)
4-Phenylbutanone ^b		17 ± 1 13 ± 1 ^d	51 ± 3 75 ± 1 ^d	>99.5 (R)

^a Unless stated otherwise, for the carbonyl acceptor scope, 10 mM RMBA was used as the amine donor. Conversions are based on the formation of acetophenone as determined by HPLC and calculated from a calibration curve. All reactions were carried out in triplicate. Standard errors include the SE of the calibration curve.^{22,23} Any conversion <1% may be due to the exchange of PMP for PLP following the first half reaction with RMBA. Enantiomeric excesses were determined after 24 h by chiral GC-FID or chiral RP-HPLC (butanone, tetrahydrofuran-3-one, β -hydroxy-pyruvate).

^b 10% (v/v) DMSO

^c 50 mM isopropylamine

^d TsRTA_G205C

^e 0.1 mg/mL enzyme

“<0.5” = calculated conversion smaller than uncertainty from calibration curve.

“n.d.” = no product detected.

TsRTA accepts the commonly used amine donors RMBA, IPA, and D-alanine (D-Ala), with the other enantiomers of MBA and Ala not being accepted. As no pyruvate removal or recycling system was used, the final conversion with D-Ala and benzaldehyde was only 30%, however this was reached within 30 min. Only low conversion was obtained with β -Ala, and only traces with *o*-xylylenediamine.²⁴ However, *p*-nitrophenethylamine²⁵ was readily accepted, making TsRTA suitable for colourimetric screening during directed evolution,²¹ as was the ‘smart’ amine donor cadaverine²⁶ (**Table 5-3**).

To further investigate the benefit of the increased thermostability of TsRTA_G205C, the intensification of biotransformations with phenoxyacetone with one equivalent of RMBA as the amino donor were investigated for RMBA concentration of 10 mM to 300 mM (incomplete dissolution of phenoxyacetone for 100 mM and higher), employing a catalyst loading of 0.025 mol%. Both variants showed similar performance at the 10 mM scale, but the wild type exhibited a faster drop in conversion at increasing scales, reaching 13% conversion at 300 mM vs 31% for TsRTA_G205C (**Figure 5-13**). In addition, the wild type reached its final conversion within 30 min for concentrations of 50 mM and higher, while TsRTA_G205C showed increased conversions after 30 min even at the 300 mM scale.

Table 5-3: Amine substrate scope (10 mM scale) of wild-type TsRTA.^a

$\text{R}^1-\overset{\text{O}}{\parallel}-\text{R}^2 + \text{R}^3-\overset{\text{NH}_2}{\text{C}}-\text{R}^4 \xrightleftharpoons[\text{5\% (v/v) DMSO, KP}_i\text{ buffer (50 mM), 0.1 mM PLP, pH 8, 37 }^\circ\text{C}]{\text{TsRTA (0.5 mg/mL)}} \text{R}^1-\overset{\text{NH}_2}{\text{C}}-\text{R}^2 + \text{R}^3-\overset{\text{O}}{\parallel}-\text{R}^4$		Conversion (%)	
Substrate		after 30 min	after 24 h
(<i>R</i>)-Methylbenzylamine ^d		75 ± 1	96 ± 1
(<i>S</i>)-Methylbenzylamine ^b		<0.5	<0.5
(±)-Methylbenzylamine ^{bc}		82 ± 3	94 ± 1
L-Alanine		n.d.	n.d.
D-Alanine		29 ± 4	30 ± 4
(±)-Alanine ^c		30 ± 4	31 ± 4
β-Alanine		6 ± 5	12 ± 4
Isopropylamine		28 ± 4	88 ± 4
Cadaverine		23 ± 4	44 ± 3
<i>o</i> -Xylylenediamine		n.d.	7 ± 5
<i>p</i> -Nitrophenethylamine		23 ± 4	53 ± 6

^a Unless stated otherwise, 10 mM benzaldehyde was used as the carbonyl acceptor. Conversions are based on the formation of either acetophenone or benzylamine, as determined by HPLC and calculated from a calibration curve. All reactions were carried out in triplicate. Standard errors include the SE of the calibration curve.^{22,23}

^b 10 mM pyruvate

^c 20 mM

"<0.5" = calculated conversion smaller than uncertainty from calibration curve.

"n.d." = no product detected.

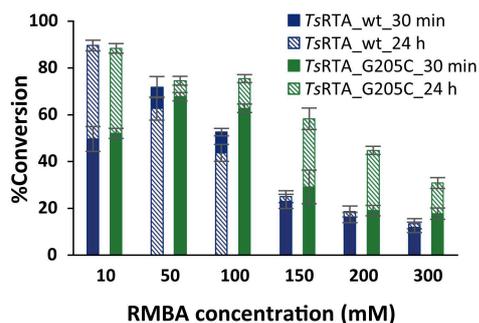


Figure 5-13: Intensification of biotransformations employing equimolar amounts of RMBA and phenoxyacetone, at 37 °C. Reactions contained 0.025 mol% of transaminase, PLP (0.1 mM), DMSO (10% v/v for 10-200 mM, 15% for 300 mM), and K_P -buffer (50 mM, pH 8). Samples were taken after 30 min and 24 h. Conversions are based on the formation of phenoxypropan-2-amine as determined by HPLC and calculated from a calibration curve. All reactions were carried out in triplicate. Error bars represent standard errors and include the SE of the calibration curve.^{22,23}

5.6 Immobilization of TsRTA

Two types of epoxide functionalized methacrylate polymer beads with a 40-60 nm pore diameter were chosen for immobilization studies: HFA403/S and EP403/S, with the former having a longer linker between the epoxide and the polymer support. Three loadings of enzyme were trialed per resin (1, 2, and 5 mg/g_{resin}). In each case complete immobilization was achieved (as determined by the absence of activity and absorbance at 280 nm in the supernatant). Following capping of unreacted epoxide groups, the recovered activities of the immobilized TsRTA (imm-TsRTA) were determined, calculated as the specific activity observed divided by the theoretical specific activity assuming complete immobilization of fully active enzyme (Table 5-4). Highest recovered activities were achieved at the lowest loading and recovered activities were higher with HFA403/S. However, with increasing loading the difference between resins decreased. The resins with the highest loading were then boiled in SDS-PAGE loading buffer, to strip any subunits that weren't covalently attached (**Figure 5-14**). Here, EP403/S performed significantly better with the shorter linker apparently allowing for a better attachment of all subunits. Furthermore, protein leaching from HFA403/S was also observed during the metal desorption step. For this reason, EP403/S was chosen for further immobilization studies.

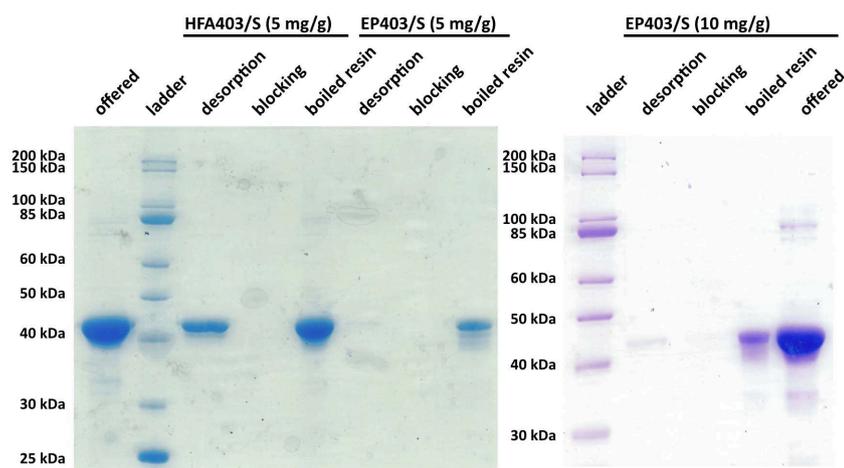
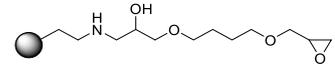


Figure 5-14: SDS-PAGE of the desorption and blocking steps following immobilization. For each lane, the volume loaded onto the gel has been normalized with respect to the total amount of protein offered to the resin, which is given for comparison.

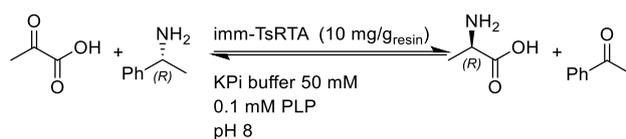
Given the good recovered activity at $5 \text{ mg/g}_{\text{resin}}$ and low subunit loss, a higher loading of $10 \text{ mg/g}_{\text{resin}}$ was trialled. During the previous immobilization it was observed that the resin was yellow prior to the blocking step, but completely colourless after. Incubation with PLP restored that colour. Thus, for this immobilization PLP was supplied in the blocking buffer. A higher recovered activity than was previously achieved suggests that this stabilized the enzyme. Boiling of the $10 \text{ mg/g}_{\text{resin}}$ catalyst still showed excellent resistance toward subunit loss (**Figure 5-14**).

Table 5-4: Specific activities and recovered activities of imm-TsRTA with various enzyme loading on two different resins. All activities were measured in duplicate, with the mean and SE stated. All immobilization reactions were carried out with identical conditions, except the $10 \text{ mg/g}_{\text{resin}}$ loading on EP403/S, where PLP (0.1 mM) was added to the blocking buffer.

	HFA403/S		EP403/S	
				
Enzyme	Specific activity	Recovered	Specific activity	Recovered
Loading	(U/g _{resin})	activity (%)	(U/g _{resin})	activity (%)
$1 \text{ mg/g}_{\text{resin}}$	1.73 ± 0.02	67 ± 2	1.34 ± 0.02	51 ± 1
$2 \text{ mg/g}_{\text{resin}}$	2.755 ± 0.001	53 ± 1	2.17 ± 0.07	42 ± 2
$5 \text{ mg/g}_{\text{resin}}$	5.24 ± 0.09	41 ± 1	4.8 ± 0.4	37 ± 3
$10 \text{ mg/g}_{\text{resin}}$	N/A	N/A	14.9 ± 0.2	58 ± 2

The operational stability of that catalyst was then investigated in flow (packed-bed reactor), using RMBA and pyruvate as substrates (Scheme 5-1). At 10 mM substrate loading, 5 min residence time, 37 °C, and 10% (v/v) DMSO as co-solvent, no loss in the specific activity

of the resin ($13 \pm 1 \text{ U/g}_{\text{resin}}$) was observed after 30 cycles and conversion remained constant (**Figure 5-15a**, one cycle = one CV). That same resin was then used at 45°C with otherwise identical conditions for an additional 20 cycles. No loss in conversion was observed (**Figure 5-15b**). Next, the substrate loading was increased to 20 mM with 20% (v/v) DMSO (45°C). After 20 cycles the same conversion as at the start was reached (**Figure 5-15c**), and the catalyst had maintained its specific activity ($15.9 \pm 0.7 \text{ U/g}_{\text{resin}}$).



Scheme 5-1: Flow reactions used to test the operational stability of imm-TsRTA. 1 equivalent of RMBA was used. 5 min residence time for pyruvate, 3.5 h residence time for butanone. Stereochemistry of alanine not determined, inferred from substrate scope (**Table 5-3**).

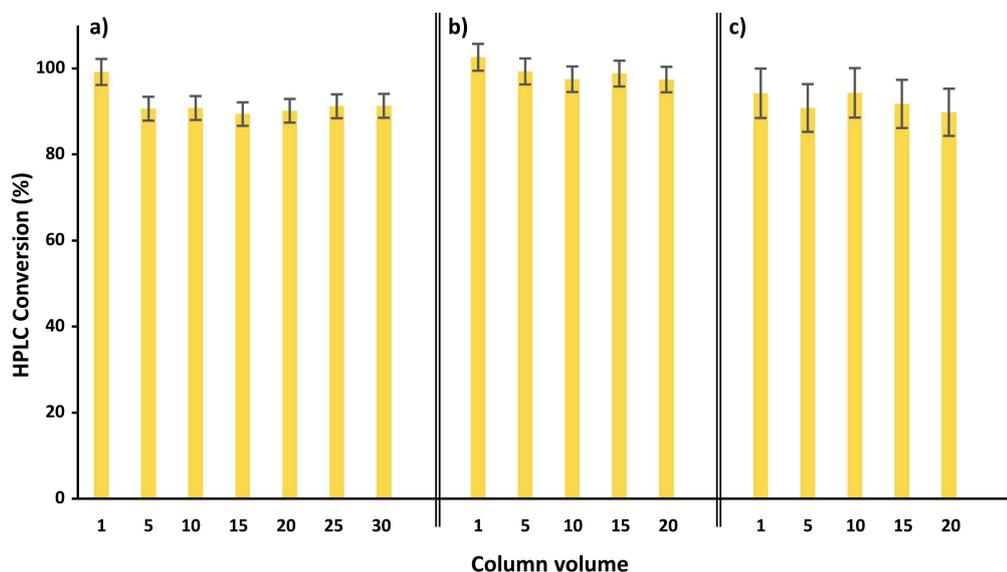


Figure 5-15: Operational stability of imm-TsRTA under increasingly harsh conditions: a) 10 mM RMBA and pyruvate, 10% (v/v) DMSO, 37°C , b) 10 mM RMBA and pyruvate, 10% (v/v) DMSO, 45°C , c) 20 mM RMBA and pyruvate, 20% (v/v) DMSO, 45°C . Potassium phosphate buffer (50 mM), 0.1 mM PLP, pH 8. 5 min residence time. (Scheme 5-1) Conversion based on the depletion of RMBA, as determined by HPLC and calculated from a calibration curve. Error bars represent the standard error from that calibration curve.

5.7 Conclusion

The predominant tetrameric composition in solution of two RTAs, TsRTA and its homologue AtRTA, has been discovered. Upon comparisons made between the crystal structure of TsRTA and the crystal structures of RTAs deposited in the PDB, a likely interface for the dimer-dimer interaction was also identified. This interface was then probed by introducing a cysteine residue mimicking ATA-117-Rd11, which stabilized both RTAs and thus rationalized the role of this mutation in the directed evolution of ATA-117-Rd11. This dimer-dimer interaction was observed in all but one deposited crystal structure of RTAs, although the position of the equilibrium between the dimeric and tetrameric forms in solution is not known in most cases. SEC studies of the RTA from *Nectria haematococca* have been reported to be consistent with a dimeric form.¹² While the wild-type TsRTA was only marginally more stable than AtRTA, this difference was significantly amplified in the mutant variants. Thus, we propose that two mechanisms, one where tetramer dissociation precedes unfolding and one where unfolding occurs in the tetramer state, might play a role in the inactivation of these RTAs. Clearly, the full inactivation kinetics of both enzymes would need to be studied in detail and in particular the relative kinetic and thermodynamic stability⁴ of the tetrameric and dimeric forms for the wild-type enzymes should be elucidated.

The full understanding of the quaternary structure of RTAs is particularly important with regard to rational approaches to enzyme engineering. The stabilization of the tetrameric form of RTAs through the mutation described herein, which could not have been predicted from a dimeric model, appears to be a promising strategy to stabilize RTAs (the residue mutated herein being present in all reported RTA structures). Additional mutations stabilizing this interface may also be possible, creating more stable and therefore more evolvable²⁷ and industrially useful catalysts.

Additionally, TsRTA readily immobilized on the epoxy-functionalized methacrylate resin EP403/S, and showed excellent operational stability at up to 20 mM substrate loading, 20% (v/v) DMSO, and 45 °C. Further studies should include a comparison to the immobilized TsRTA_G205C variant, to ascertain whether the stabilizing disulfide bridges offer any additional stability once the enzyme is immobilized, as immobilization itself presumably also prevents dissociation of the tetramer.

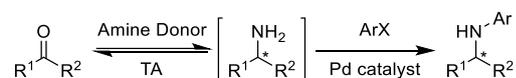
5.8 References

1. Höhne, M., Schätzle, S., Jochens, H., Robins, K. & Bornscheuer, U. T. Rational assignment of key motifs for function guides in silico enzyme identification. *Nat. Chem. Biol.* **6**, 807–813 (2010).
2. Guo, F. & Berglund, P. Transaminase biocatalysis: optimization and application. *Green Chem.* **19**, 333–360 (2017).
3. Slabu, I., Galman, J. L., Lloyd, R. C. & Turner, N. J. Discovery, Engineering, and Synthetic Application of Transaminase Biocatalysts. *ACS Catal.* **7**, 8263–8284 (2017).
4. Bommarius, A. S. & Paye, M. F. Stabilizing biocatalysts. *Chem. Soc. Rev.* **42**, 6534–6565 (2013).
5. Littlechild, J. A. *et al.* Natural methods of protein stabilization: Thermostable biocatalysts. *Biochem. Soc. Trans.* **35**, 1558–1563 (2007).
6. Xia, J., Zhao, J. X., Sang, J., Chen, G. J. & Du, Z. J. *Halofilum ochraceum* gen. Nov., sp. nov., a gammaproteobacterium isolated from a marine solar saltern. *Int. J. Syst. Evol. Microbiol.* **67**, 932–938 (2017).
7. Łyskowski, A. *et al.* Crystal structure of an (R)-selective ω -transaminase from *Aspergillus terreus*. *PLoS one* **9**, e87350 (2014).
8. Heckmann, C. M. Construction of Expression Vectors for Selective His-tag Removal. (University of Nottingham, 2017).
9. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
10. Schätzle, S. *et al.* Enzymatic asymmetric synthesis of enantiomerically pure aliphatic, aromatic and arylaliphatic amines with (R)-selective amine transaminases. *Adv. Synth. Catal.* **353**, 2439–2445 (2011).
11. Davies, M. T. A universal buffer solution for use in ultra-violet spectrophotometry. *Anal.* **84**, 248–251 (1959).
12. Sayer, C. *et al.* The substrate specificity, enantioselectivity and structure of the (R)-selective amine: Pyruvate transaminase from *Nectria haematococca*. *FEBS J.* **281**, 2240–2253 (2014).
13. Telzerow, A. *et al.* Amine Transaminase from *Exophiala xenobiotica* – Crystal Structure and Engineering of a Fold IV Transaminase that Naturally Converts Biaryl Ketones. *ACS Catal.* **9**, 1140–1148 (2019).
14. Thomsen, M. *et al.* Crystallographic characterization of the (R)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **70**, 1086–1093 (2014).
15. Isupov, M. N. *et al.* Thermostable Branched-Chain Amino Acid Transaminases From the Archaea *Geoglobus acetivorans* and *Archaeoglobus fulgidus*: Biochemical and Structural Characterization. *Front. Bioeng. Biotechnol.* **7**, 1–16 (2019).
16. Börner, T. *et al.* Explaining Operational Instability of Amine Transaminases: Substrate-Induced Inactivation Mechanism and Influence of Quaternary Structure on Enzyme-Cofactor Intermediate Stability. *ACS Catal.* **7**, 1259–1269 (2017).
17. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2256–2268 (2004).
18. Laskowski, R. A., Watson, J. D. & Thornton, J. M. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res.* **33**, W89-93 (2005).
19. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
20. Cuetos, A. *et al.* Catalytic Promiscuity of Transaminases: Preparation of Enantioenriched β -Fluoroamines by Formal Tandem Hydrodefluorination/Deamination. *Angew. Chemie Int. Ed.* **55**, 3144–3147 (2016).
21. Planchestainer, M., Hegarty, E., Heckmann, C. M., Gourlay, L. J. & Paradisi, F. Widely applicable

- background depletion step enables transaminase evolution through solid-phase screening. *Chem. Sci.* **10**, 5952–5958 (2019).
22. Theodorou, D., Zannikou, Y. & Zannikos, F. Estimation of the standard uncertainty of a calibration curve: Application to sulfur mass concentration determination in fuels. *Accredit. Qual. Assur.* **17**, 275–281 (2012).
 23. *Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement.* (2012). doi:0 948926 15 5.
 24. Green, A. P., Turner, N. J. & O'Reilly, E. Chiral amine synthesis using w-transaminases: An amine donor that displaces equilibria and enables high-throughput screening. *Angew. Chemie Int. Ed.* **53**, 10714–10717 (2014).
 25. Baud, D., Ladkau, N., Moody, T. S., Ward, J. M. & Hailes, H. C. A rapid, sensitive colorimetric assay for the high-throughput screening of transaminases in liquid or solid-phase. *Chem. Commun.* **51**, 17225–17228 (2015).
 26. Gomm, A., Lewis, W., Green, A. P. & O'Reilly, E. A New Generation of Smart Amine Donors for Transaminase-Mediated Biotransformations. *Chem. Eur. J* **22**, 12692–12695 (2016).
 27. Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5869–5874 (2006).

6 Accessing chiral *N*-aryl amines in a sequential transaminase-Buchwald-Hartwig amination cascade

In the previous chapters much of the focus was on small, primary chiral amines. However, many important molecules, such as XL888 mentioned in **chapter 3**, contain chiral *N*-aryl amines. One important method of preparing *N*-aryl amines is the Buchwald-Hartwig amination (BHA), a (typically) palladium catalysed coupling of aryl halides and amines, which could be used to derivatize the primary amines obtained from transaminases, ideally without the need for purification of intermediates (**Scheme 6-1**).

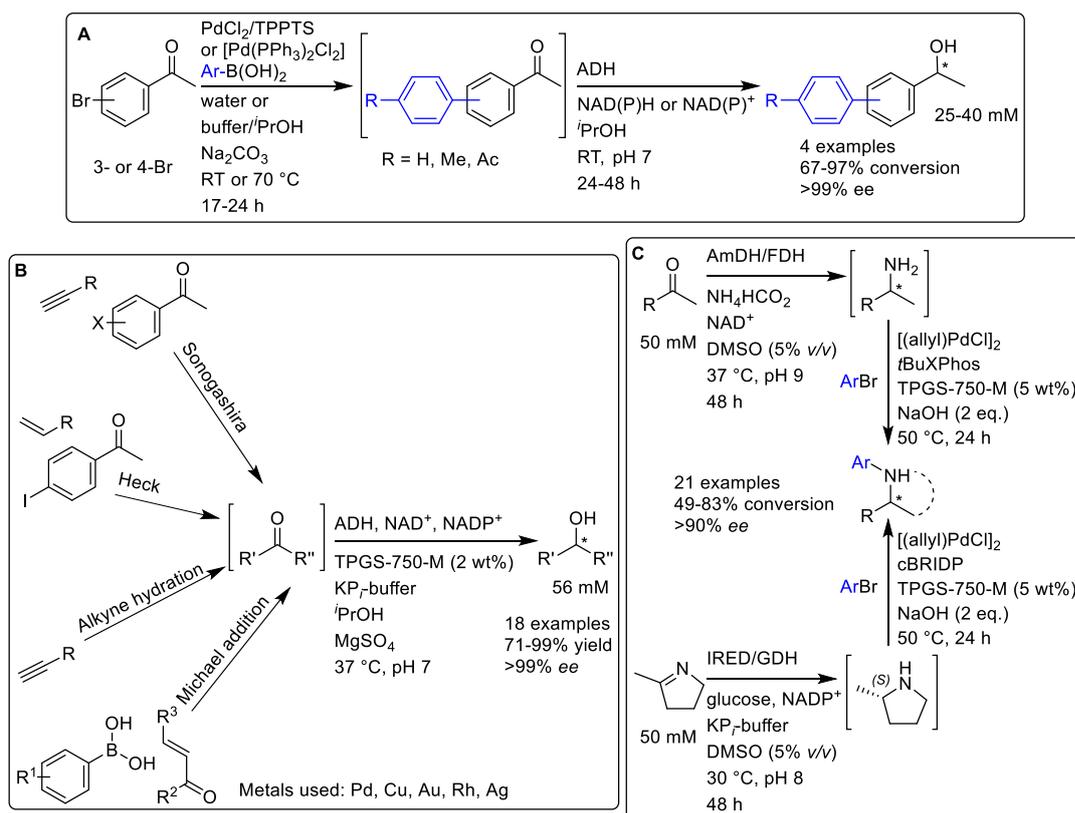


Scheme 6-1: Synthesis of chiral *N*-aryl amines: A prochiral ketone is transaminated using a transaminase biocatalyst and a sacrificial amine donor, which is then arylated in a BHA, ideally without isolation of the intermediate primary amine. X = Cl, Br, or I. R¹, R² = H, alkyl, or aryl.

Combining chemical and enzymatic steps poses several key challenges, specifically in the compatibility of the reaction environments. Either the enzyme has to be able to tolerate the conditions required by the chemical reaction (organic solvents, reagents such as heavy metals, high temperatures, high or low pH, etc.) or the chemical step has to work in the environment required by the enzyme (aqueous medium, buffer salts, cofactors, mild temperature, mild pH, etc.). Additionally, biotransformations typically employ much more diluted conditions than chemical transformations. These challenges notwithstanding, there are several examples in the literature of such chemo-enzymatic cascades which overcome some of these challenges by separating the reactions in time (one-pot two-step). Two such examples have already been mentioned in the introduction (**Scheme 1-6**): Bornscheuer's group¹ used a Suzuki-Miyaura coupling to produce a biaryl ketone as the substrate for a subsequently added transaminase. Latham *et al.*² used halogenases to produce a variety of aryl halides which were then used in a Suzuki coupling. In both cases, substrate concentrations were low (<5 mM) and therefore of limited synthetic use.

Additional examples include Suzuki cross-couplings in water to produce biaryl ketones as substrates for alcohol dehydrogenases (ADHs) by the Gröger group (**Scheme 6-2A**),^{3,4} and a variety of transition-metal catalysed transformations (Sonogashira and Heck cross-couplings, Michael additions, and alkyne hydrations) to again form substrates for ADHs by the Lipshutz group (**Scheme 6-2B**).⁵ In the former examples, careful design of the palladium catalysed reaction allowed that reaction medium to be compatible with the ADH (after pH

adjustment and co-factor addition following the completion of the Suzuki cross-coupling). In the latter example, the use of the detergent DL- α -tocopherol methoxypolyethylene glycol succinate (TPGS-750-M) facilitated the transition metal catalysed reactions, which were then diluted approx. 10-fold to make them compatible with the ADH. In both cases the concentrations achieved after the second biocatalytic step were somewhat higher (25-56 mM).



Scheme 6-2: **A:** Suzuki-coupling in water to produce biaryl ketones as substrates for ADHs to give chiral biaryl alcohols in a one-pot system. The use of (tris(3-sulfonatophenyl)phosphine hydrate, sodium salt (TPPTS) as the ligand allowed the Suzuki coupling to proceed at ambient temperature.^{3,4} **B:** A variety of transition metal catalysed reactions to give pro-chiral ketones, which are then reduced *in-situ* to chiral alcohols using ADH.⁵ **C:** Chiral amines are produced from prochiral ketones (top) or a cyclic imine (bottom), which are then arylated *in-situ* using a palladium catalyst.⁶

In the case of a BHA subsequent to a transaminase-catalysed synthesis of a primary amine, an additional challenge emerges. Typically, transaminases require an excess of a sacrificial amine donor for high conversions, thus two amines will be present in the BHA step, giving the desired product and an unwanted side-product. Indeed, at the outset of this project no example of an enzyme catalysed amine formation followed by a BHA existed in the literature. Recently, the Turner group published the first work of an enzyme-BHA cascade, using the detergent TPGS-750-M (**Scheme 6-2C**);⁶ however, while they demonstrated one-pot two-step strategies using IREDs and AmDHs, they reported that the

use of a transaminase required extraction of the amine (separating the amine product from the highly water-soluble alanine used as the amine donor) and could not be carried out in a one-pot reaction. In this chapter, a sequential cascade involving a transaminase followed by a BHA on a 100 mM scale (in the biotransformation) is described, which tolerates the presence of alanine in the BHA step.

6.1 Initial optimizations using benzylamine as a model substrate

Initial work was carried out by Alice Keil,⁷ focussing on the coupling of benzylamine to 3-bromotoluene, since benzylamine can be obtained almost quantitatively with one equivalent of SMBA in the transaminase reaction. Initial attempts employed *t*BuXPhos (**Figure 6-1**) as the ligand and [(Cinnamyl)PdCl]₂ as the palladium source in the presence of TPGS-750-M;⁸ however, this was quickly switched to a biphasic system with toluene and using tetrabutylammonium bromide (TBAB) as a phase transfer catalyst, which allowed for easier sampling of the reaction (TPGS-750-M reactions were highly heterogeneous).⁹ While typical substrate loadings for the BHA are 0.5-2 M,⁸⁻¹¹ the substrate concentration in the biotransformation was much lower (50 mM). Indeed, while excellent conversions (90%) were obtained at a typical scale (1 mmol, 1 mL of each phase), under the more dilute conditions when using a biotransformation (50 μmol, 1 mL of each phase) conversion dropped to 38%. In Alice's work, immobilized HEWT was used, to avoid interference of enzyme with the palladium catalysed reaction as reported by Latham *et al.*²

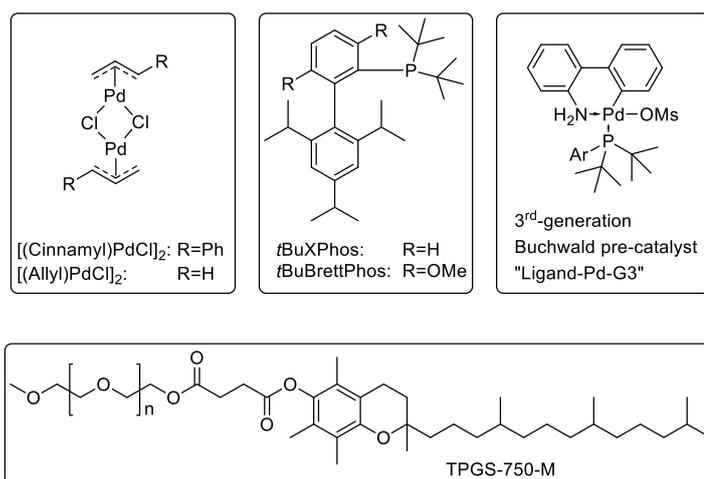
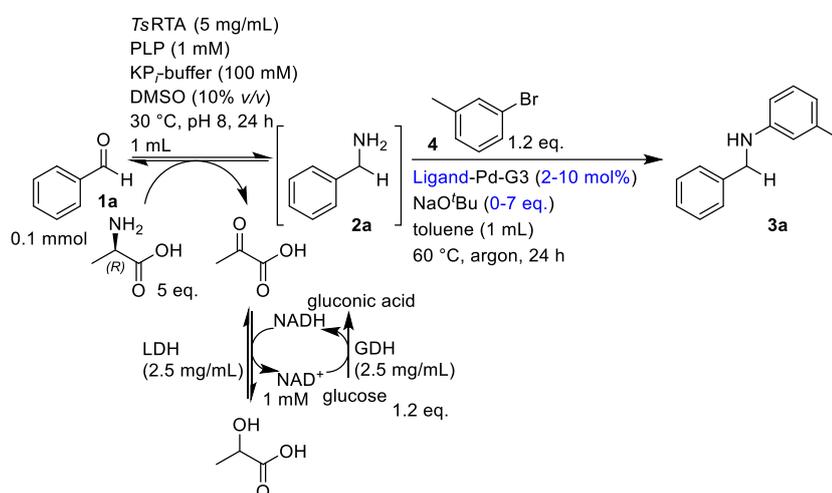


Figure 6-1: Top: Palladium sources and ligands used in this chapter. Bottom: the surfactant TPGS-750-M ($n \approx 16$).

To further optimize the BHA coming from a biotransformation, the palladium source was changed to third-generation Buchwald pre-catalysts, and a *t*BuBrettPhos ligand (**Figure 6-1**), which had also previously been shown to work well in water,⁸ was considered alongside *t*BuXPhos. Additionally, reactions containing alanine as an additive were trialed.

In preliminary tests, conversions of the BHA step were generally low (<40%) and this was shown to be the case also for downscaled reactions containing no buffer or any other additives. However, *t*BuBrettPhos-Pd-G3 showed conversion (TLC) also in the presence of alanine but this was initially not observed for *t*BuXPhos-Pd-G3. Thus, the optimization focussed on this pre-catalyst, employing a biotransformation containing *Ts*RTA (**Chapter 5**), D-Alanine, and an LDH/GDH pyruvate removal system (**Scheme 6-3**). Encouraged by the results of Cosgrove *et al.*,⁶ the phase transfer catalyst TBAB was removed, and soluble enzymes were used, simplifying the set-up. Since the BHA clearly performs better at higher substrate concentrations, the substrate loading in the biotransformation was increased two-fold to 100 mM (at this loading the starting ketone was not fully soluble). Results from this optimization are shown in **Table 6-1**.



Scheme 6-3: Model system used for the optimization of the BHA, showing all the components and by-products from the transaminase-catalysed biotransformation. Conditions optimized are shown in blue.

The result of the optimization showed that a higher catalyst loading is required (5 mol%) (**Table 6-1**, entries 1–3) as well as additional equivalents of base (**Table 6-1**, entries 2&4–6). The higher amount of base is due to 5 eq. getting consumed in an acid-base reaction with the protonated amine. Using the optimized conditions, swapping *t*BuBrettPhos-Pd-G3 back to *t*BuXPhos-Pd-G3 resulted in almost identical conversions (**Table 6-1**, entries 2&7). Curiously, Cosgrove *et al.*⁶ who were using *t*BuXPhos and [(allyl)PdCl]₂ reported no conversion in the presence of alanine.

Table 6-1: Optimization of the BHA step, coming from a biotransformation of benzaldehyde with *TsRTA* (Scheme 6-3).

Entry	Ligand	Catalyst loading (mol%)	NaO ^t Bu (equivalents)	%Conversion ^a
1	<i>t</i> BuBrettPhos	2	7	28 (28)
2	<i>t</i> BuBrettPhos	5	7	81-92 (81-92) ^b
3	<i>t</i> BuBrettPhos	10	7	99 (97)
4	<i>t</i> BuBrettPhos	5	0	0 (0)
5	<i>t</i> BuBrettPhos	5	2	14 (14)
6	<i>t</i> BuBrettPhos	5	5	46 (45)
7	<i>t</i> BuXPhos	5	7	72-77 (71-77) ^b

^a Conversion of the BHA step followed by overall conversion in parentheses.

Conversions determined by RP-HPLC, comparing peak areas of benzaldehyde, benzylamine, and 3-benzylaminotoluene, corrected for their response factor, after the BHA step.

^b Range observed across two independent experiments.

To investigate whether the choice of a biphasic system, with the catalyst primarily contained in the organic phase and alanine in the aqueous phase, the third-generation pre-catalyst, or the presence or absence of TBAB were crucial for this success, a set of control experiments was carried out. Reactions containing TBAB and either *t*BuBrettPhos-Pd-G3 or *t*BuXPhos-Pd-G3 gave identical results (Table 6-2, entries 1&2). Next, experiments were set-up using *t*BuXPhos and [(allyl)PdCl]₂ either in a biphasic system without TBAB or adding an aqueous solution of TPGS-750-M (5 wt%) instead of toluene (Table 6-2, entries 3&4). Again, conversions were similar with the TPGS-750-M reaction giving a slightly lower conversion. This leads to the conclusion that neither the choice of pre-catalyst, nor the choice of a biphasic system (with or without phase-transfer catalyst) over the micellar system is the key to the success of this reaction. In the absence of the detailed conditions of Cosgrove *et al.*,⁶ one may speculate that the reason for their negative result lies within their choice of base, assuming they employed 2 equivalents of sodium hydroxide as for their other BHAs (see Table 6-1, entries 2&4–6).

Table 6-2: Control reactions investigating the effect of reaction conditions in the presence of alanine.

Entry	TBAB (mol%)	Pd catalyst	Solvent	%Conversion ^a
1 ^b	5	<i>t</i> BuXPhos-Pd-G3	Buffer + toluene	90–96 ^e
2 ^b	5	<i>t</i> BuBrettPhos-Pd-G3	Buffer + toluene	91–95 ^e
3 ^c	0	<i>t</i> BuXPhos + [(allyl)PdCl] ₂	Buffer + toluene	89
4 ^d	0	<i>t</i> BuXPhos + [(allyl)PdCl] ₂	Buffer + TPGS-750-M	75

^a Conversions determined by RP-HPLC, comparing peak areas of benzylamine, and 3-benzylaminotoluene, corrected for their response factor.

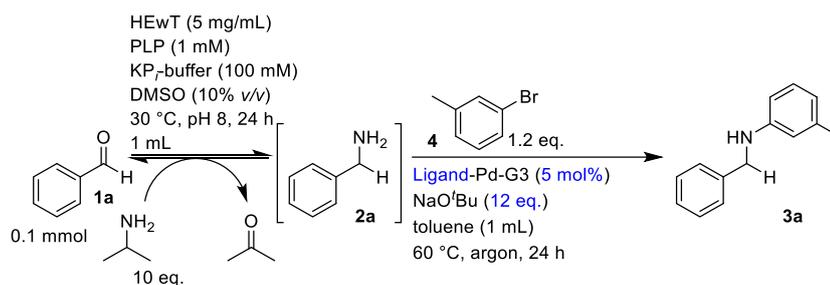
^b Benzylamine (1 mmol), 3-Br-toluene (1.2 eq), TBAB (5 mol%), DL-Ala (1 eq.), NaO^tBu (2 eq.), pre-catalyst (1 mol%), KP_T-buffer (50 mM, pH 8; 0.9 mL), DMSO (0.1 mL), toluene (1 mL), 60 °C, 20–22 h.

^c Benzylamine (1 mmol), 3-Br-toluene (1.2 eq), DL-Ala (1 eq.), NaO^tBu (2 eq.), [(allyl)PdCl]₂ (1 mol%), *t*BuXPhos (1.2 mol%), KP_T-buffer (50 mM, pH 8; 0.9 mL), DMSO (0.1 mL), toluene (1 mL), 60 °C, 24 h.

^d Benzylamine (1 mmol), 3-Br-toluene (1.2 eq), DL-Ala (1 eq.), NaO^tBu (2 eq.), [(allyl)PdCl]₂ (1 mol%), *t*BuXPhos (1.2 mol%), KP_T-buffer (50 mM, pH 8; 0.9 mL), DMSO (0.1 mL), TPGS-750-M (5% (w/v) in water; 1 mL), 60 °C, 24 h.

^e Range observed across two independent experiments.

Next, the use of *t*BuBrettPhos-Pd-G3 and *t*BuXPhos-Pd-G3 in the biphasic system for biotransformations employing IPA as the amine donor was explored (**Scheme 6-4**). Based on the optimized conditions previously developed for alanine, 5 mol% of catalyst and *n*+2 equivalents of NaO^tBu (where *n* denotes the equivalents of amine donor employed in the biotransformation) were used. The results are summarized in **Table 6-3**. While *t*BuBrettPhos showed a higher ratio of the desired product (**3a**) to the undesired IPA product, conversions between the two ligands were comparable.



Scheme 6-4: Model system for the combined cascade when using IPA as the amine donor, showing all the components and by-products from the transaminase-catalysed biotransformation. No acetone removal to shift the equilibrium was used. Conditions (previously) optimized are shown in blue.

Table 6-3: Comparison of *t*BuBrettPhos-Pd-G3 and *t*BuXPhos-Pd-G3 when using a biotransformation employing IPA as the amine donor (**Scheme 6-4**). Data quoted as a range across two independent experiments.

Entry	Ligand	%Conversion ^a	Ratio 3a:IPA product ^b
1	<i>t</i> BuBrettPhos	65-85 (64-85)	1.7-2.0
2	<i>t</i> BuXPhos	76-85 (76-85)	0.9-1.1

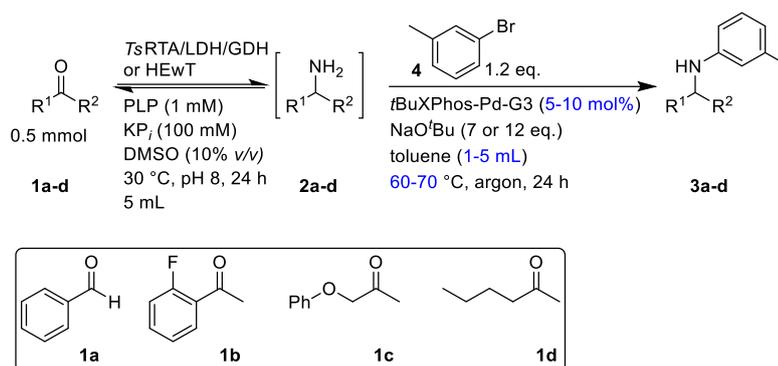
^a Conversion of the BHA step followed by overall conversion in parentheses. Conversions determined by RP-HPLC, comparing peak areas of benzaldehyde, benzylamine, and 3-benzylaminotoluene, corrected for their response factor, after the BHA step.

^b Ratio determined by RP-HPLC, comparing response-factor-corrected peak areas for BnNHTol and IPATol.

6.2 Scale-up and ketone substrate scope

To improve the synthetic potential, reactions were scaled up to a 0.5 mmol scale, by increasing the volume of the biotransformation 5-fold, using the significantly cheaper *t*BuXPhos-Pd-G3 precatalyst (£139 vs £270 for 1 g, Sigma Aldrich) as the differences between both pre-catalysts were minimal, as seen above. Conversions dropped compared to the 0.1 mmol scale, but were improved by decreasing the ratio of toluene to the aqueous biotransformation (1:5 v:v) and by removing precipitated protein from the biotransformation by centrifugation (**Table 6-4** entries 1–3). These scaled-up conditions were then employed starting from a variety of pro-chiral ketones (**Scheme 6-5**).

Moving from benzaldehyde to the structurally similar *o*-fluoroacetophenone (**1b**) (benzylic carbonyl/amine), conversions in the BHA step were reduced significantly (**Table 6-4**, entry 4). Increasing the temperature to 70 °C had negligible effect on the reaction (entry 5), whereas increasing the catalyst loading to 10 mol% resulted in increased conversion to approx. 50% (entry 6). Phenoxyacetone (**1c**) and hexan-2-one (**1d**) (aliphatic carbonyl/amine; with and without a distal aromatic group) showed similar conversions using the increased catalyst loading (entries 7&8). Starting from all three prochiral substrates the system gave exclusively the (*R*)-enantiomer of the final product, with no racemisation being observed during the BHA step.



Scheme 6-5: Scaled-up synthesis of *N*-arylamines. Conditions that are being varied are shown in blue. TsRTA/LDH/GDH: *D*-Ala (5 eq.), *D*-Glc (1.2 eq.), NAD⁺ (1 mM), LDH (2.5 mg/mL), GDH (2.5 mg/mL), TsRTA (5 mg/mL). HEwT: IPA (10 eq.), HEwT (10 mg/mL).

Table 6-4: Scale-up of the cascade to 0.5 mmol and ketone substrate scope using TsRTA and *D*-Ala (LDH/GDH) (Scheme 6-5).

Entry	Carbonyl	Catalyst loading (mol%)	organic:aqueous (v:v)	Temperature BHA (°C)	%Conversion ^a	ee (%)
1	1a	5	5:5	60	46–62 (45–62) ^b	
2	1a	5	1:5	60	75–83 (73–82) ^b	
3	1a	5	1:5 ^d	60	93–94 (93–93) ^b	
4	1b	5	1:5 ^d	60	29–47 (28–45) ^b	
5	1b	5	1:5 ^d	70	37–39 (36–37) ^b	
6	1b	10	1:5 ^d	60	53–55 (51–53) ^b	>99.5 (<i>R</i>) ^e
7	1c	10	1:5 ^d	60	56–59 (56–59) ^b	>99.5 (<i>R</i>) ^e
8	1d	10	1:5 ^d	60	47–50 (46–49) ^c	>99.5 (<i>R</i>) ^f

^a Conversion of the BHA step followed by overall conversion in parentheses. Range across two independent experiments.

^b Conversions determined by RP-HPLC, comparing peak areas of ketone, primary amine, and *N*-arylamine, corrected for their response factor, after the BHA step.

^c Conversions determined by GC-FID, comparing peak areas of ketone, primary amine (acetylated), and *N*-arylamine, corrected for their response factor, after the BHA step.

^d Precipitated protein removed by centrifugation after biotransformation.

^e Enantiomeric excess determined by chiral RP-HPLC, following acetylation. Stereochemistry assigned based on known absolute configuration of the primary amine intermediate.

^f Enantiomeric excess determined by chiral GC-FID. Stereochemistry assigned based on known absolute configuration of the primary amine intermediate.

Using these conditions for biotransformations employing the (*S*)-selective transaminase (STA) HEwT and using IPA (10 eq.) as the donor (increasing the amount of NaO^tBu as before), the good selectivity in the case of benzylamine was maintained (Table 6-5, entry 1). However, for the chiral amines, conversions dropped significantly and production of the unwanted IPA coupling product increased, with the ratio of desired to undesired product matching approximately their concentration in the reaction (Table 6-5 entries 2-4). Again, good *ees* (89 – >99.5%) were achieved in the final product, yielding the

(*S*)-enantiomer. The *ees* of the final product match those observed in the primary amine intermediate obtained with HEwT (data not shown).

Table 6-5: Scaled-up ketone substrate scope (0.5 mmol) using HEwT and IPA (**Scheme 6-5**). Precipitated protein removed by centrifugation after biotransformation.

Entry	Carbonyl	Catalyst loading (mol%)	%Conversion ^a	<i>ee</i> (%)	Ratio desired product/IPA product
1	1a	5	63–68 (63–68) ^b		1.3–1.5 ^f
2	1b	10	25–28 (22–25) ^b	>99.5 (<i>S</i>) ^d	0.14–0.15 ^f
3	1c	10	27–29 (27–29) ^b	97 (<i>S</i>) ^d	0.23–0.25 ^f
4	1d	10	24–28 (22–26) ^c	89 (<i>S</i>) ^e	0.08–0.09 ^f

^a Conversion of the BHA step followed by overall conversion in parentheses. Range across two independent experiments.

^b Conversions determined by RP-HPLC, comparing peak areas of ketone, primary amine, and *N*-arylamine, corrected for their response factor, after the BHA step.

^c Conversions determined by GC-FID, comparing peak areas of ketone, primary amine (acetylated), and *N*-arylamine, corrected for their response factor, after the BHA step.

^d Enantiomeric excess determined by chiral RP-HPLC, following acetylation. Stereochemistry assigned based on known absolute configuration of primary amine intermediate.

^e Enantiomeric excess determined by chiral GC-FID. Stereochemistry assigned based on known absolute configuration of primary amine intermediate.

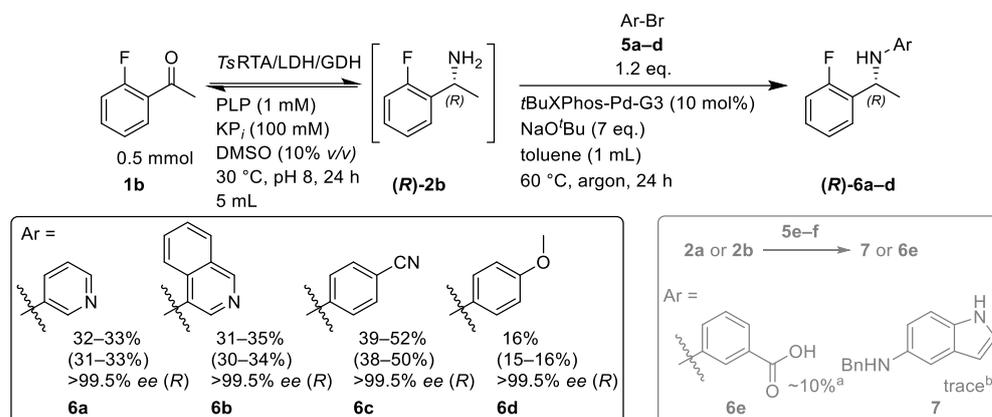
^f Ratio of peak areas, corrected by their response factor. Determined using same method as conversion (GC-FID or RP-HPLC).

Both the decrease in conversion between benzylamine and the chiral amines, as well as the selectivity observed between benzylamine and IPA, imply that increased steric bulk of the branched α -carbon reduces the efficiency of the BHA catalyst. Buchwald and co-workers recently showed that removal of the *p*-*i*Pr-group in a BrettPhos-derived ligand can expand the substrate scope to α -tertiary amines (in dioxane or THF).¹⁰ However, this ligand (GPhos) is currently prohibitively expensive, with the price at the time of writing being £407 for 250 mg (Sigma Aldrich). In comparison, the price for *t*BuXPhos and *t*Bu-BrettPhos is £37.40 and £315.00 for 1 g, respectively. A similar strategy may prove successful in improving yields/lowering the required catalyst loading in the current system. Additionally, in most of the reactions containing alanine (but not all), the formation of varying amounts of “palladium black” was observed (not quantified), but rarely in the reactions containing IPA. Eliminating this formation is probably also key to further improve conversions, either through ligand design/screening or possibly improvements in the air-free technique may be required.

6.3 Aryl halide substrate scope

To verify the utility of this reaction system, exploring the substrate scope with respect to the aryl halide is also necessary. To this end, several additional aryl bromides have been employed in the BHA step (**Scheme 6-6**), using the same method as above. Three of the

substrates (**5a,b,f**) chosen contained nitrogen heterocycles, common features in drug molecules and considered privileged scaffolds.¹² 4-Bromobenzonitrile (**5c**) and 4-bromoanisole (**5d**) were chosen as examples of phenyls with electron withdrawing and electron donating groups, respectively. Finally, 3-bromobenzoic acid (**5e**) was chosen for its acidic functional group.



Scheme 6-6: Aryl halide substrate scope, starting from prochiral ketone **1b**. Precipitated protein removed by centrifugation after biotransformation. Conversions of the BHA step followed by overall conversion in parentheses. Range across two independent experiments. Conversions determined by RP-HPLC, comparing peak areas of ketone, primary amine, and *N*-arylamine, corrected for their response factor, after the BHA step. Enantiomeric excess determined by chiral RP-HPLC, following trifluoroacetylation (**6a-c**) or acetylation (**6d**). Stereochemistry assigned based on known absolute configuration of primary amine intermediate. TsRTA/LDH/GDH: *D*-Ala (5 eq.), *D*-Glc (1.2 eq.), NAD⁺ (1 mM), LDH (2.5 mg/mL), GDH (2.5 mg/mL), TsRTA (5 mg/mL). In grey: attempted synthesis of standards **6e** and **7**; NaO^{*t*}Bu (2 eq.), KP_{*f*}-buffer (50 mM, pH 8; 0.9 mL), DMSO (0.1 mL), toluene (1 mL), 60 °C, argon, 23 h. ^a**2b** (1 mmol), **5e** (1.2 eq.), tBuXphos-Pd-G3 (5 mol%). ^b**2a** (1 mmol), **5f** (1.2 eq.), tBuXphos-Pd-G3 (4.5 mol%).

Aryl bromides **5a-c** all reacted well, giving conversions slightly lower than 4-bromotoluene (**4**) used above. However, the electron rich **5d** gave much lower conversions of only 16%, coinciding with a substantial formation of palladium black. This was unexpected, as during the synthesis of the racemic standards (**Section 2.15.3**) it actually showed conversions on par with **4** and **5a** (93 vs 96 and 94%, respectively) and performed better than **5b** or **5c** (79 and 85%, respectively). It thus appears that the poor conversions are due to a combined effect of the electron-rich substrate and the other components of the biotransformation. As before, the BHA step did not alter the excellent enantioselectivity obtained during the biotransformation.

On the other hand, attempts to synthesise the standards of **6e** and **7** (from **5e** and **5f**, respectively) were unsuccessful (**Scheme 6-6**, grey insert). In the case of the former, only approx. 10% conversion were achieved (assuming the largest unidentified peak in the HPLC trace corresponds to the product and estimating the response factor based on the starting

materials). Partitioning of the acid into the aqueous phase and the amine into the organic phase due to the alkaline conditions was observed, which is a likely reason for the low conversions. In the case of the latter, only traces of the desired coupling product as well as low amounts of coupling through the indole nitrogen (dimers and trimers) were observed by LC-MS. Changing the conditions to more closely resemble the system used by the Goss group,¹³ who successfully coupled anilines to unprotected indoles (THF instead of toluene, potassium carbonate as the base), prevented coupling through the indole nitrogen but did not improve the formation of the desired coupling product. Thus, their synthesis in the combined cascade was not attempted. Further work on this is currently ongoing.

6.4 Conclusions

The synthesis of chiral *N*-arylamines in a sequential transaminase–BHA cascade has been described. In its current form, conversions of approx. 50% can be achieved starting from diverse pro-chiral ketones, without the need for purification or even extraction of the intermediate. The current system can tolerate a large excess of alanine, although the formation of palladium black mainly in the presence of alanine suggests that a better catalyst (better ligand) may be desirable. Coupling to alanine is not observed. Unbranched α -carbons (benzylamine) are preferred. Again, better catalysts or ligands may improve the conversions in the case of α -secondary amines.

While a one-pot two-step reaction was possible, removing precipitated protein from the biotransformation in a quick centrifugation improved conversions. It may well be that removal of the soluble enzymes as well would result in further improvements. However, doing so by e.g. ultrafiltration would further complicate the reaction set-up, and eliminate its main advantage which is the minimal processing that is required prior to the BHA step, significantly increasing the time-efficiency of the synthesis. Indeed, given the high equivalents of base and catalyst (and the expensive ligand) required for the sequential system, it is likely that separation of the intermediate may be preferred in a production route. However, in a drug discovery route generating diverse and enantiopure compounds quickly is key. In this case, the high costs and modest conversions may be acceptable trade-offs for high throughput.

6.5 References

1. Dawood, A. W. H., Bassut, J., de Souza, R. O. M. A. & Bornscheuer, U. Combination of the Suzuki-Miyaura-cross coupling reaction with engineered transaminases. *Chem. Eur. J* **24**, 16009–16013 (2018).
2. Latham, J. *et al.* Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C-H activation. *Nat. Commun.* **7**, 1–8 (2016).
3. Burda, E., Hummel, W. & Gröger, H. Modular chemoenzymatic one-pot syntheses in aqueous media: Combination of a palladium-catalyzed cross-coupling with an asymmetric biotransformation. *Angew. Chemie Int. Ed.* **47**, 9551–9554 (2008).
4. Borchert, S., Burda, E., Schatz, J., Hummel, W. & Gröger, H. Combination of a Suzuki cross-coupling reaction using a water-soluble palladium catalyst with an asymmetric enzymatic reduction towards a one-pot process in aqueous medium at room temperature. *J. Mol. Catal. B: Enzym.* **84**, 89–93 (2012).
5. Cortes-Clerget, M. *et al.* Bridging the gap between transition metal- and bio-catalysis via aqueous micellar catalysis. *Nat. Commun.* **10**, 1–10 (2019).
6. Cosgrove, S. C., Thompson, M. P., Ahmed, S. T., Parmeggiani, F. & Turner, N. J. One-Pot Synthesis of Chiral N-Arylamines by Combining Biocatalytic Aminations with Buchwald–Hartwig N-Arylation. *Angew. Chemie Int. Ed.* **59**, 18156–18160 (2020).
7. Keil, A. A sequential chemo-enzymatic transamination—Buchwald-Hartwig cross coupling cascade. (University of Nottingham, 2019).
8. Wagner, P. *et al.* t-BuXPhos: a highly efficient ligand for Buchwald–Hartwig coupling in water. *Green Chem.* **16**, 4178 (2014).
9. Naber, J. R. & Buchwald, S. L. Packed-bed reactors for continuous-flow C-N cross-coupling. *Angew. Chemie Int. Ed.* **49**, 9469–9474 (2010).
10. McCann, S. D., Reichert, E. C., Arrechea, P. L. & Buchwald, S. L. Development of an aryl amination catalyst with broad scope guided by consideration of catalyst stability. *J. Am. Chem. Soc.* **142**, 15027–15037 (2020).
11. Noël, T. *et al.* Palladium-catalyzed amination reactions in flow: Overcoming the challenges of clogging via acoustic irradiation. *Chem. Sci.* **2**, 287–290 (2011).
12. Vitaku, E., Smith, D. T. & Njardarson, J. T. Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals. *J. Med. Chem.* **57**, 10257–10274 (2014).
13. Renault, Y. J. G. *et al.* Buchwald Hartwig diversification of unprotected halotryptophans, halotryptophan containing tripeptides and the natural product baretin in aqueous conditions. *Chem. Commun.* **55**, 13653–13656 (2019).

7 Conclusions and Outlook

The application of transaminases in the synthesis of chiral amine building blocks has been broadened. By screening a panel of transaminases, with process intensification in mind, a suitable STA and RTA were identified for the synthesis of (*S*)- and (*R*)-2-aminobutane, respectively. Both enzymes were immobilized and used in flow continuously for 1 week in a multigram-scale synthesis. While conversions were only moderate (approx. 70%), *ees* were excellent (>99.5%, following the introduction of a single point mutation in the STA). Indeed, state-of-art transition metal catalysts can only achieve very low *ees* and similar conversions. Additionally, lipases, the most common biocatalytic way to access small chiral amines, also show only poor enantioselectivity for this substrate. Thus, while higher conversions are obviously desirable, those achieved here are quite competitive. A significant challenge in this synthesis that needs to be overcome is product isolation. While pure product was obtained from a distillation, the recovery was only approx. 50% (on a lab-scale). Further improvements (possibly using an industrial distillation set-up) are clearly needed for this to be a viable production route. Additionally, recycling of unreacted butanone and IPA may also be desirable.

Attempts to synthesis 2,2-dimethylhexan-3-amine were not successful. Here, screening of the same panel of transaminases only revealed candidates with at best traces of conversion (which could not be confirmed). Attempts of directed evolution, using a rational CASTing approach on two candidates, also were not successful. This was due to a combination of irreproducibility in the screening (particularly false-positives), throughput, and the approach to generating the libraries for that compound. Clearly, both better high-throughput screening of transaminases as well as smaller more sophisticated libraries are needed. In particular, focussing exclusively on the active site may have been too limited and future work should also take into account the path of the substrate through the entrance of the active site. Molecular dynamic simulations, or tools such as CaverDock, may prove useful.

Applying transaminases to the synthesis of tetrahydrofuran-3-amine revealed an unexpected variation of enantiomeric outcome of the reaction, depending on the reaction conditions. Using the same enzyme, *ees* of the product ranged from 19% (*R*) to 70% (*S*). This behaviour was initially observed during intensification experiments, where increasing substrate loading resulted in increasing (*R*)-selectivity (or decreasing (*S*)-selectivity). However, careful experiments then demonstrated that increasing ionic strength favours production of the (*R*)-enantiomer, while increasing organic content, as well as increasing pH,

favour the (*S*)-enantiomer. The reason for this behaviour is not yet understood, but subtle structural changes of the enzyme may play a role.

Recruitment of new RTAs serendipitously revealed the tetrameric structure of these enzymes. While several crystal-structures of RTAs have been reported, their quaternary structure in solution is often unknown. Indeed, comparison of the crystal-structures of the two RTAs studied in this work to those reported of other RTAs suggests that most RTAs have at least the potential to assemble into a tetramer. Both RTAs studied here exist in an equilibrium between a dimeric structure and a tetrameric structure, with the tetramer being in higher proportion. However, different RTAs may have different equilibria and favour the dimer. Using this information, rational mutations were introduced forming disulfide bridges across the tetrameric interface which stabilized the RTAs (only when correctly assembled into the tetramer), highlighting the importance of that interface. Clearly, identifying whether any given RTA is tetrameric or dimeric is important before attempting any rational mutagenesis, in particular if increased stability of the catalyst is desired.

Finally, transaminase-catalysed (formal) reductive amination has been combined with Buchwald-Hartwig amination, giving access to chiral *N*-arylamines which are common features in important molecules. By eliminating the need for isolation of the primary amine intermediate, libraries of compounds can be accessed quickly. While the system works with either alanine or IPA as the amine donor, the former gives a cleaner reaction as no cross-coupling to alanine is observed. In its current form, overall conversions are only modest (50%); however, further improvements in the choice of palladium catalyst or the use of immobilized enzymes (facilitating their easy removal) may increase conversions.

Clearly, both the applicability and robustness of transaminases has been expanded in this work. However, more work needs to be done to include an even broader set of substrates. There are other small pro-chiral ketones, such as 4,4,4-trifluorobutanone, where competing reactions (most likely imine formation, but possibly also hydration) prevent the transaminase from forming the desired product. Reaction and/or enzyme engineering are needed to overcome this barrier. The switch in enantioselectivity with tetrahydrofuran-3-one needs to be understood and exploited. The role of the tetramer for RTA stability, and its prevalence, also need to be explored further. There may be additional mutations that can stabilize this interface, without introducing a kinetic barrier to tetramer assembly. Finally, better palladium catalysts or better reaction conditions need to be developed to increase the performance of the transaminase-BHA system. Indeed, for biocatalysis to be utilized to

its full potential, barriers between traditional synthesis and biocatalysis need be overcome. This probably needs to happen through two combined approaches; namely, expansion of the toolbox of biocatalysts that can work under more “chemical” conditions (such as non-aqueous environments, high and low pH, etc.) and the development of chemical catalysts that can tolerate typical conditions and components in biotransformations (e.g. water, enzymes, buffers, mild pH, co-factors).