

**Engineering Native Amine Dehydrogenases
for the Production of Chiral Amines**

Megan Bennett

PhD

University of York

Chemistry

October 2022

Abstract

Chiral amines are fundamental building blocks in synthetic chemistry routinely used in the production of pharmaceuticals, agrochemicals or in the food industry. In recent years, given the need for more renewable and sustainable approaches to the synthesis of these chiral amines, biocatalysts have been investigated. The consideration of biocatalytic approaches has largely led to enantioselective routes to produce chiral amines often with better yields and less harsh reaction conditions and components. More recently, a class of enzymes, found by sequence driven searches of metagenomic databases and samples, known as native Amine Dehydrogenases (nat-AmDHs) have been used for the asymmetric reductive amination of a range of carbonyl compounds. Herein a selection of nat-AmDHs have been structurally and biochemically studied using a range of techniques such as X-ray crystallography, UV-vis spectrometry, and GC-FID analysis. The canonical structural functionalities of these enzymes were found to be consistent throughout. However, some aspects of these previously uncharacterised nat-AmDHs displayed interesting new properties. Namely the increased active site space in the binding pocket of MATOUAmDH2 which permitted the rational engineering of this enzyme to bind larger, more sterically hindered, substrates to produce pharmaceutically relevant amines such as 2-aminonorborene. These results provide new knowledge and protein scaffolds for the further engineering of nat-AmDHs for the production of primary and secondary chiral amines of both biological and chemical significance.

List of Contents

Abstract	3
List of Contents	5
List of Figures	9
List of Tables	17
Acknowledgements	19
Author's Declaration	20
1. Introduction	21
1.1 Synthesis and application of chiral amines	21
1.1.1 Reductive amination	23
1.1.2 Racemic resolution	30
1.2 Biocatalysts for the production of chiral amines	32
1.2.1 Transaminases	33
1.2.2 Imine Reductases and Reductive Aminases	38
1.2.2.1 <i>AspRedAm</i>	40
1.2.3 Amino Acid Dehydrogenases and Amine Dehydrogenases	44
1.2.3.1 Opine Dehydrogenases	49
1.2.3.2 Improving the promiscuity of Amine Dehydrogenases	51
1.3 Native Amine Dehydrogenases	53
1.3.1 <i>CfusAmDH</i> , <i>MsmeAmDH</i> and <i>MicroAmDH</i>	55
2. Aims of the project	58
3. Methods	60
3.1 Molecular Biology	60
3.1.1 Primer and Recombinant Plasmid Design	60

3.1.1.1 Site Directed Mutagenesis Design	61
3.1.2 Polymerase Chain Reaction (PCR)	62
3.1.2.1 Site Directed Mutagenesis PCR	64
3.1.2.2 DpnI Digestion of PCR Products	64
3.1.3 In-Fusion® Cloning and Stellar™ Competent Transformations	65
3.1.4 Confirmation of Fragments via Gel Electrophoresis	67
3.1.4.1 Double Digests	67
3.1.4.2 Minipreps and Sequencing	67
3.1.5 BL21(DE3) Transformations	68
3.1.6 Expression of Recombinant Genes	69
3.2 Protein Production and Purification	69
3.2.1 Protein Production	69
3.2.2 Protein Purification	70
3.2.2.1 Nickel Affinity Chromatography	70
3.2.2.2 Size Exclusion Chromatography (SEC)	71
3.2.3 SDS-PAGE Analysis	73
3.3 Crystallisation and structural studies	74
3.3.1 Generating crystals	74
3.3.1.1 Initial Screening	75
3.3.1.2 Optimisations	75
3.3.1.3 Fishing	76
3.3.2 Co-crystallisation	76
3.3.2.1 Ligand Soaking	77
3.3.3 Data Collection and Processing	77
3.3.4 Molecular Replacement and Refinement	78
3.3.5 Deposition and Visualisation	78

3.4 Kinetics	79
3.4.1 Spectroscopically Monitoring NAD(P)H Oxidation	79
3.4.2 pET22b(+) MATOUAmDH2 Kinetics	80
3.4.3 pETYSBLIC-3C MATOUAmDH2 and Mutant Kinetics	80
3.4.4 Data Processing	81
3.5 Biotransformations	81
3.5.1 pET22b(+) MATOUAmDH2 Biotransformations	81
3.5.2 pETYSYBLIC-3C MATOUAmDH2 and Mutant Biotransformations	83
3.5.2.1 Cyclic Substrates Based Reactions	84
3.5.2.2 Aliphatic Substrates and Ammonia Reactions	84
3.5.2.3 β -Tetralone and Ammonia Reactions	85
3.5.3 Peak Identification and Data Processing	85
4. Results and Discussion	86
4.1 Characterisation of pET22b(+) MATOUAmDH2	86
4.1.1 Discussion	97
4.2 Co-crystallisation of MATOUAmDH2 with cyclohexylamine	102
4.2.1 Discussion	112
4.3 Engineering MATOUAmDH2 for biocatalysis	115
4.3.1 Discussion	133
4.4 Other nat-AmDHs of interest	142
4.4.1 <i>Cfus</i> AmDH W145A	142
4.4.2 <i>Tther</i> AmDH	148
4.4.3 <i>Micro</i> AmDH	152
4.4.4 Discussion	155
5. Summary and Conclusions	159
Appendices	161

Chemical Glossary	170
Abbreviations	176
References	179

List of Figures

1.1 Synthesis and application of chiral amines

Figure 1: Various uses of chiral amines in pharmaceutical, food and agricultural chemistry.

Figure 2: Chemical synthesis of chiral amines from functionalised centres.

Figure 3: Chemical synthesis of chiral amines from unfunctionalised carbonyl centres.

Figure 4: Synthesis of rivastigmine adapted from Interquim and Zentiva.

Figure 5: Synthesis of rivastigmine adapted from Zhejiang Hisun Pharmaceutical.

Figure 6: Synthesis of oseltamivir from ethylbenzoate.

Figure 7: Synthesis of oseltamivir from 5e.

Figure 8: The reductive amination step in the synthesis of maraviroc.

Figure 9: The active site of DDP-4 in complex with (*R*)-sitagliptin.

Figure 10: The synthesis of (*R*)-sitagliptin.

Figure 11: Chiral resolution of racemic amines *via* fractional crystallisation.

Figure 12: Kinetic resolution of racemic amines via acylation-based reactions.

1.2 Biocatalysts for the production of chiral amines

Figure 13: PLP-dependent reactions catalysed by transaminases.

Figure 14: Synthesis of (*R*)-sitagliptin catalysed by (*R*)-selective transaminase ATA-117.

Figure 15: PLP-PMP mechanism for transaminases.

Figure 16: The synthesis of (*S*)-rivastigmine using Vf.ω-TA.

Figure 17: The synthesis of (1*S*)-1-(1,1'-biphenyl-2-yl)ethanamine using Vf.ω-TA.

Figure 18: Biological assembly and active site of Vf.ω-TA in complex with PMP.

Figure 19: Reactions catalysed by Imine Reductases (IREDs).

Figure 20: Exemplar 'reductive amination' reactions enabled by Imine Reductases (IREDs).

Figure 21: Reaction catalysed by *AspRedAm*.

Figure 22: Exemplar reductive amination reactions catalysed by *AspRedAm*.

Figure 23: Biological assembly and active site of *AspRedAm* in complex with NADP(H) and (*R*)-rasagiline.

Figure 24: Monomeric superimposition of *AolRED* and *AspRedAm* in complex with associated ligands and NADP(H).

Figure 25: Reactions catalysed by Amino Acid Dehydrogenases (AADHs).

Figure 26: Catalytic monomer and active site of *LeuDh* in complex with NAD(H) and L-leucine.

Figure 27: Catalytic monomer and active site of *PheDH* in complex with NAD(H) and L-phenylalanine.

Figure 28: Exemplar reductive amination reactions catalysed by engineered Amino Acid Dehydrogenases (AADHs) to form artificial AmDHs.

Figure 29: Examples of reactions catalysed by native and engineered Opine Dehydrogenases (OpDHs).

Figure 30: Catalytic monomer and active site of *OpDH* in complex with NAD(H) and L-arginine.

Figure 31: Exemplar reductive amination reactions catalysed by *Ch1-AmDH* and *Rs-AmDH*.

Figure 32: The one-enzyme synthesis of benzaldehyde from benzylic acid using *LE-AmDH-v1*.

1.3 Native Amine Dehydrogenases

Figure 33: Reactions catalysed by 2,4-diaminopentanoate Dehydrogenase (2,4-DAPDH) and *AmDH4* (4OP-AmDH).

Figure 34: Open and closed forms of *AmDH4* and its active site in complex with NAD⁺ and (4*S*)-4AP.

Figure 35: Reaction catalysed by native Amine Dehydrogenases (nat-AmDHs).

Figure 36: Carbonyl substrate scope of native Amine Dehydrogenases (nat-AmDHs).

Figure 37: Active sites of native AmDHs *CfusAmDH* and *MsmeAmDH*.

3.1 Molecular Biology

Figure 38: Recombinant plasmid map of pETYSBLIC-3C with the MATOUAmDH2 insert.

Figure 39: Structurally informed alanine-based mutant design for MATOUAmDH2.

Figure 40: Schematic of the generation of MATOUAmDH2 based recombinant plasmids via PCR, DpnI digestion and In-Fusion® Cloning.

3.2 Protein Production and Purification

Figure 41: Scheme of the steps involved in Nickel Affinity Chromatography.

Figure 42: Scheme of the steps involved in Size Exclusion Chromatography (SEC).

3.3 Crystallisation and structural studies

Figure 43: Diagram depicting a typical 48-well optimisation screen design.

3.4 Kinetics

Figure 44: Experimental design for monitoring reductive amination using NAD(P)H-dependent UV-vis spectrometry.

3.5 Biotransformations

Figure 45: Programmed GC-FID oven temperatures for analysing cyclohexanone products and reactants.

Figure 46: Programmed GC-FID oven temperatures for analysing cyclic products and reactants.

Figure 47: Programmed GC-FID oven temperatures for analysing aliphatic products and reactants.

Figure 48: Programmed GC-FID oven temperatures for analysing 2-tetralone products and reactants.

4.1 Characterisation of pET22b(+) MATOUAmDH2

Figure 49: Amino acid sequence alignment of MATOUAmDH2, *Cfus*AmDH and *Msme*AmDH.

Figure 50: Nickel Affinity Chromatogram for MATOUAmDH2 expressed in the pET22b(+) vector.

Figure 51: Size Exclusion Chromatogram for MATOUAmDH2 expressed in the pET22b(+) vector.

Figure 52: SDS-PAGE analysis of protein fractions from pET22b(+) MATOUAmDH2 protein purification.

Figure 53: Kinetic plot for pET22b(+) MATOUAmDH2 with cyclohexanone at 50°C.

Figure 54: GC-FID traces for pET22b(+) MATOUAmDH2 reactions with ammonia and cyclohexanone.

Figure 55: GC-FID traces for pET22b(+) MATOUAmDH2 reactions with methylamine and cyclohexanone.

Figure 56: Reaction progressions of biotransformations using pET22b(+) MATOUAmDH2 with ammonia and cyclohexanone.

Figure 57: Reaction progression of a biotransformation using pET22b(+) MATOUAmDH2 with methylamine and cyclohexanone.

Figure 58: Crystal forms from initial screening of pET22b(+) MATOUAmDH2 in complex with NADP⁺.

Figure 59: Crystals and diffraction patterns from optimisations of pET22b(+) MATOUAmDH2 in complex with NADP⁺.

Figure 60: Structural properties of pET22b(+) MATOUAmDH2 in complex with NADP⁺.

Figure 61: Secondary structure visualisation of pET22b(+) MATOUAmDH2.

Figure 62: Visualisation of the binding pocket of pET22b(+) MATOUAmDH2 and associated NADP⁺ omit map.

Figure 63: Crystals and diffraction patterns from trials of pET22b(+) MATOUAmDH2 in complex with NADP⁺ and pentylamine.

Figure 64: Open and closed forms of pET22b(+) MATOUAmDH2 in complex with NADP⁺.

Figure 65: Comparison of ligand free-pET22b(+) MATOUAmDH2 and *Cfus*AmDH active sites.

4.2 Co-crystallisation of MATOUAmDH2 with cyclohexylamine

Figure 66: PCR product bands from MATOUAmDH2 insert amplification analysed via gel electrophoresis.

Figure 67: PCR product bands from pETYSBLIC-3C vector amplification, linearisation and digestion, analysed via gel electrophoresis.

Figure 68: Product bands from the double restriction digest of the recombinant MATOUAmDH2:pETYSBLIC-3C cloning plasmid.

Figure 69: Nickel Affinity Chromatogram for MATOUAmDH2 expressed in the pETYSBLIC-3C vector.

Figure 70: Size Exclusion Chromatogram for MATOUAmDH2 expressed in the pETYSBLIC-3C vector.

Figure 71: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C MATOUAmDH2 protein purification.

Figure 72: Crystal forms from initial screening of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and either cyclohexylamine or 2-aminocyclohexanone.

Figure 72: Crystal forms from 48-well optimisations of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and either cyclohexylamine or 2-aminocyclohexanone.

Figure 74: Crystal forms from 96-well optimisations of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and either cyclohexylamine or 2-aminocyclohexanone.

Figure 75: Secondary structure visualisation of cyclohexylamine bound pETYSBLIC-3C MATOUAmDH2.

Figure 76: Structure and binding pocket of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and cyclohexylamine.

Figure 77: Omit maps for both NADP⁺ and cyclohexylamine in respect to the pETYSBLIC-3C MATOUAmDH2 structure.

Figure 78: In-depth analysis of the substrate cage and active site residues in MATOUAmDH2.

4.3 Engineering MATOUAmDH2 for biocatalysis

Figure 79: PCR product bands from an exemplar L144A thermogradient trial using both types of primers.

Figure 80: PCR product bands from MATOUAmDH2 alanine screen site directed mutagenesis, analysed via gel electrophoresis.

Figure 81: Nickel Affinity Chromatogram for MATOUAmDH2 alanine mutants expressed in the pETYSBLIC-3C vector.

Figure 82: Size Exclusion Chromatogram for MATOUAmDH2 alanine mutants expressed in the pETYSBLIC-3C vector.

Figure 83: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C MATOUAmDH2 F143A, L144A and L169A protein purification.

Figure 84: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C MATOUAmDH2 L180A, M215A and T312A protein purification.

Figure 85: Kinetic plot for pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone.

Figure 86: Biotransformations using pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone.

Figure 87: PCR product bands from MATOUAmDH2 site directed mutagenesis at L180 and M215 positions, analysed via gel electrophoresis.

Figure 88: SDS-PAGE analysis of L180 mutant lysates.

Figure 89: SDS-PAGE analysis of M215 mutant lysates.

Figure 90: Biotransformations using pETYSBLIC-3C MATOUAmDH2 L180 and M215 mutants.

Figure 91: Conversions of a range of ketones using pETYSBLIC-3C MATOUAmDH2 L180 and M215 mutants.

Figure 92: Nickel Affinity Chromatogram for M215S MATOUAmDH2 expressed in the pETYSBLIC-3C vector.

Figure 93: Size Exclusion Chromatogram for M215S MATOUAmDH2 expressed in the pETYSBLIC-3C vector.

Figure 94: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C M215S MATOUAmDH2 protein purification.

Figure 95: Crystal forms from M215S MATOUAmDH2:pETYSBLIC-3C in complex with NADP⁺ and 2-aminonorbornane.

Figure 96: Dimer of M215S bound to 2-aminonorbornane and NADP⁺.

Figure 97: Omit maps for both NADP⁺ and 2-aminonorbornane bound to M215S.

Figure 98: GC-FID traces of substrate and product peaks for the 1,2-cyclohexanedione:ammonia based reactions monitored over 24 h.

Figure 99: Underestimated biotransformations calculated from the combined peak of 1,2-cyclohexanedione in reactions using ammonia and M215 mutants.

Figure 100: K43 orientations between chains of M215S and WT MATOUAmDH2.

Figure 101: Binding pocket of M215S MATOUAmDH2 in complex with NADP⁺ and 2-aminonorbornane.

Figure 102: Proposed reductive amination mechanism of norcamphor in the active site of MATOUAmDH2.

4.4 Other nat-AmDHs of interest

Figure 103: Crystal forms from *CfusAmDH* W145A in complex with NADP⁺ or NAD⁺.

Figure 104: Monomeric forms of *CfusAmDH* W145A bound to either NAD⁺ or NADP⁺ and *n*-pentylamine.

Figure 105: Biological assembly forms of *CfusAmDH* W145A bound to either NAD⁺ or NADP⁺ and *n*-pentylamine.

Figure 106: Omit maps for the cofactors and ligands of both *CfusAmDH* W145A structures.

Figure 107: Nickel Affinity Chromatogram for *TtherAmDH* expressed in the pET22b(+) vector.

Figure 108: Size Exclusion Chromatogram for *TtherAmDH* expressed in the pET22b(+) vector.

Figure 109: SDS-PAGE analysis of protein fractions from pET22b(+) *TtherAmDH* protein purification.

Figure 110: Crystal forms from *Tther*AmDH in complex with NADP⁺.

Figure 111: Biological assembly and domain assignment of apo-*Tther*AmDH

Figure 112: Nickel Affinity Chromatogram for *Micro*AmDH expressed in the pET22b(+) vector

Figure 113: Size Exclusion Chromatogram for *Micro*AmDH expressed in the pET22b(+) vector.

Figure 114: SDS-PAGE analysis of protein fractions from pET22b(+) *Micro*AmDH protein purification.

Figure 115: Crystal forms from *Micro*AmDH in complex with NADP⁺ and pentylamine/hexylamine.

Figure 116: A structural comparison of *Tther*AmDH and AmDH4.

Figure 117: A structural comparison of *Cfus*AmDH wild-type and W145A structures.

Figure 118: Cofactor stabilisation and interactions of *Cfus*AmDH wild-type and W145A structures.

Appendices

Figure 119a: Chemical structure of NAD⁺ and NADH.

Figure 119b: Chemical structure of NADP⁺ and NADPH.

Figure 120: Initial kinetic gradients for MATOUAmDH2 in the pET22b(+) vector.

Figure 121: Initial kinetic gradients for MATOUAmDH2 alanine screen mutants

Figure 122: GC-FID traces for pET22b(+) cyclohexanone and cyclohexylamine standards.

Figure 123: GC-FID traces for pET22b(+) cyclohexanone and *N*-methylcyclohexylamine standards.

Figure 124: GC-FID traces for pETYSBLIC-3C cyclohexanone and cyclohexylamine standards.

Figure 125: GC-FID traces for pETYSBLIC-3C cyclohexanone and *N*-methylcyclohexylamine standards.

Figure 126: GC-FID traces for 1,2-cyclohexanedione and 2-aminocyclohexanone standards.

Figure 127: GC-FID traces for norcamphor and 2-aminonorbornane standards.

Figure 128: GC-FID traces for 2-teralone and 2-aminotetralin standards.

Figure 129: GC-FID traces for 2-pentanone and 2-aminopentane standards.

Figure 130: GC-FID traces for 2-heptanone and 2-aminoheptane standards.

Figure 131: Sequence alignment of MATOUAmDH2 and *Cfus*AmDH.

Figure 132: Sequence alignment of MATOUAmDH2 and *Msme*AmDH.

Figure 131: Sequence alignment of *Tther*AmDH and AmDH4.

List of Tables

3.1 Molecular Biology

Table 1 – PCR components for the amplification of all MATOUAmDH2 based inserts

Table 2 – PCR components for the amplification and linearisation of the pETYSBLIC-3C vector

Table 3 – PCR cycling conditions for the MATOUAmDH2 insert

Table 4 – PCR cycling conditions for the pETYSBLIC-3C vector

Table 5 – PCR cycling conditions for MATOUAmDH2 mutants

3.2 Protein Production and Purification

Table 6 – Ingredients for 10 mL of 2X SDS Loading Dye

Table 7 – Ingredients for 1 L of ready to use Running Buffer

Table 8 – Ingredients to make 2x 12% SDS gels

4.1 Characterisation of pET22b(+) MATOUAmDH2

Table 9 – Kinetic parameters for pET22b(+) MATOUAmDH2 with cyclohexanone

Table 10 – Data collection and refinement statistics for the 2.3 Å pET22b(+) MATOUAmDH2 in complex with NADP⁺.

4.2 Co-crystallisation of MATOUAmDH2 with cyclohexylamine

Table 11 – Data collection and refinement statistics for the 2.08 Å pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and cyclohexylamine.

4.3 Engineering MATOUAmDH2 for biocatalysis

Table 12 – Kinetic parameters for pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone

Table 13 – Data collection and refinement statistics for the 1.61 Å MATOUAmDH2 M215S in complex with NADP⁺ and 2-aminonorbornane.

4.4 Other nat-AmDHs of interest

Table 14 – Data collection and refinement statistics for the 1.64 Å *Cfus*AmDH W145A structure in complex with NAD⁺.

Table 15 – Data collection and refinement statistics for the 1.50 Å *Cfus*AmDH W145A structure in complex with NADP⁺ and *n*-pentylamine.

Table 16 – Data collection and refinement statistics for the 2.71 Å apo-*Tther*AmDH structure.

Appendices

Table 17 – List of Primers for pETYSBLIC-3C subcloning and mutagenesis

Table 18 – Protein sequences of nat-AmDHs

Table 19 – Retention times of standards analysed by GC-FID

Acknowledgments

To start I would like to thank my supervisor, Professor Gideon Grogan for the opportunity to carry out my PhD. For the last three years I have benefitted immensely from his advice, direction, and encouragement of which have all shaped me in becoming the scientist I am today. My experiences and opportunities during my time in the group has been irreplaceable, something I will forever be grateful for. The Grogan group in general have always been lovely, always offering advice or helpful anecdotes when needed. I will forever cherish the memories and friendships we've all made together in the last three years. Particularly those who have always listened to my ranting, either in the office, round our own dinner tables or at many pubs – thank you for the endless support and care.

My time wouldn't have been half as special if it wasn't for the rest of the YSBL team members. The community here is like no other; thank you all for broadening my scientific interests and the fruitful discussions, science related or otherwise. Especially thanks to the YSBL technicians, Louise, Simon, Jules and Tim, without you all YSBL wouldn't run nearly as smoothly. Thanks to Mr Sam Hart and Dr Johan P. Turkenburg for assistance with X-ray data collection which made up a large part of my last three years. To Laurine Ducrot and Dr Carine Vergne-Vaxelaire, our collaboration has been invaluable, a lot of this work wouldn't have been possible without you both – thank you.

I want to say a huge thank you to my family and best friends back in London. Without them I'm not sure the last three years would have been as enjoyable. Thank you for always being a phone call or train away. I am, and have always been, so grateful to have such amazing people in my life. Thank you for always providing a de-stress weekend when needed and listening to me constantly talk about science! To my parents Vicky and Andy and grandparents Julie and Ron, thank you for always encouraging me to pursue my dreams and aspirations – without you all I wouldn't have achieved anything I have today, I am forever indebted.

Lastly, Theo, thank you for reminding me where I am heading and never failing to motivate me to get there. You've always instilled my self-assurance and you never fail to reassure me when times get a bit tough. Thank you for making every day so much easier and never, ever, failing to make me smile. Also thank you for always politely and inquisitively answering my ridiculous chemistry questions!

Author's Declaration

I declare that this thesis is a presentation of original work, unless stated otherwise, this work has been performed by myself. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Laurine Ducrot and Dr Carine Vergne-Vaxelaire at Genoscope provided the following items:

- MATOUAmDH2 in the pET22b(+) plasmid
- *MicroAmDH* in the pET22b(+) plasmid
- *TtherAmDH* in the pET22b(+) plasmid
- A purified sample of *CfusAmDH* W145A protein
- 2-aminocyclohexanone hydrochloride

Balázs Pogrányi synthesised the following:

- 2-aminotetralin

The work reported herein or parts of, have been published in the following Journals and are available for accessing:

L. Ducrot, M. Bennett, G. Grogan and C. Vergne-Vaxelaire, *Adv. Synth. Catal.*, 2020, **363**, 328–351.

L. Ducrot, M. Bennett, A. A. Caparco, J. A. Champion, A. S. Bommarius, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire, *Front. Catal.*, 2021, **1**, DOI:10.3389/fctls.2021.781284.

M. Bennett, L. Ducrot, C. Vergne-Vaxelaire and G. Grogan, *ChemBioChem*, 2022, **23**, 6–11.

L. Ducrot, M. Bennett, G. André-Leroux, E. Elisée, S. Marynberg, A. Fossey-Jouenne, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire, *ChemCatChem*, 2022, DOI:10.1002/cctc.202200880.

1. Introduction

1.1 Synthesis and application of chiral amines

Chiral amines are amine functional groups in which the carbon atom, directly attached to the nitrogen, is the site of chirality. Chiral amines are important functionalities, routinely being used in small molecule pharmaceuticals, food chemistry and agrochemical industries. It is approximated that more than 40% of small molecule drugs contain one or more chiral amines with defined chirality being a desirable of around 80% of new pipeline pharmaceuticals.¹ The chemical and biological significance of these chiral amine-based pharmaceuticals are extensive; often used in the treatment of numerous diseases and conditions.²⁻⁴ For example, Rivastigmine **1** is used to treat early onset Alzheimer's disease and Nubeqa® **2** and Brukinsa® **3** are used in the treatment of various cancers and lymphomas to more well-known drugs such as the antibiotic Amoxicillin **4**. Figure 1 displays the applications when using chiral amines as building blocks in various chemical industries.

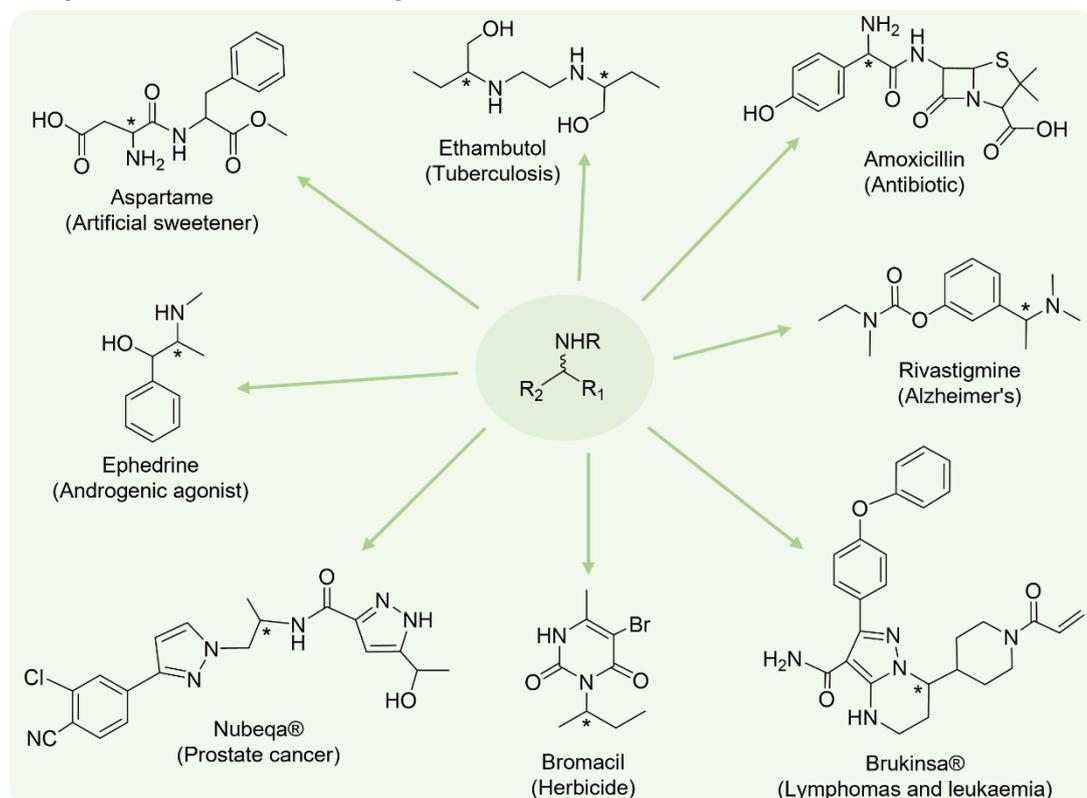
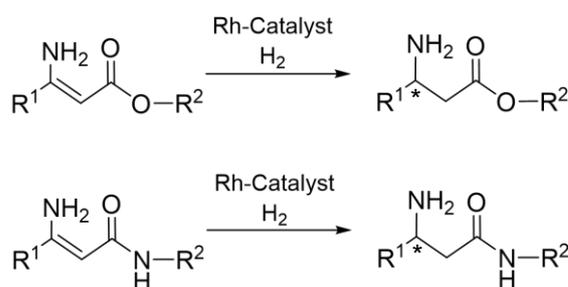


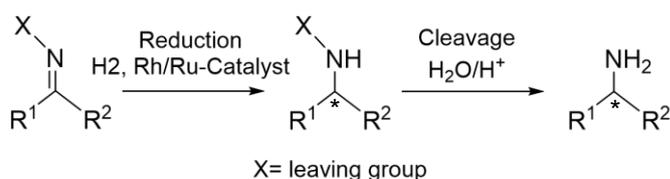
Figure 1: Various uses of chiral amines in pharmaceutical, food and agricultural chemistry. This figure displays a few uses of chiral amines as building blocks for chemical synthesis of important compounds. All compounds are named accordingly with references to their uses being bracketed. Chiral centres of interest are marked with an asterisk.

Traditionally, synthesis of chiral amines relies on the asymmetric reduction of heavily functionalised substituents. For example, the catalytic hydrogenation of enamines⁵, reduction of C=N bonds^{1,6,7} and the use of intermolecular C-H insertion.⁸ In the case of an enamine hydrogenation, typically a β -enamine ester or amide, prepared from the β -keto acid or amide precursor, is reduced to the resulting amine by employing a rhodium based catalyst (Figure 2).⁵ In the case of these reactions specific functionality is required if no leaving group is present. For C=N bond reduction typically the imine precursor is reduced by catalytic hydrogenation facilitated by a transition metal catalyst, such as a rhodium or ruthenium-based catalyst, and the resulting amine is formed (Figure 2).^{6,7} For these types of reactions in general the synthesis of the imine precursor and cleavage of the leaving group are the bottlenecks. Intermolecular C-H insertion requires an oxidising agent, catalytic hydrogen, and a heavy metal catalyst such as rhodium-based tether such as $\text{Rh}_2(\text{esp})_2$. An amine source is required to generate formation of a new C-H bond (Figure 2).⁸

Reduction of β -enamine esters and amides



C=N bond reduction and cleavage



Intermolecular C-H amination

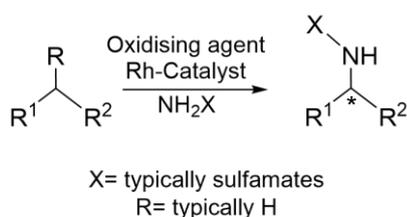


Figure 2: Chemical synthesis of chiral amines from functionalised centres. This schematic displays the generalised chemical synthesis of chiral amines from either enamines, imines, or intermolecular C-H insertion. Groups of interest are highlighted in the legends. R^1 and R^2 reflect various substituted groups. Potential chiral centres of interest are marked with an asterisk.

1.1.1 Reductive amination

Reductive amination is achieved by the coupling of a carbonyl centre with an amine centre to form a amine product (Figure 3).⁹ It is thought that currently >25% of C-N bond synthesis relies on the reductive amination of carbonyl centres.¹⁰ This reaction is extremely important in the pharmaceuticals industry; often the synthesis of small molecule chiral amine drugs require at least one or more reductive amination step. Initially, the hemiaminal, if formed from a secondary amine, can, through condensation, form an iminium ion intermediate (Figure 3).⁹ In the case of primary amines, the hemiaminal typically forms an imine intermediate (Figure 3).⁹ These intermediates can be further reduced to the resulting amine or in some cases the reduced amine can be produced directly from the hemiaminal (Figure 3).⁹ Typically the reductive amination step is achieved by using reducing borohydride complexes such as NaBH(OAc)₃ or NaBH₄ and/or achieved through hydrogenation with hydrogen gas or catalytic hydrogen coupled to various catalysts including transition metal-based agents such as those containing palladium or titanium centres.¹¹⁻²² These reactions often provide good yields and enantiomeric excesses; but this is largely influenced by the steric effects of the R groups on both the carbonyl and amine donor. Frequently the amine donors required are complex and require extensive leaving groups. Often hydrogenolysis, again typically using transition metal catalysts, is required to produce the primary amine by cleaving the leaving groups (Figure 3).^{9,23-25} This also requires the separation of the minor and major products for hydrogenolysis to produce the desired enantiomer, as is the case with racemic resolution that does not precede hydrogenolysis (*Introduction 1.1.2*).^{1,9,26,27}

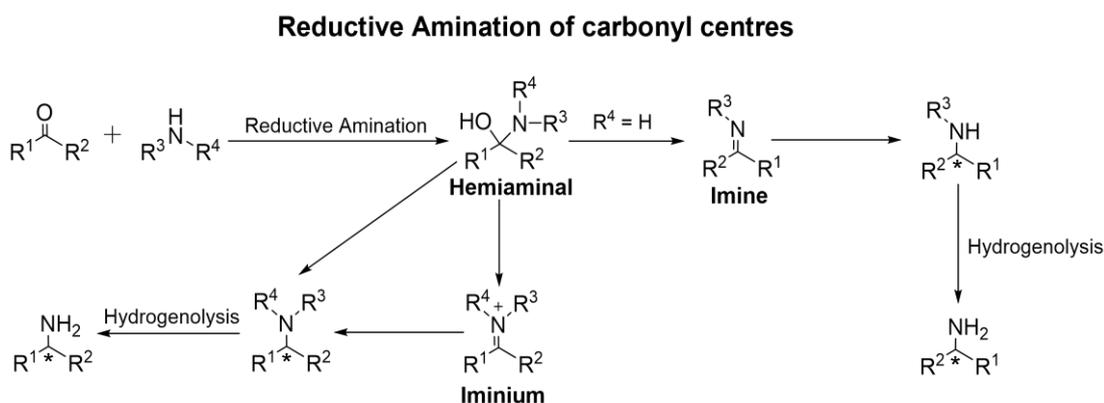


Figure 3: Chemical synthesis of chiral amines from unfunctionalised carbonyl centres. This schematic displays the generalised synthesis of chiral secondary and tertiary amines from unfunctionalised carbonyls through hemiaminal formation. R¹, R² and R³ reflect various substituted groups where R⁴ can be a substituent or a hydrogen atom (to form secondary amines via imine reduction). Potential hydrolysis is also highlighted for the ultimate production of primary (chiral) amines. Potential chiral centres of interest are marked with an asterisk.

Rivastigmine (**1**), as mentioned previously (*Introduction 1.1*), is used in the treatment of Alzheimer's disease.²⁸ Rivastigmine can be chemical synthesised by various means of reductive amination, of which include the use of reducing agents such as NaBH₄ and transition metal catalysts^{11,12,15} (Figure 4) or catalytic hydrogenation using transition metal catalysts and hydrogen gas (Figure 5).^{16,17} The initial reductive amination step in the synthesis of **1**, adapted from Interquim and Zentiva, required 3-acetylanisole **1a** and either a methylamine or dimethylamine in the presence of NaBH₄ and a Ti(OiPr)₄ catalyst followed by a demethylation to 3-[1-(methylamino)ethyl]phenol **1.1b** or 3-[1-(dimethylamino)ethyl]phenol **1.2b** respectively (Figure 4).^{11,12,15} The racemic separation of **1.1b** and **1.2b** is carried out using chiral resolution via crystallisation (*Introduction 1.1.2*) to (*S*)-enantiomers **1.1c** and **1.2c** respectively. Subsequently **1.1c** or **1.2c** undergo acylation with carbamoyl chloride either yield (*S*)-**1d** or (*S*)-**1** respectively, (*S*)-**1d** can then be used in reductive amination of formaldehyde in the presence of NaBH₃CN to produce (*S*)-**1**. The final yields achieved were 75% for methylamine-based reaction and 80% for the dimethylamine-based reaction.^{11,12,15}

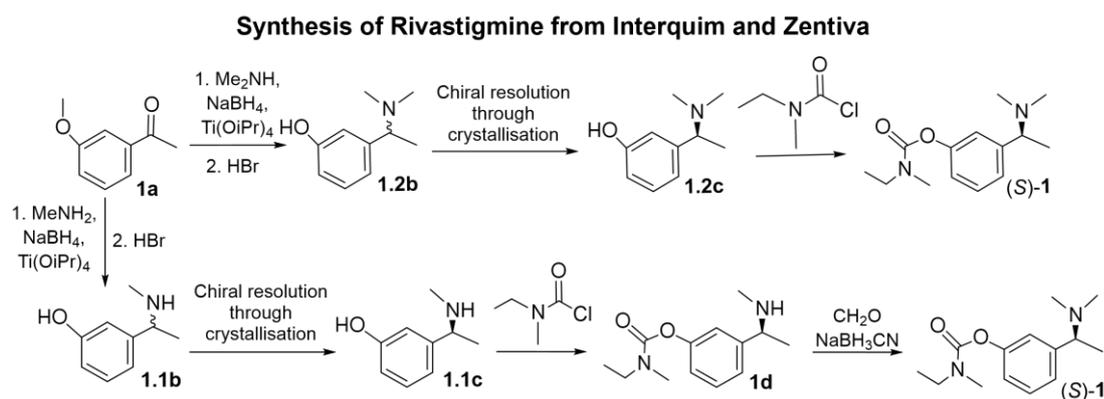


Figure 4: Synthesis of rivastigmine adapted from Interquim and Zentiva. This schematic displays the two methods of the synthesis of (*S*)-rivastigmine from 3-acetylanisole using either methylamine or dimethylamine to form **1.1b** and **1.2b** respectively.

Subsequently, a synthetic method was developed by Zhejiang Hisun Pharmaceutical that did not require the racemic separation of intermediates (as is previously presented in Figure 4).^{16,17} In this method the reductive amination step required **1a** in the presence of chiral phenylethylamine, hydrogen gas and a Ti(OiPr)₄ catalyst to form **1e** (Figure 5).^{16,17} **1e** could then be used in the methylation and demethylation steps to ultimately yield **1f** which undergoes acylation with carbamoyl chloride to form **1g**. **1g** undergoes hydrogenolysis catalysed by a palladium-based catalyst to form **1d** which in turn can be methylated to ultimately form (*S*)-**1**.^{16,17} The overall yields for this reaction were reported to be between 57-61%.^{16,17}

Synthesis of Rivastigmine from Zhejiang Hisun Pharmaceutical

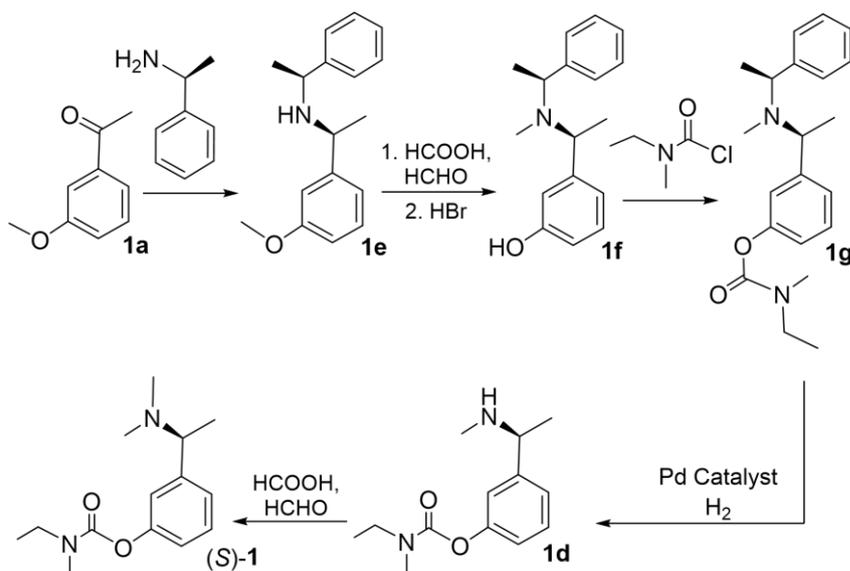


Figure 5: Synthesis of rivastigmine adapted from Zhejiang Hisun Pharmaceutical. This schematic displays the synthesis of (S)-rivastigmine from 3-acetylanisole using chiral phenylethylamine as designed by Zhejiang Hisun Pharmaceutical.

Another example of reductive amination used in the pharmaceutical industry is seen during the synthesis of oseltamivir **5**. Oseltamivir, more commonly known as Tamiflu, is routinely used as an antiviral drug during the treatment of influenza viruses. Some adapted exemplar reductive amination procedures to produce oseltamivir include the use of downstream reductive agents NaBH₄ and/or transition metal (co)catalysts such as MoO₃, Rh₂O₃ and Al₂O₃ (Figure 6 and 7).^{18–20} Synthesis of **5** typically starts from derived benzoic acid ethyl esters such as ethylbenzoate **5a** or **5e** as in Figure 6 and 7 respectively.^{18–20} In the case of the ethylbenzoate-based reaction, **5a** reacts in a five step procedure including the use of both enzymatic and transition-metal based catalysis steps, to produce ketone intermediate **5b** (Figure 6).¹⁸ **5b** then subsequently undergoes reductive amination using hydroxylamine hydrochloride to produce the imine intermediate **5c** which is further reduced using hydrogen gas coupled with either Rh₂O₃ or Al₂O₃ based catalysts (Figure 6).¹⁸ The resulting amine **5d** can be converted to (6S)-**5** by further coupling reactions including the use of 3-pentanol and a deprotection step.¹⁸ The overall yield from this reported reaction was recorded as 60%.¹⁸ As mentioned, alternative methods for the synthesis of oseltamivir have been reported.^{19,20} The reductive amination, again using hydroxylamine, of a ketone intermediate **5f** (generated from **5e**) to form imine **5g**, could be achieved (Figure 7).^{19,20} The resulting imine **5g** can be further reduced by NaBH₄ and a MoO₃ catalyst to form amine **5h** (Figure 7). **5h** can be used in downstream reactions usually requiring

3 or 4 steps to produce the resulting compound (6S)-5.^{19,20} The overall yields for these reported reactions were much lower at around 2.9-5%.^{19,20}

Synthesis of Oseltamivir from ethylbenzoate

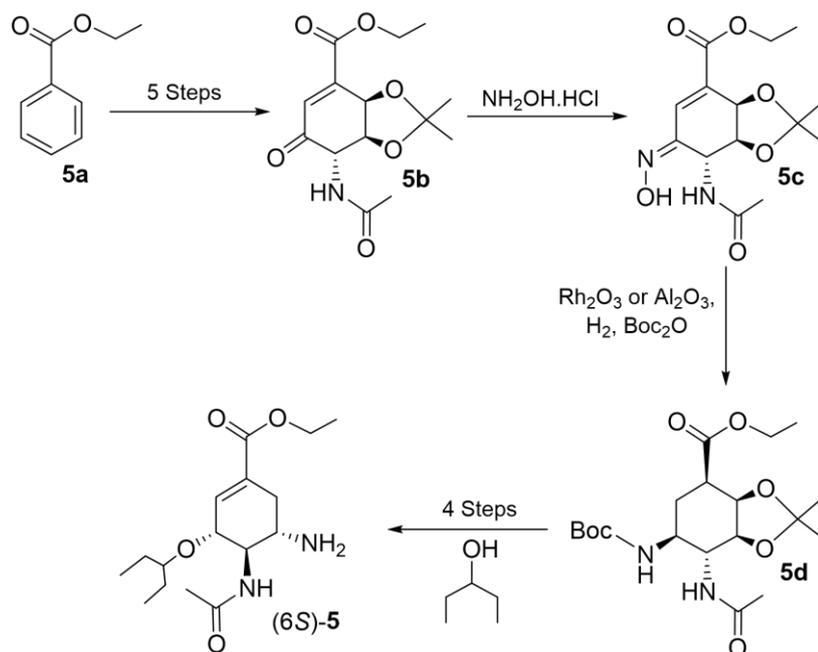


Figure 6: Synthesis of oseltamivir from ethylbenzoate. This schematic displays the synthesis of (6S)-oseltamivir from ethylbenzoate as designed by Hudlicky *et al.*

Synthesis of Oseltamivir from 5e

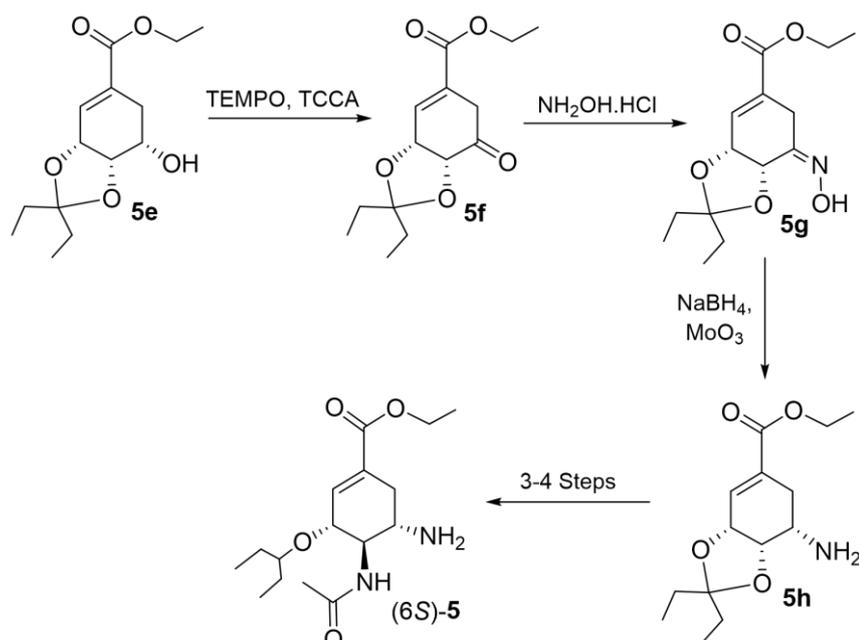


Figure 7: Synthesis of oseltamivir from 5e. This schematic displays the synthesis of (6S)-oseltamivir from 5e as designed by Kongkathip *et al.*

A further example of how reductive amination is employed during the synthesis of small molecule pharmaceuticals concerns the production of maraviroc **6**. Maraviroc is an antiviral used in the treatment of retroviruses such as the human immunodeficiency virus (HIV).²⁹ Ultimately the protocol for the reductive amination part of the multi-step maraviroc synthesis has been reduced to two steps.^{21,22} The reductive amination part is characterised by the reduction of aldehyde **6a** using an amine tosylate **6b** and reducing agent NaBH(OAc)₃ and then quenched to produce **6** (Figure 8).^{21,22}

Synthesis of Maraviroc

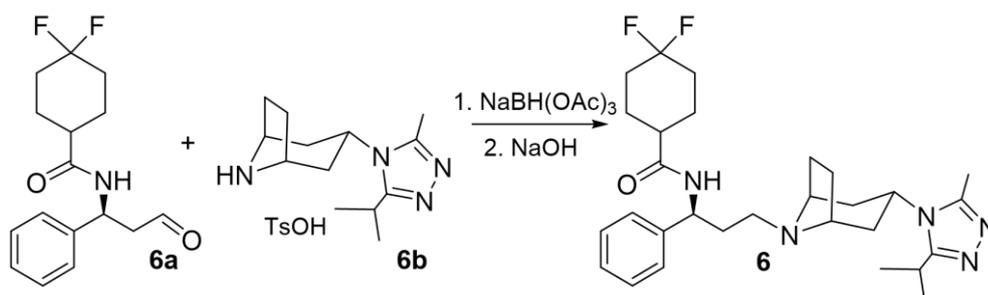


Figure 8: The reductive amination step in the synthesis of maraviroc. This schematic displays the synthesis of maraviroc from **6a** and **6b** as designed by Córdova et al. *N.B.* the chirality at the *exo* centre is not displayed.

Reviewing the synthesis of sitagliptin **7** further highlights the use of complex transition metal catalysts in the synthesis of pharmaceutical chiral amines. Sitagliptin is routinely used in the treatment of type 2 diabetes.^{30,31} Sitagliptin acts as a small molecule inhibitor of the serine protease dipeptidyl peptidase 4 (DPP-4).^{30,31} *In vivo*, DPP-4 is responsible for degrading glucagon-like peptide 1 (GLP-1); an incretin hormone that is released in response to food intake and has roles in stimulating insulin biosynthesis coupled with inhibition of glucagon release.^{30,31} Therefore, administration of sitagliptin would delay the degradation of GLP-1 and subsequently allow for longer lasting effects resulting in more insulin production and reductions in glucagon levels. Like most inhibitors that target protein sites, the enantiomeric properties of sitagliptin dictate its mode of action. (*R*)-sitagliptin is the required enantiomer to elucidate its desired pharmacodynamic and pharmacokinetic properties *in vivo*.^{30,31} This is because (*R*)-sitagliptin is stabilised by a hydrogen bonding and ionic interactions network at many contact points within the proteins binding pocket, namely canonical glutamate, tyrosine, serine, and arginine residues which tend to be conserved across many DPP-4 proteins from various host organisms (Figure 9).^{30,32} Thus, an enantioselective route during the synthesis of (*R*)-sitagliptin is required, often requiring the use of complex chiral transition metal catalysts.

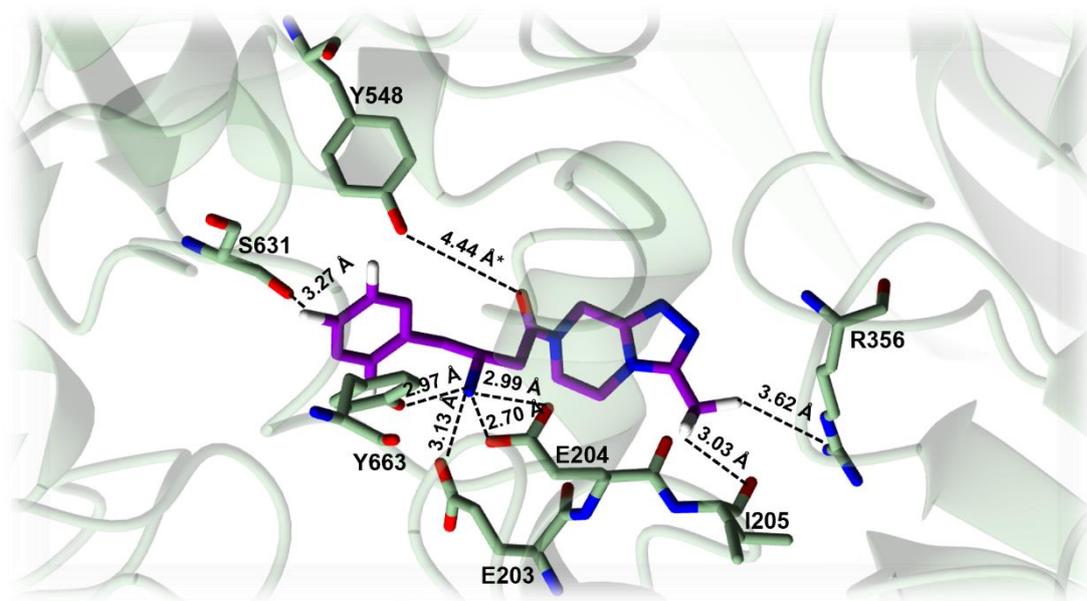


Figure 9: The active site of DDP-4 in complex with (*R*)-sitagliptin. All images have been visualised using CCP4mg (v.2.10.11). This image represents the active site residues in DDP-4 from *Rattus norvegicus* (green) (PDB: 4FFW) in complex with (*R*)-sitagliptin (purple). All residues are labelled accordingly. The dotted line indicates the distances between interactions which are measured in Ångstroms. In the human model of DDP-4 (PDB: 1X70) I205 is replaced by S209 which is functionally conserved and interchangeable. *In the human model this hydrogen bonding interaction is mediated via a water molecule; there is no water present here in the *R. norvegicus* structure, however it is only a 2.90 Å structure.

Initially, Merck provided a synthetic one pot synthesis for (*R*)-**7** (Figure 10).¹³ By combining (2,4,5-Trifluorophenyl)acetic acid **7a**, and 2,2-Dimethyl-1,3-dioxane-4,6-dione (Meldrum's acid) **7b**, and then reacting the intermediate with a trifluoro based amine salt the resulting ketoamide **7c** was formed.¹³ **7c** can then undergo condensation using ammonium acetate to form the enamine **7d**. **7d** subsequently is reduced to the resulting chiral amine (*R*)-**7** using hydrogen gas, catalysed by the chiral rhodium-based catalyst and *t*-Bu-Josiphos ligand.¹³ The reported yield from this reaction was 82% with an e.e. of 99%.¹³ This method was also adapted by Steinhuebel *et al* whereby using the ketoamide **7c**, hydrogen gas and a ruthenium-based catalyst with the (*R*)-dm-segphos® ligand a yield of 91% and 99.5% e.e. for (*R*)-**7** could be achieved (Figure 10).¹⁴

Synthesis of Sitagliptin

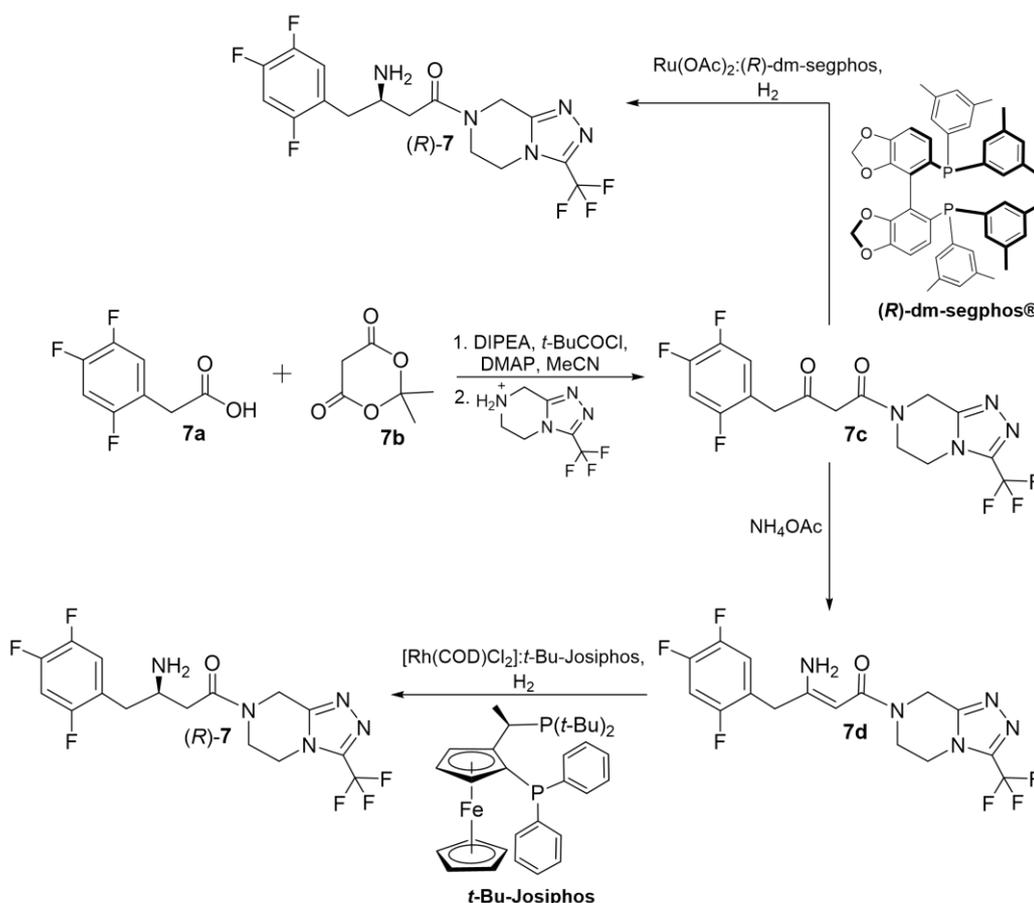


Figure 10: The synthesis of (R)-sitagliptin. This schematic displays the reductive amination steps in the synthesis of (R)-sitagliptin requiring the use of rhodium and ruthenium-based co-ligand catalysts designed by Merck and Steinhuebel *et al* respectively.

These chemical syntheses as described in *Introduction 1.1 and 1.1.1* all pose problems. In the case of the synthesis from enamines the resulting product and substrates are highly functionalized, limiting the scope of potential products formed.⁵ In the case of C=N bond reduction the imine formation is the limiting step, often requiring long incubation and distillation times.^{1,6,7} Also, the leaving group needs to be optimised in such a way that it does not affect the reduction of the imine through steric effects to give undesirable enantiomers.^{1,6,7} In the case of C-H amination through intermolecular insertion typically the resulting reactions give poor yields and enantiomeric excesses.⁸ Additionally, the amine donors tend to be highly functionalised i.e., use of sulfamates, limiting the potential product scope.⁸ The formation of amines by the reductive amination of unfunctionalised ketones overcomes the potential substrate and product scope problem. However, the remaining issues for all syntheses described above seems to be in the resolution of enantiomers either before or after the formation of the final product(s).

1.1.2 Racemic resolution

Defined chirality, especially in the pharmaceutical industry is extremely important; this remains the case with chiral amine containing drugs. Some examples noted above were during the synthesis of (*S*)-rivastigmine **1** and (*R*)-sitagliptin **7**.^{11–17} One of the most well-known examples was seen when the drug thalidomide **8** was prescribed to pregnant women as a remedy for morning sickness. Thalidomide works by inhibiting TNF- α production by degrading TNF- α mRNA, impacting many cytokine-regulated physiologies such as inflammations, oxidative stresses, and fevers.³³ Whilst both enantiomers are potent TNF- α inhibitors (*S*)-thalidomide **8.1** was teratogenic and could be interconverted from the (*R*)-enantiomer **8.2** *in vivo*, impacting the development of the embryo or fetus.^{34–37} Because thalidomide was prescribed in a racemic mixture many children were born with limb and/or brain abnormalities. Hence the importance of determining, and the ability to, isolate pure enantiomeric products. There are currently several methods for the resolution of chiral amine enantiomers. The most routinely used methods are either through chiral resolution or kinetic resolution. For chiral resolution, fractional crystallisation of the racemic amine solution with a chiral carboxylate forms a crystalline material trapping one enantiomer whilst the other is left in solution (Figure 11).¹ The stereochemistry of the crystalline enantiomer is determined by the initial stereochemistry of the carboxylic acid salt.

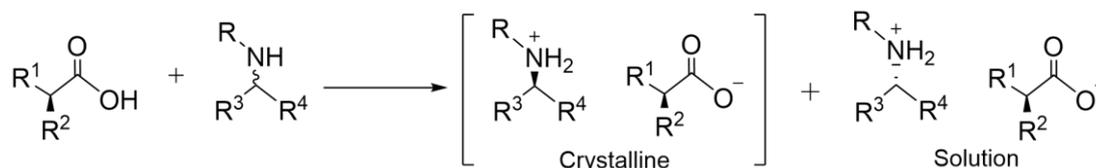
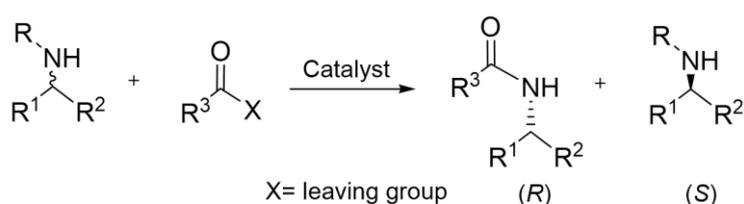


Figure 11: Chiral resolution of racemic amines via fractional crystallisation. This schematic displays the generalised resolution of a racemic amine using chiral carboxylates to trap one enantiomer in a crystalline phase. R₁, R₂, R₃ and R₄ reflect various substituted groups where R can either be H or a substituted group. N.B. The stereochemistry of the crystalline amine is determined by the stereochemistry of the initial carboxylic acid salt used.

In the case of kinetic resolution, more traditional synthesis steps are used, usually an acylating agent will be utilised. A racemic mixture of the amine can be incubated with an acylating agent and a catalyst where the (*R*)-enantiomer undergoes acylation to form the amide and the (*S*)-enantiomer of the amine is released (Figure 12).^{1,26} Subsequent work has been focused on using biocatalysts for the dynamic kinetic resolution of chiral amines. By utilising enzymes native abilities for stereoselectivity and promiscuity a range of enantiomerically pure amines have been synthesised. By pairing enzymes such as lipases and acylating agents, a one-step synthesis for the separation of both enantiomers can be achieved by coupling the acylation and then

hydrolysis of the amide bond (Figure 12).²⁷ Natively these lipases break down ester bonds of triglycerides *in vivo*, this requires the formation of a typical enzyme-acyl intermediate with resulting products being released through nucleophilic attack by water.³⁸ However, these lipases have been engineered to accept a wide range of nucleophiles such as alcohols or amines to hydrolyse ester or amide bonds.^{27,38} As the reaction is reversible lipases can be utilised (in the 'reverse reaction') to accept racemic mixtures of amines and acylating agents to form the (*R*)-acylation product and (*S*)-amine; the (*R*)-acylation product can then be further hydrolysed to give the (*R*)-amine (Figure 12).^{27,38}

Acylation of racemic amines



Acylation and hydrolysis of racemic amines using lipases

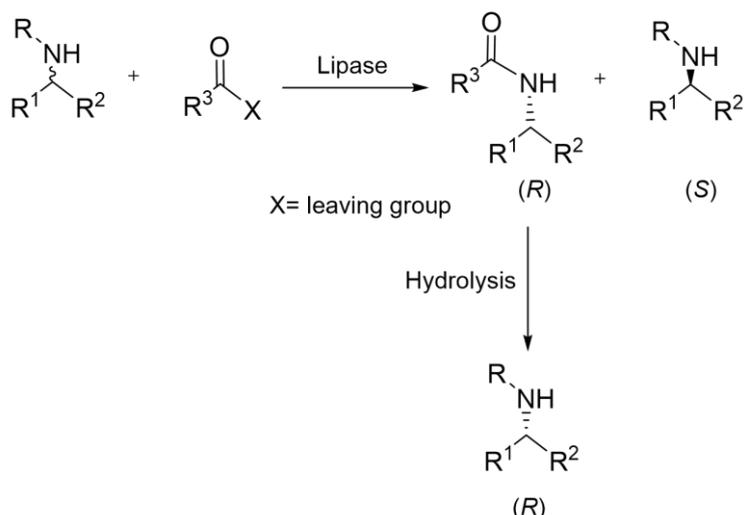


Figure 12: Kinetic resolution of racemic amines via acylation-based reactions. This schematic displays the generalised resolution of a racemic amine using acylating agents and organic or enzymatic catalysts, such as lipases, for the formation of (*R*)-amides and (*S*)-amine products. R¹, R², R³ and R⁴ reflect various substituted groups where R can either be H or a substituted group. X refers to the leaving group on the acylating agent as denoted in the legend.

In addition to the problems faced in the resolution of enantiomers, the traditional chemical syntheses of chiral amines also raise a problem concerning the use of transition metal catalysts. Often these catalysts employ expensive and non-renewable transition metals such as ruthenium, rhodium, palladium, iridium, molybdenum, and titanium. Additionally, the ligands that coordinate these transition metal centres are

often expensive and complex, sometimes requiring multi-step syntheses themselves, for example the use of $\text{Rh}_2(\text{esp})_2^8$, (*R*)-dm-segphos[®]¹⁴, *t*-Bu-Josiphos¹³, BINAP, and cyclooctadiene based ligands.^{39,23} Often these ligand-metal catalysts can degrade during the reaction progression, show low efficiency, cannot be recovered and/or aren't tailored for controllable selectivity.²³ One way to overcome these problems is to immobilise the chiral catalyst to polymers offering improved enantioselectivity and recovery with no compromise in activity.⁴⁰ However, designing a ligand tether proposes a new challenge in this field.

1.2 Biocatalysts for the production of chiral amines

Due to the limitations in more traditional chemical syntheses concerning the production of chiral amines recent attention has turned to finding enzymes capable of creating chiral amines. Enzymes as biocatalysts tend to offer improved enantioselectivity and substrate specificity over transition metal-based catalysts and organocatalysts.^{41,42} In addition to this, enzymes can be engineered to improve substrate scopes and promiscuities offering the ability to synthesise a wide range of chiral amine functionalities.^{41,42} If libraries express well in their host organism(s), and can be scaled to model expression organism(s), it also provides a relatively cost-efficient route compared to the synthesis of transition metal-based catalysts. Furthermore, enzymes offer a more renewable and sustainable route typically with lower reaction temperatures and pressures and less harsh conditions.^{41,42} The idea of using biocatalysts in chemical industries is not a new concept. Many pharmaceutical molecules target host proteins by methods of action, for example sitagliptin (*Introduction 1.1.1*).³⁰⁻³² The natural propensity of target molecules to bind to their target protein(s), coupled with an enzyme's natural substrate and catalytic promiscuity, is often exploited to find enzymes capable of binding synthetic intermediates. This is widely seen in the administration of pro-drugs which are metabolised *in vivo* into the active form, often by enzymes such as cytochrome P450s.⁴³⁻⁴⁵ Often target molecules are engineered to mimic natural products of enzymes which highlights the importance placed on enzymes as biocatalysts in synthetic processes.⁴⁶⁻⁴⁸ However, enzymes are often screened and engineered to accept and synthesise new and varying chemical functionalities.^{41,42} Discussed herein are a selection of biocatalysts capable of producing chiral amines and their associated structural and mechanistic features are detailed.

1.2.1 Transaminases

Traditionally transaminases (TAs) *in vivo* are used in amino acid recycling and degradation. Native TAs accept functionalised α -amino acids transferring the amino group from a donor α -amino acid to an acceptor α -keto acid known as transamination (Figure 13).⁴⁹ ω -TAs are a sub-group of TAs which catalyse the transamination of a donor and acceptor molecule where at least one of the counterparts do not need to be an α -amino acid or α -keto acid (Figure 13).^{50,51} Therefore, ω -TAs can catalyse transamination of a range of carbonyl acceptor groups and produce a wider range of chiral amines not limited to the production of α -amino acids.

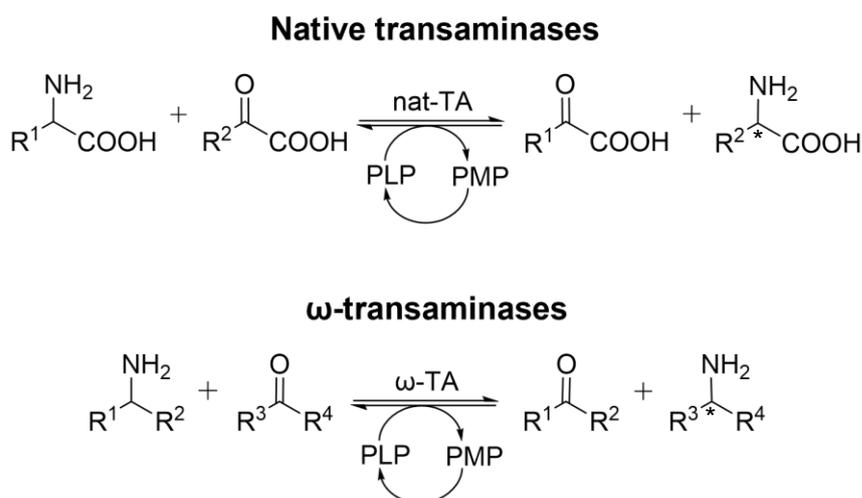


Figure 13: PLP-dependent reactions catalysed by transaminases. This schematic displays the generalised reactions of native TAs and ω -TAs using the PLP cofactor. R^1 , R^2 , R^3 and R^4 reflect various substituted groups. Potential chiral centres of interest are marked with an asterisk.

As mentioned in *Introduction 1.1.1*, sitagliptin (**7**) synthesis requires an enantioselective route to produce the enantiomerically pure, optically active (*R*)-enantiomer. Hence, taking advantage of enzymes such as transaminases, for enantioselective reductive amination was an attractive prospect. Merck and Codexis developed a route for the biocatalytic preparation of enantiomerically pure (*R*)-**7** using isopropylamine and acetone as amine donors and acceptors respectively (Figure 14).⁵² An (*R*)-selective transaminase, ATA-117, which shows 99.7% sequence identity to a parental (*R*)-transaminase from *Arthrobacter sp.* KNK168⁵³, was selected and engineered to accept the ketone intermediate **7c**.⁵² Ultimately after 11 rounds of directed evolution using ATA-117, 27 mutations were made, 10 of which were not catalytic but found to interact with the substrate in the binding pocket.⁵² The other 17 were found outside of the binding pocket (Figure 14).⁵² This variant gave extremely good enantioselectivity with an e.e. of >99.95% and an overall yield of 92%.⁵²

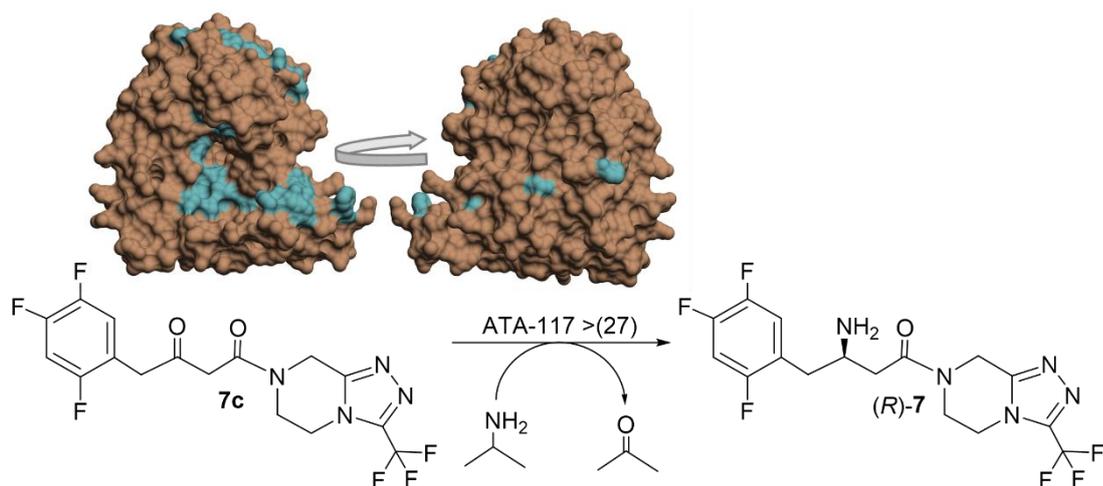


Figure 14: Synthesis of (*R*)-sitagliptin catalysed by (*R*)-selective transaminase ATA-117. The scheme represents the reductive amination of **7c** ketone intermediate to (*R*)-**7** catalysed by ATA-117 using isopropylamine and acetone as amine donor and acceptors respectively. > denotes point mutations, of which there are 27 in the final construct. The crystal structure of ATA-117 >(27) has been visualised using CCP4mg (v.2.10.11) and the overall surface structure of the protein is highlighted in brown with mutations being highlighted in blue, both the 'front' with the binding pocket apparent, and the 'back' of the protein is represented.

Transaminases require the pyridoxal 5'-phosphate cofactor (PLP) which in turn is reduced to form pyridoxamine 5'-phosphate (PMP) (Figure 13 & 15). One limitation of transaminases lies in the fact that the PLP cofactor requires a primary amine-based donor in its mechanism to successfully generate PMP and then regenerate PLP (Figure 15). PLP, when within the transaminase, is coordinated by a catalytic lysine.^{51,54} The amine donor then attacks this E-PLP complex generating an aldimine.^{51,54} The pyridine ring of PLP acts as an electron sink to form the resulting ketimine.^{49,51,54} The resulting ketimine is hydrolysed and the resulting carbonyl is formed and (E-)PMP is generated.^{51,54} PMP can then transfer the amine group to the acceptor molecule to generate the resulting chiral amine and regenerating (E-)PLP.^{51,54} Hence, transfer to the carbonyl acceptor can only result in the production of primary amines and therefore limits the functionalisation of potential products.

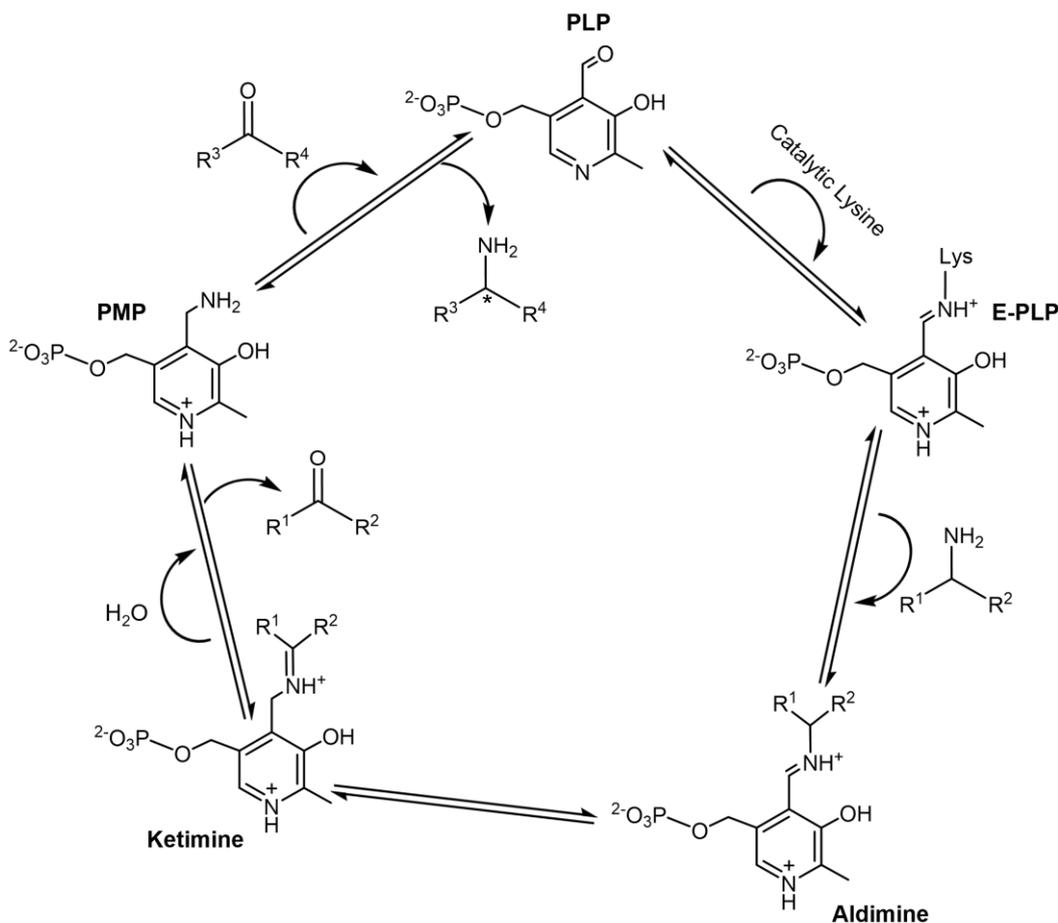


Figure 15: PLP-PMP mechanism for transaminases. This schematic displays the generalised mechanism of the catalytic enzyme-PLP complex (E-PLP) to generate the resulting chiral amine via amine group transfer from the donor to the acceptor molecule. Generation of PMP and regeneration of PLP is also depicted. R¹, R², R³ and R⁴ reflect various substituted groups. Potential chiral centres of interest are marked with an asterisk.

Another potential limitation with the use of ω -TAs is the build-up of coproducts as the reaction progresses. The ketone by-product that is produced over-time inhibits the forward reaction and further generation of the desired chiral amine. There are several methods for the removal of ketone by-products but tend to be cumbersome.^{55,56} Most studies conducted into this area are typically concerned with the removal of pyruvate; this is the keto-acid by-product formed when alanine is routinely used as the amine donor in the reaction. Recent work has been carried out into the use of a multi-enzyme system for the removal of pyruvate.⁵⁶ Lactate dehydrogenase (LDH) can be employed to catalyse the free pyruvate to lactate. Since LDH requires NADH as a cofactor, glucose dehydrogenase (GDH) is also required to regenerate this cofactor and allow the multi-enzyme reactions to continue. Consequently ω -TA mediated reactions can lead to either low yields or inefficient and relatively costly reactions with the need to use multi-enzyme/cofactor systems. However, the application of this ω -TA-LDH-GDH system is still an attractive prospect. A commercially available ω -TA from *Vibrio*

fluvialis (Vf.ω-TA) in tandem with the LDH-GDH recycling system was used to produce the Alzheimer's drug (S)-rivastigmine **1** with yields larger than 80% and >99% e.e (Figure 16).⁵⁷

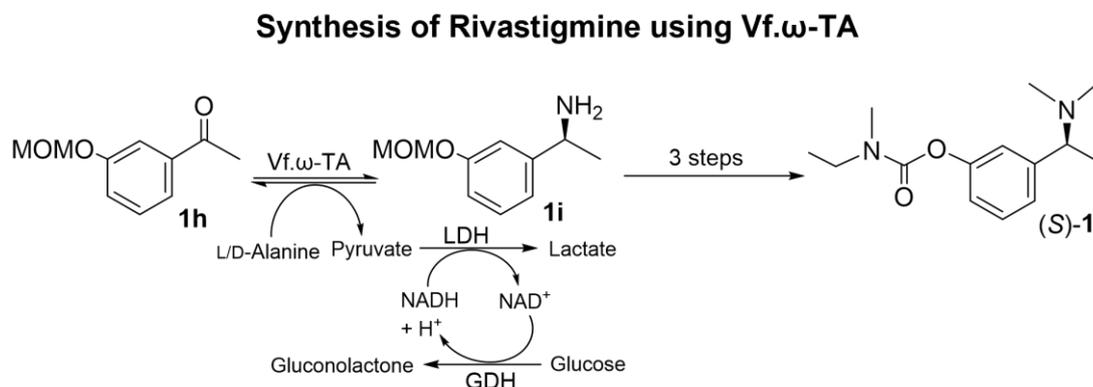


Figure 16: The synthesis of (S)-rivastigmine using Vf.ω-TA. This schematic displays the reductive aminase step of the synthesis of (S)-rivastigmine using a transaminase from *Vibrio fluvialis*. Highlighted is the recycling system set up using glucose dehydrogenase (GDH) and lactate dehydrogenase (LDH) enzymes for the removal of pyruvate.

Furthermore, another drawback of native ω-TAs is seen in their narrow substrate scope. Often native ω-TAs do not accept carbonyl centres with bulky substituents. However, in recent years, given the advances in enzyme engineering, these ω-TAs have been able to be rationally designed to accept a wide range of substrates. In one example the rational design of the ω-TA from *V. fluvialis* for the synthesis of (1S)-1-(1,1'-biphenyl-2-yl)ethanamine **9**, using an isopropylamine donor and acetone acceptor, was examined (Figure 17).⁵⁸

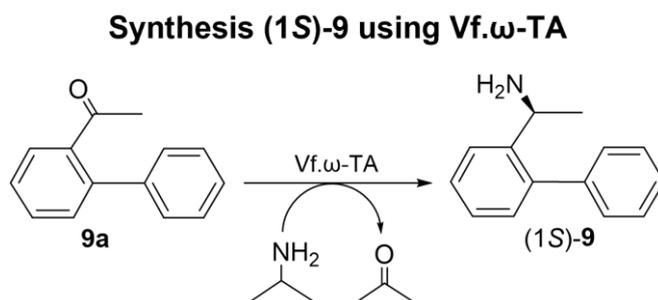


Figure 17: The synthesis of (1S)-1-(1,1'-biphenyl-2-yl)ethanamine using Vf.ω-TA. This schematic displays the reductive aminase step of the synthesis of (1S)-9 using a transaminase from *Vibrio fluvialis*. Highlighted is the isopropylamine donor and the acetone acceptor compounds that were used.

When looking at the wild-type enzyme (PDB: 4E3Q) it is clear that the binding pocket is relatively small, too small to accept a bulky ketone such as 2-acetylphenyl **9a** (Figure 18).⁵⁸ This is likely why the WT enzyme shows no activity towards 2-acetylphenyl. From docking and molecular dynamic simulations, it was clear that W57 is orientated into the middle of the binding pocket (Figure 18).⁵⁸ Therefore, the first round of rational design largely focused on mutation at this position to smaller, non-polar residues such as alanine or glycine.⁵⁸ In addition to this, the binding pocket is lined with positively charged residues such as R88, K163, and R415 (Figure 18). These were all mutated to non-polar residues in the second round of screening to increase the hydrophobicity of the binding pocket and negate any additional effects from the non-polar ketone substituent.⁵⁸ Ultimately seven mutations were made to give the best mutant: W57F/ R88H/ V153S/ K163F/ I259M/ R415A/ V422A.⁵⁸ W57F not only increases the size in the active site but also interacts with the ketone and side chain of I259M to form a π -stacking network with a S/ π interaction from the sulfur atom.⁵⁸ R88H, K163F and R415A all decrease the overall positive charge of the active site or near the vicinity of the active site (R88H).⁵⁸ V153S was adopted to slightly decrease size at the subunit interface without additional charge being added.⁵⁸ V422A is thought to interact with the side chain of the ketone via a hydrophobic interaction.⁵⁸ This mutant gave a reaction rate increase of 1716-fold with 99% e.e. for the production of **9** (Figure 17).⁵⁸ This case study highlights the promising advances made in the engineering of ω -TAs to produce a large range of chiral amines with improvements in efficiencies, substrate scopes, enantioselectivities and yields.

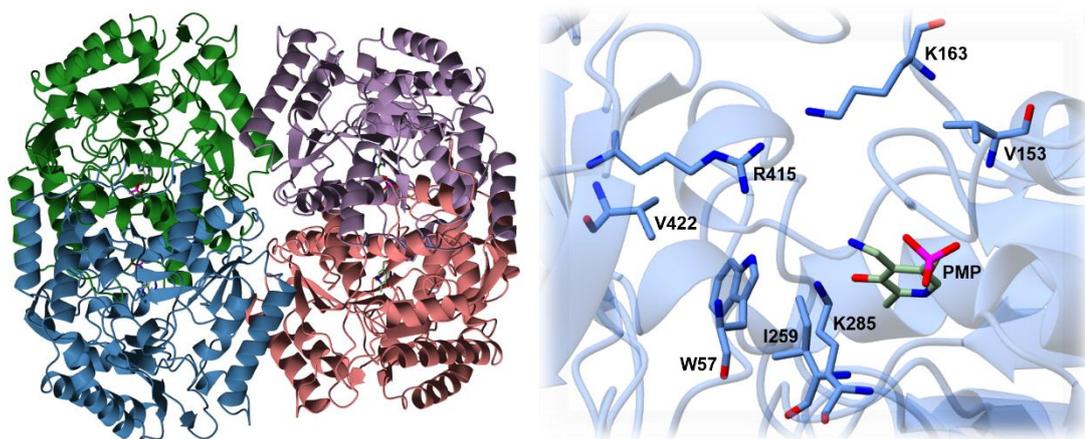
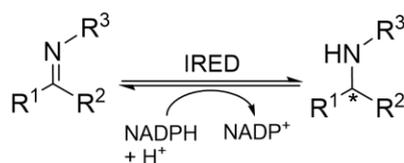


Figure 18: Biological assembly and active site of Vf. ω -TA in complex with PMP. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the biological assembly of the wild type ω -TA from *V. fluvialis* (PDB: 4E3Q). The right-hand panel represents the associated active site residues from chain A (blue) of the wild-type enzyme as well as the catalytic lysine linked to PLP/PMP (K285). The PMP cofactor is represented in green in all cases.

1.2.2 Imine Reductases and Reductive Aminases

Conventionally imine reductases (IREDs) are used for the reduction of preformed imines to amines usually requiring an NADPH cofactor (Figure 19).^{59,60} Substituents on the imine dictate which at which face on the iminium ion the hydride attacks from and therefore an attractive feature of IREDs is their ability to tailor enantioselectivity. One example of an particularly enantioselective IRED is the (*S*)-selective *Bcl*RED from *Bacillus cereus*. Furthermore, the (*S*)-selective *Bcl*RED homologue from *Nocardiosis halophila*, *Nhl*RED, displayed switching preferences in enantioselectivity for a range substituted piperideines and pyrrolines.⁶¹ *Nhl*RED, when supplied with 2-methyl piperideine formed the (*S*)-amine whereas when supplied with 2-*n*-propyl piperideine an inversion of selectivity was seen forming the (*R*)-amine.⁶¹ An additional example was seen during the screening of 20 IRED proteins carried out by Roche researchers.⁶² Among them IR_14, annotated as an (*R*)-selective IRED from *Nocardia cyriacigeorgica*, showed differential selectivity; for monocyclic and bicyclic aliphatic imines (*R*)-amines were produced.⁶² For the larger substituted bicyclic benzylic aromatic imines (*S*)-amines were observed.⁶² This reflects not only the varied enantioselectivity between IREDs but the ability to tailor one IRED to produce either desired enantiomer. This presents an attractive chance for the further engineering of IRED scaffolds to enable enantiomeric switching centred around imine functionality.

Native imine reductases (IREDs)



'Reductive Amination' using imine reductases (IREDs)

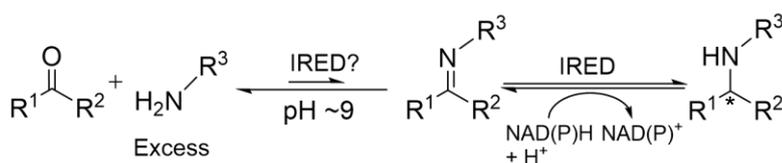


Figure 19: Reactions catalysed by Imine Reductases (IREDs). This schematic displays the generalised reactions of native IREDs and native or engineered IREDs for 'reductive aminations' of prochiral ketones. R¹, R² and R³ reflect various substituted groups. Potential chiral centres of interest are marked with an asterisk.

A drawback of using native IREDs for the formation of chiral amines is due to their substrate scope being largely limited to prochiral imines. This limits the potential

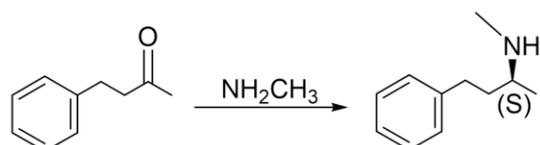
plethora of chiral amine functionalities able to be produced. Subsequently, attention was directed into finding IREDs capable of accepting carbonyl centres such as prochiral ketones (Figure 20). A range of studies have investigated the ability of IREDs to enable the formation of secondary amines from prochiral ketones when using a primary amine donor. Both Huber and co-workers⁶³ and Scheller⁶⁴ and co-workers were able to produce (*S*)-2(methylamino)-4-phenyl-butane **10a** and (*R*)-1-phenylethylamine **11a** (and (*R*)-secondary amine product **11b**) from their respective ketones (**10** and **11**) when using IREDs from *Streptomyces* sp. GF3587 and *Streptosporangium roseus* respectively (Figure 20). For *Streptomyces* sp. GF3587 IRED 8.8% conversion and an *e.e.* of 76% was recorded.⁶³ For the *S. roseus* IRED the secondary amine product yielded 39% conversion and an 87% *e.e.*, for the primary amine product only 16% conversion was recorded with 97% *e.e.*⁶⁴ Likewise, Roche carried out a large screen of a library of 28 IREDs towards prochiral ketone substrates.⁶⁵ Across the board most IREDs from this library showed little or no activity when ammonia or butylamine was the amine donor, showing better conversions when using methylamine.⁶⁵ Two IREDs IR_20 and IR_11 performed the best towards (*R*)-3-methylcyclohexanone **12** with ammonia and methylamine donors respectively (Figure 20).⁶⁵ Ultimately a 50% yield, 94% *d.e* and 96% conversion for IR_20 was achieved and a 55% yield, 96% *e.e.* and 95% conversion for IR_11 was recorded.⁶⁵ These initial findings were then applied to the production of small molecule pharmaceuticals. It was demonstrated that indanone **13a** could be reduced to both (*R*)- and (*S*)-rasagiline **13.1/13.2**, a drug used to treat Parkinson's disease, using two IREDs; IR_14 from the Roche generated library and IR-Sip from *Streptomyces ipomoeae* 91-03 (Figure 21).⁶⁶ In the presence of propargylamine IR_14 was able to produce (*R*)-rasagiline **13.1** with a 71% conversion, 58% yield and 90% *e.e.*, IR-Sip was able to produce (*S*)-rasagiline **13.2** with a 91% conversion, 81% yield and 72 *e.e.*⁶⁶

However, what was clear from all these examples was the relatively low conversions with ammonia.⁶¹⁻⁶⁵ This limits the possibility for the effective production of any primary amines limiting the potential functionalities of these chiral centres. Additionally, to see any significant conversion in these reactions an excess of the amine donor is required typically varying from 12.5 to 50 equivalents.⁶¹⁻⁶⁵ In addition to this high pH values of around 9.0 are required to produce substantial conversions and yields.⁶¹⁻⁶⁵ It is clear from these factors that high pH and high excess of amine donors is what ultimately drives the formation of the imine or iminium ion intermediate. Without these factors the IRED does not seem to produce any resulting chiral amine, suggesting that IREDs

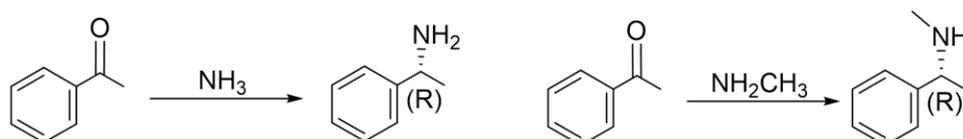
do not in fact catalyse the first ‘reductive amination’ step and rather catalyse the reduction of the preformed imine in solution as dictated by the reaction conditions. Therefore, the rate limiting step of this reaction is the spontaneous formation of an imine in solution.

‘Reductive Amination’ examples using imine reductases (IREDs)

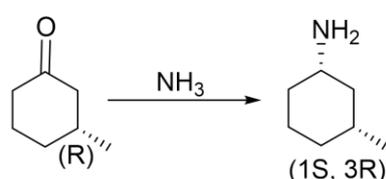
IREd from *Streptomyces* sp. GF3546



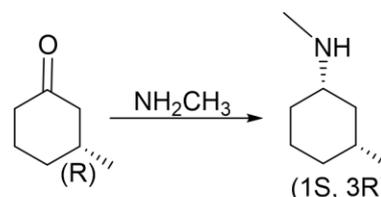
IREd from *Streptosporangium roseus*



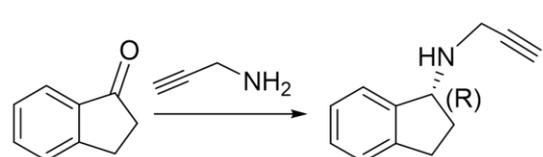
IR_20 from *Streptomyces tsukubaensis*



IR_11 from *Verrucospora maris*



IR_14 from *Nocardia cyriacigeorgica*



IR_Sip from *Streptomyces ipomoeae*

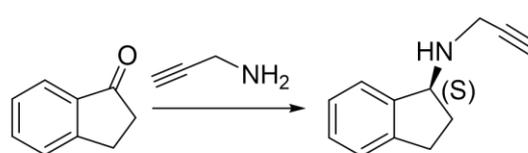


Figure 20: Exemplar ‘reductive amination’ reactions enabled by Imine Reductases (IREDs). This schematic displays examples of ‘reductive amination’ reactions enabled by a range of IREds for prochiral ketones using different amine donors. Stereochemistry for each reaction is assigned to each molecule.

1.2.2.1 AspRedAm

As a result of the above mentioned IREd findings, there was a pressing need to find an enzyme capable of carrying out a true reductive amination reaction, catalysing both the formation of an imine from a prochiral carbonyl centre, coupled with the imine reduction. Recently a sequence similarity screen led to the discovery of a reductive aminase from *Aspergillus oryzae* (*A. oryzae*) named AspRedAm, a eukaryotic fungal

IRED homologue.⁶⁷ *AspRedAm* (UniProtKB: Q2TW47) shares a sequence similarity of 37.6% to an (S)-IRED from *Amycolatopsis orientalis* (*AoIRED*) (UniProtKB: R4SNK4).⁶⁸ *AoIRED* is a structurally characterised IRED used to catalyse a range of prochiral imines with greater efficiency than their *Streptomyces* counterparts.⁶⁸ *AspRedAm* can catalyse the formation of an imine from a prochiral ketone and an amine donor, along with the reduction of an imine moiety to the resulting chiral amine, leading to the breakthrough discovery of the first true reductive aminase (Figure 21).⁶⁷

Reductive Amination using *AspRedAm*

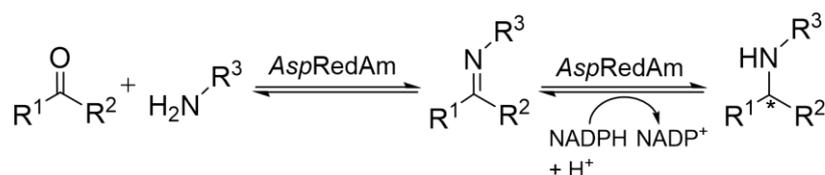


Figure 21: Reaction catalysed by *AspRedAm*. This schematic displays the generalised reaction of a native reductive aminase from *A. oryzae* (*AspRedAm*) with prochiral ketones and a range of amine donors. R¹, R² and R³ reflect various substituted groups. Potential chiral centres of interest are marked with an asterisk.

AspRedAm was screened for a range of both carbonyl substrates and amine donor libraries, usually using 1-7 mM of carbonyl to 2-5 mM amine donor with some in larger excess and some in stoichiometric amounts.⁶⁷ It was evident from the results that *AspRedAm* catalysed a range of reductive aminations with various cyclic, bicyclic and aliphatic straight chained carbonyls using various amine donors of length and functionalities, even at pH 7.0 (Figure 22).⁶⁷ *AspRedAm* typically worked the best with smaller primary amine donors (excluding ammonia) such as methylamine, allylamine, cyclopropylamine and propargylamine (Figure 22).⁶⁷ *AspRedAm* also displayed a preference for cyclic ketones or 5 and 6 membered aliphatic straight chain ketones. However *AspRedAm* did catalyse substrates as large as 1-tetralone **14** or 2-tetralone **15** for the production of amines as large as *N*-(prop-2-yn-1-yl)-1,2,3,4-tetrahydronaphthalene-1-amine **14a** and *N*-(prop-2-yn-1-yl)-1,2,3,4-tetrahydronaphthalen-2-amine **15a** respectively (Figure 22).⁶⁷ One attractive feature of *AspRedAm* was the ability to catalyse some reductive amination reactions with the amine donors and ketones in 1:1 stoichiometric quantities or 2-4 equivalents of the amine donor (Figure 22).⁶⁷ Larger substrates typically required higher loading of the amine donor(s) (Figure 22).⁶⁷ This highlights *AspRedAm* as a promising protein scaffold, without the need for an excess of amine donor, which could be extended to access an even larger scope of ketones through enzyme engineering.

Reductive Amination examples using AspRedAm

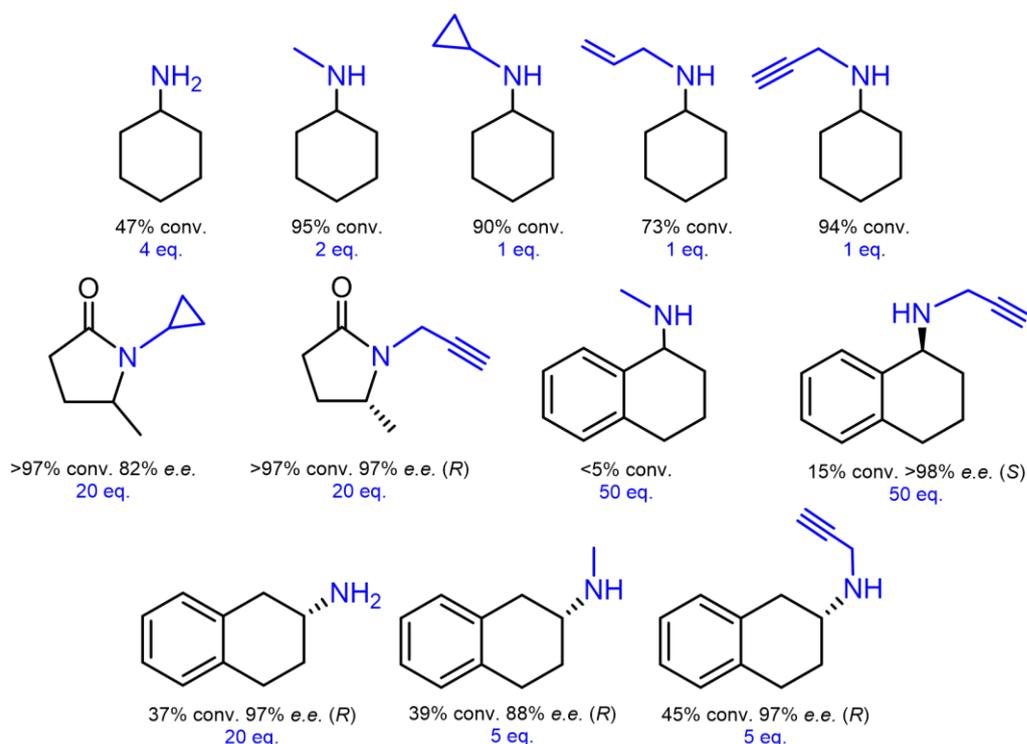


Figure 22: Exemplar reductive amination reactions catalysed by AspRedAm. This schematic displays a few examples of reductive amination reactions of cyclic and bicyclic ketones using a range of amine donors and AspRedAm. Equivalents of amine donors are highlighted and labelled in blue. Conversions and enantiomeric excesses are labelled where appropriate. Stereochemistry was assigned if determined.

Similar to IR_14^{65,66}, AspRedAm was applied to the synthesis of (*R*)-rasagiline **13.1** using propargylamine (50 equivalents) to give 64% conversion and 95% e.e. using the starting material indanone **13a**.⁶⁷ Although both reactions required an excess of the amine donor to yield similar conversions and e.e. values, AspRedAm was able to achieve this conversion after 24 h⁶⁷ whereas IR_14 achieved this conversion after 7 d.^{65,66} This is a significant reduction in process time – a good starting point for engineering AspRedAm for high turnovers of prochiral ketones with shorter reaction times.

AspRedAm, like IREDs such as AoIRED⁶⁸, displays the typical canonical ‘IRED fold’ in which the biological assembly structure represents a functional dimer; two monomers associate at an interface of N and C terminal domains (Figure 24). Each monomer is made up of a large N-terminal Rossmann domain and C-terminal helical bundle each terminal being connected by an interdomain helix (Figure 24). However, the AspRedAm, unlike IREDs seems to adopt a more ‘closed’ form around the ligand and cofactor during dynamic binding of a ligand (Figure 25). IREDs like AoIRED seem

to adopt a more 'open' substrate cage, even when in complex with ligands (Figure 24).⁶⁸ Taking a look at the active site revealed key residues crucial for reductive amination. Y177 is positioned at an ideal distance (3.27 Å in chain B) away from the amine group of the co-crystallised ligand (*R*)-rasagiline **13.1**, suggesting a possible role in proton donation or stabilisation of the substrate/ligand in the active site (Figure 23). Y177A resulted in a 30-fold decrease in reductive aminase activity when compared to the wild-type enzyme.⁶⁷ D169 is thought to play a role in proton dependent catalysis. In an IRED from *Streptomyces kanamyceticus* (PDB: 3ZHB) this position is occupied by D187; mutation to D187A or D187N resulted in loss of activity.⁶⁹ Likewise, D169A and D169N resulted in a >200-fold reduction in k_{cat}/K_m when compared to the wild-type enzyme for the reductive amination production of N-methylcyclohexylamine.⁶⁷ Tracing this back to other enzymes such as dehydrogenases reveals that typically this aspartate residue is usually a lysine residue which aids in the proton donation of the alcohol intermediate formed in these reactions.⁷⁰ Interestingly in *Asp*RED the residue in this position is N171.⁶⁸ N93 is thought to hydrogen bond to D169A and stabilise this catalytic residue (Figure 23). M239 and Q240 are brought in from the other chain on binding of the substrate/ligand to close the active site with W210 and Y217 closing the hydrophobic binding pocket (Figure 23). These findings are the structural basis that could be used for the rational design of a RedAm and/or an engineered enzyme capable of the reductive amination of a wide range of prochiral ketones including pharmaceutically and biologically significant moieties.

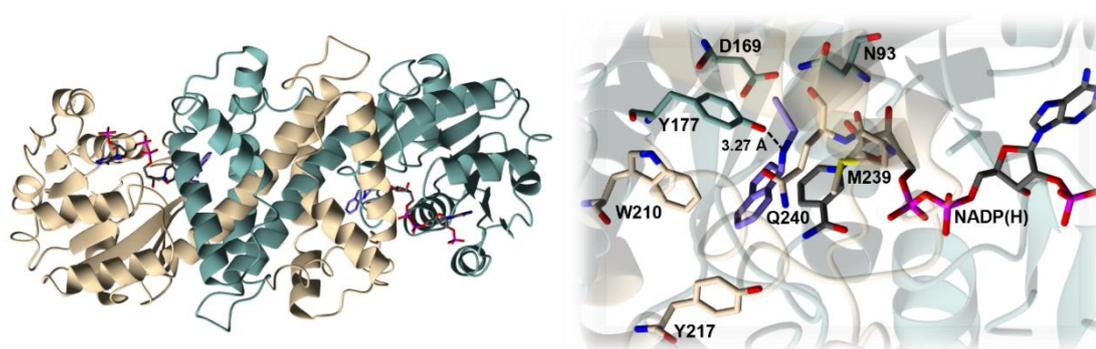


Figure 23: Biological assembly and active site of *Asp*RedAm in complex with NADP(H) and (*R*)-rasagiline. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the biological assembly of the wild type *Asp*RedAm (PDB: 5G6S) dimer using chains B (blue) and C (yellow) in complex with NADP(H) grey and (*R*)-rasagiline (purple). The right-hand panel represents the associated active site residues from chain B (blue) and chain C (yellow) of the wild-type enzyme. The dotted line indicates the distance between the Y177 hydroxyl and the amine group of (*R*)-rasagiline (purple), the distance is measured in Ångstroms.

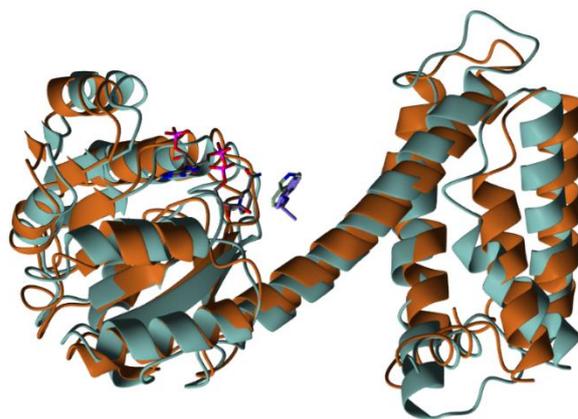
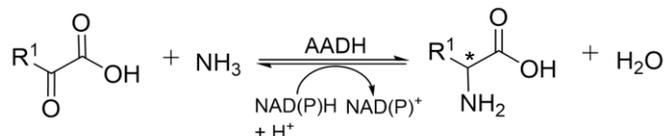


Figure 24: Monomeric superimposition of AoIRED and AspRedAm in complex with associated ligands and NADP(H). All images have been visualised using CCP4mg (v.2.10.11). NADP(H) is highlighted in grey in all cases. SSM superimposition of AoIRED chain A (PDB: 5FWN) (orange) in complex with (*R*)-methyltetrahydroisoquinoline with AspRedAm chain B (PDB: 5G6S) (blue) in complex with (*R*)-rasagiline (purple). RMSD values are 2.26 Å across 265 C α atoms.

1.2.3 Amino Acid Dehydrogenases and Amine Dehydrogenases

As with transaminases, native amino acid dehydrogenases (AADHs) are used *in vivo* for the recycling and degradation of L- and D-amino acids to oxo acids or the forward (major) reaction to produce functionalised amino acids using ammonia (Figure 25).^{71,72} Typically, these enzymes utilise NAD(P)H as their cofactor for hydride transfer depending on the role of the enzyme. Normally *in vivo*, L-AADHs used are used for catabolic processes and utilise NAD(H) whereas D-AADHs are used in biosynthetic pathways to produce branched chained amino acids such as leucine, isoleucine and valine, often requiring NADP(H).⁴⁹ These AADHs have been engineered to accept unfunctionalised carbonyl centres and not just amino acids/ α -keto acids, where they utilise ammonia, and more recently, a range of amine donors^{73–76}, creating the ‘amine dehydrogenases’ (AmDHs) family (Figure 25).^{73–83}

Native Amino Acid Dehydrogenases (AADHs)



Engineered Amino Acid Dehydrogenases (AADHs)

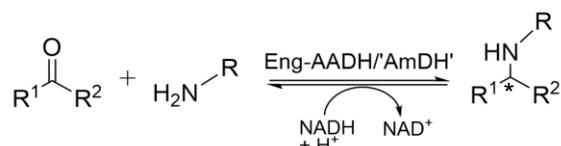


Figure 25: Reactions catalysed by Amino Acid Dehydrogenases (AADHs). This schematic displays the generalised reactions of native AADHs and engineered AADHs/‘Amine Dehydrogenases’ (AmDHs) for reductive aminations of prochiral ketones. R¹ and R² reflect various substituted groups. R typically reflects a hydrogen atom but in recent cases can be a variety of substituted groups. Potential chiral centres of interest are marked with an asterisk.

The structures of AADHs have been extensively studied over the last 30 years. Typically, the catalytic monomer of these AADHs comprises two domains. At the interface of the two domains is a cofactor binding cleft which undergoes a dynamic closure on ligand/substrate binding to engulf both cofactor and substrate in close proximity for efficient hydride transfer.^{71,72} These monomers can associate to form dimers and often the biological assembly in solution is a multimer.^{71,72} The two most studied AADHs, LeuDh and PheDH are NADH dependent AADHs involved in the interconversions that can produce L-leucine and L-phenylalanine from their respective keto/oxo acid.

LeuDh is an AADH of particular significance. Its native product leucine is involved in several metabolic and physiological pathways, specifically it triggers the mammalian target of rapamycin (mTOR), a kinase that regulates several cellular pathways such as, and not limited to, tissue and muscle synthesis, cell proliferation, transcription, and autophagy.^{84,85} In the pharmaceutical industry LeuDh has been used for the biotransformations of leucine based protease inhibitors for treatment and management in various diseases such as HIV/AIDs or cancerous metastasis.²⁴ The structure of a LeuDh homologue from *Bacillus sphaericus* was determined by X-ray crystallography (PDB: 1LEH).⁷⁸ It displays a typical 'AADH fold' where the monomer is comprised of two domains and the NADH and subsequent substrate/product binds at the interface of the two (Figure 26). It is clear from the active site that residues K68, K80 and D115 stabilise the product L-leucine in the active site as they are within close proximity for facilitating electrostatic interactions (Figure 26). K80 is thought to be involved in the activation of water for nucleophilic attack of the iminium ion intermediate generated during the reductive amination.⁷⁹ K80 is highly conserved amongst other AADHs of interest. Regarding the specific activity towards L-leucine, a K80Q mutation resulted in a 3-fold increase of the K_m value when compared to the wild-type enzyme.⁷⁹ Other residues found within the active site are thought to play a role in recognition of the substrate and or cofactor (Figure 26).

Due to the significant role K80 and D115 are thought to play in catalysis, these corresponding positions were avoided in the rational engineering of a LeuDh from *Bacillus stearothermophilus* to produce the resulting 'AmDH'.⁸¹ This engineered LeuAmDH was targeted for the reductive amination of the α -ketoisocaproate (natural keto acid to produce L-leucine) ketone analogue methyl isobutylketone (MIBK) **16** to produce (*R*)-1,3-dimethylbutylamine (1,3-DMBA) **16a** (Figure 29).⁸¹ K68M displayed low activity towards **16** but it was a good starting point for observing reductive amination properties which could be subjected to further enzyme engineering.⁸¹

Subsequent rounds of engineering yielded the mutant K68S/E114V/N262L/V291C which displayed a >3000-fold higher k_{cat} value for the reductive amination of **16** to **16a** when compared to the K68M mutant (Figure 28).⁸¹ This four-point mutant was able to produce 1,3-DMBA (**16a**) with a 92.5% conversion and 99.8% e.e.⁸¹ N262L and V291C were crucial for enlargement of the binding pocket to accommodate MIBK.

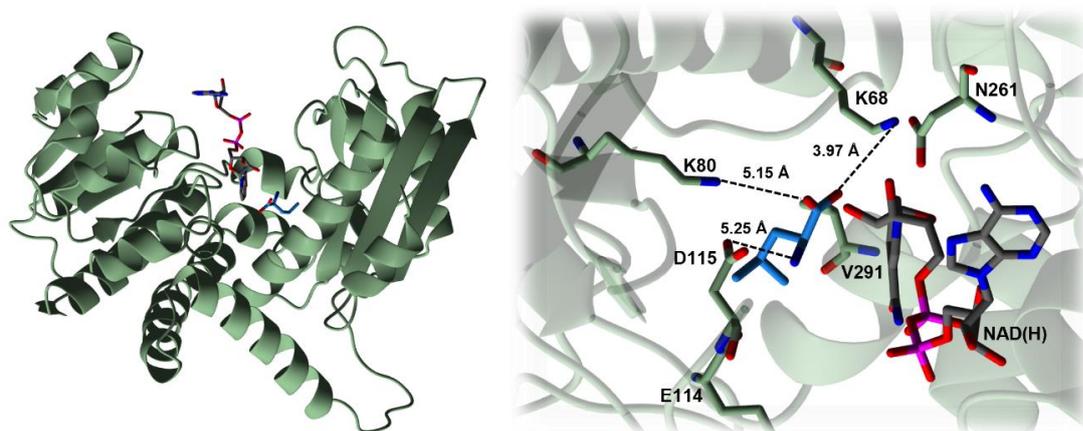


Figure 26: Catalytic monomer and active site of LeuDh in complex with NAD(H) and L-leucine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the monomeric form of the wild type LeuDh from *B. sphaericus* (PDB: 1LEH) using chains A (green) in complex with NAD(H) (grey) and L-leucine (blue). The right-hand panel represents the associated active site residues from chain A (green). The dotted lines indicate the distances of K68, K80, D115 and L-leucine (blue) based interactions, the distances are measured in Ångstroms. *N.B.* L-leucine and NAD(H) were modelled in CCP4mg from a 3D coot dictionary, its associated atom coordinate files, and the structure of PheDH (PDB: 1C1D).

PheDH, an amine dehydrogenase homologue of LeuDh, sharing 48% sequence similarity, catalyses the interconversion of L-phenylalanine and phenylpyruvate. Structurally PheDH is very similar to LeuDh and each catalytic monomer shares two domains in which the interface of both accommodates NAD(H) and its associated substrates/products (Figure 27). As with LeuDh, the dynamic closure of both domains brings the cofactor and substrate into close proximity for hydride transfer (Figure 27). A PheDH from *Rhodococcus sp.* M4 was crystallised with L-phenylalanine (PDB: 1C1D) and is the structural basis for the rational engineering of these class of AADHs and ultimately functional AmDHs.⁸²

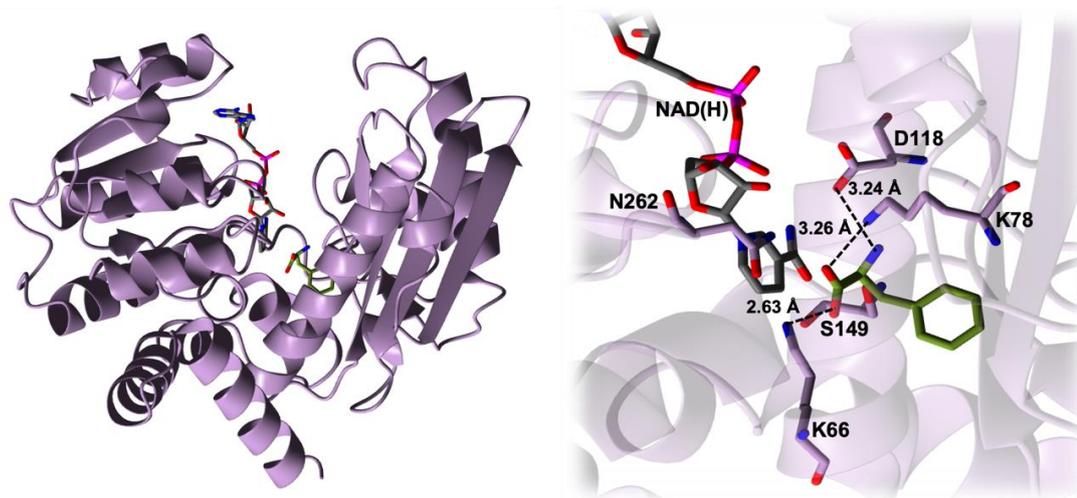
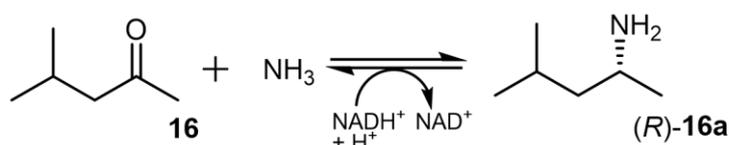


Figure 27: Catalytic monomer and active site of PheDH in complex with NAD(H) and L-phenylalanine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the monomeric form of the wild type PheDH from *Rhodococcus sp. M4* (PDB: 1C1D) using chains A (purple) in complex with NAD(H) (grey) and L-phenylalanine (green). The right-hand panel represents the associated active site residues from chain A (purple). The dotted lines indicate the distances of K66, K78, D118 and L-phenylalanine (green) based interactions, the distances are measured in Ångstroms.

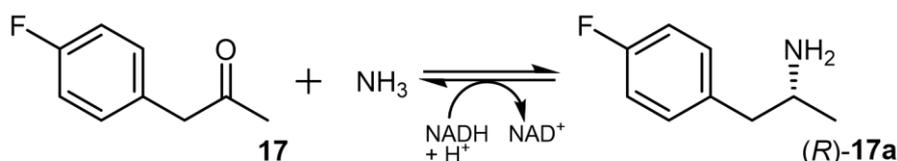
The LeuD_H mutations at positions K68S/E114V/N262L/V291C, showing promising reductive amination abilities, were mapped onto a *Bacillus badius* (*B. badius*) PheDH scaffold.⁸⁰ Again, the corresponding positions for K80 and D115 in PheDH were not included in the screening due to their crucial role in catalysis. In the first round of *B. badius* PheDH screening it was shown that the double mutant K77M/N276V (K68 and N262 in LeuD_H), had significantly improved amination properties towards MIBK **16**.⁸⁰ Both K77 and N276 interact and stabilise, respectively, with the carboxyl moiety of L-phenylalanine (Figure 27: K66/N262) but interestingly single mutations at these sites did not significantly improve its reductive amination ability.⁸⁰ This is likely due to the combination of both mutants working synergistically to interact with the carbonyl moiety of the substrate (MIBK - **16**). PheDH was screened for reductive amination using *p*-fluorophenylacetone (PFPA) **17**, an analogue of phenylpyruvate (one of the natural substrates of PheDH) (Figure 28).⁸⁰ The double mutant K77M/N276V showed much better reductive amination towards **17** when compared to **16**.⁸⁰ Screening libraries at both these positions allowed for more rigorous evaluation and engineering of PheDH towards reductive amination reactions. The ultimate variant K77S/N276L displayed a 15-fold greater k_{cat} value for reductive amination of **17**⁸⁰ when compared to that seen previously of LeuD_H with **16**.⁸¹ This variant was also screened for substrate scope using a wide range of prochiral ketones varying from smaller aliphatic compounds to larger ketones with cyclic functionalities.⁸⁰ Further work was carried out

to ultimately produce a triple mutant K66Q/S149G/N262C in *Rhodococcus sp. M4* which gave 95% conversion and >98% e.e. for the production of (*R*)-1-methyl-3-phenylpropylamine **18a** from 4-phenyl-2-butanone **18** (Figure 28).⁸³ In *R. Sp M4*-PheDH, similarly to LeuDh, catalytic residues K78 and D118 (K80 and D115 in LeuDh) are responsible for the coordination of the carboxyl moiety in the natural substrate. It was found from modelling that when binding functionalised ketones, the engineered PheDH's active site adopted a different, unfavourable orientation.⁸³ This was mainly due to the presence of the K66 mediated hydrogen bond towards the C=O of the ketone substrate. Therefore, it was theorised that K66 needed to be mutated in order to allow for the re-establishment of the hydrogen bonding network between K78 and D118. N262C is thought to bring the substrate closer to active site residues. S149G decreased side chain length and thus enlarged the hydrophobic binding pocket of engineered PheDH. Both the engineering of LeuDh and various PheDH scaffolds paves the way for the design of an enzyme capable of reducing unfunctionalised and functionalised prochiral ketones to chiral amines and the ultimately the formation of an artificial 'AmDH'.

K68S/E114V/N262L/V291C LeuDh AmDH from *B. stearothermophilus* with MIBK



K77S/N276L PheDH AmDH from *B. badius* with PFPA



K66Q/S149G/N262C PheDH AmDH from *R. sp M4* with 4-phenyl-2-butanone

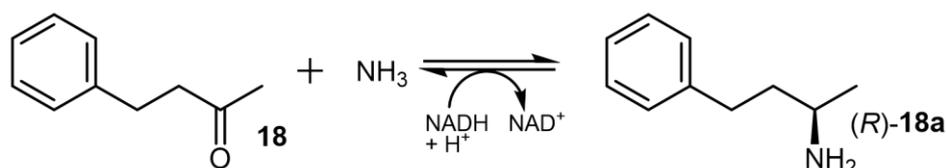


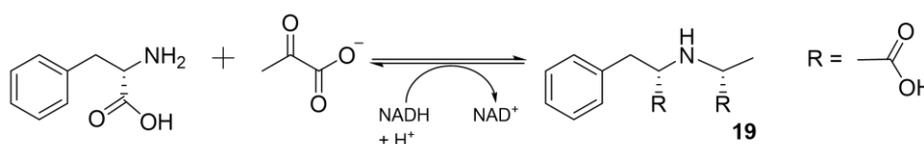
Figure 28: Exemplar reductive amination reactions catalysed by engineered Amino Acid Dehydrogenases (AADHs) to form artificial AmDHs. This schematic displays examples of reductive amination reactions catalysed by a range of engineered LeuDh and PheDH AADHs for prochiral ketones using ammonia. Stereochemistry for each reaction is assigned to each molecule.

1.2.3.1 Opine Dehydrogenases

Opine Dehydrogenases (OpDHs) are a class of NAD(P)H dependent AADHs that are responsible for the biosynthesis of opines (*N*-derivatised amino acids) from keto acids and amino acids (Figure 29).⁷² For example CENDH OpDH used for the synthesis of *N*-[1-(*R,S*)-(carboxyl)ethyl]-(*S*)-phenylalanine **19** from L-phenylalanine and pyruvate (Figure 29).⁷⁷ However, OpDHs, distinct from earlier AADHs, can couple carbonyl centres with amine donors that are larger than ammonia. This allows for the synthesis of compounds not limited to primary species, a good scaffold for engineering reductive amination reactions to produce larger, more functionalised, chiral amines. In recent years OpDHs have been engineered to accept prochiral ketones not limited to keto acids such as 2-pentanone **20**, cyclohexanone **21** and 2-teralone derivatives **22** (Figure 29).^{86,87}

Reactions catalysed by native and engineered OpDHs

OpDH catalysed formation of *N*-[1-(*R,S*)-(carboxyl)ethyl]-*S*-phenylalanine



Engineered OpDH (CENDH) and ketone based substrates

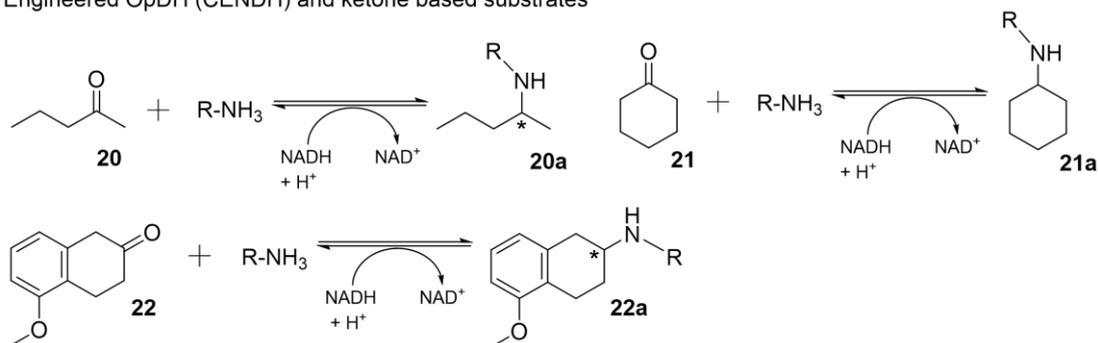


Figure 29: Examples of reactions catalysed by native and engineered Opine Dehydrogenases (OpDHs). This schematic displays the reaction examples catalysed by a range of native and engineered OpDHs to produce opines and secondary (chiral) amines respectively. R, unless specified otherwise, represents a range of alkyl substituents on the amine donors and amine products. Chiral centres of interest, if not assigned by stereochemistry, are marked with an asterisk.

The overall tertiary structure of OpDHs, as expected, displays the typical AADH fold in which the two domains of a catalytic monomer come together in which the interface between the two domains accommodates NAD(P)H and relevant substrates/products (Figure 30). In OpDHs different active site residues that are not the typical conserved

AADH residues are involved in the hydrogen bonding network to stabilise the substrate/product in the active site. The structure of an OpDH from *Pecten maximus* was obtained using X-ray crystallography in complex with NADH and L-arginine (PDB: 3C7C).⁸⁶ It was observed that H212 and E142 interact in a hydrogen bonding network with L-arginine (Figure 30).⁸⁶ However, both ligand and H212/E142 are relatively far away from the NADH cofactor (Figure 30). This suggests that the hydride transfer, to reduce the iminium ion during the reaction, must take place through other active site residues, stabilising L-arginine in a closer proximity to NADH but still within the monomeric unit. Structurally speaking, studies on this OpDH are just beginning and so no clear explanation is yet offered but nevertheless was decided to be an appropriate target for enzyme engineering. CENDH variants produced such as A111M/ K156T/ N198H/ Y259M/ Y280L/ R292V/ Y293H were crucial for the coupling of a range of amine donors to carbonyl centres, for example the coupling of cyclohexanone **21** and butylamine (Figure 29) to yield an e.e. value of >95%.⁸⁷ The residues subjected to mutagenesis were all found to be within the active site cleft and so may be the basis of the hydride transfer network, key for the reductive catalysis. These preliminary studies are helping to create the foundation for a OpDH based protein scaffold, to be specifically engineered to produce secondary and possibly primary and tertiary chiral amines, something which engineered AADHs (artificial AmDHs) at the time had not yet accessed.^{77–82}

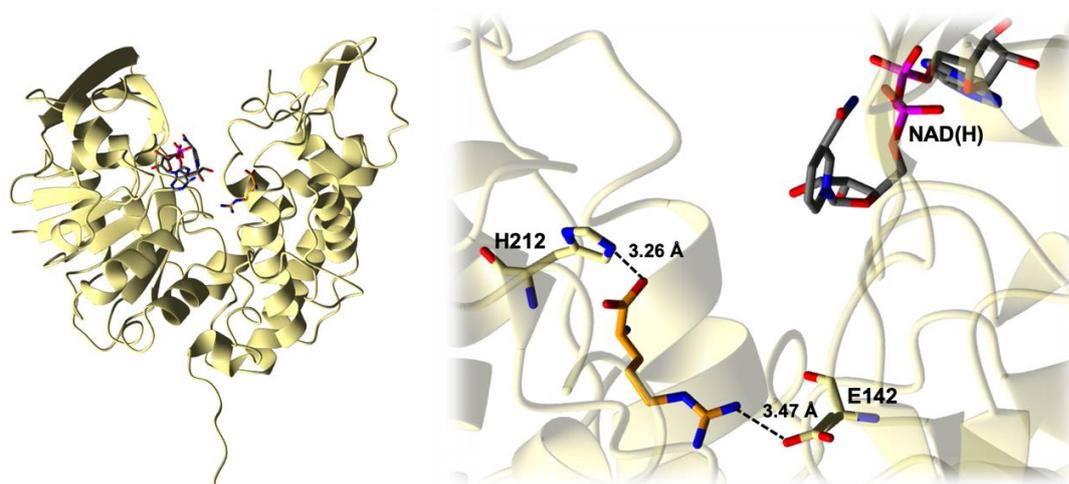


Figure 30: Catalytic monomer and active site of OpDH in complex with NAD(H) and L-arginine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the monomeric form of the wild type OpDH from *Pecten maximus* (PDB: 3C7C) using chains A (yellow) in complex with NAD(H) (grey) and L-arginine (orange). The right-hand panel represents the E142 and H212 hydrogen bonding network from chain A (yellow). The dotted lines indicate the distances of E142 and H212 and L-arginine (orange) based interactions, the distances are measured in Ångstroms.

1.2.3.2 Improving the promiscuity of Amine Dehydrogenases

In more recent years, thanks to the engineering of AADHs⁷⁷⁻⁸² and the discovery of OpDHs^{86,87}, artificial AmDHs have been engineered to accept a wide range of amine donors not limited to ammonia.⁷³⁻⁷⁶ Currently this means AmDHs now have the ability to synthesise larger secondary and tertiary chiral amines as well as primary amines. Two AmDHs Rs-AmDH and Ch1-AmDH, were screened for substrate and amine donor spectra.⁷⁶ Rs-AmDH is a variant derived from the L-phenylalanine dehydrogenase from *Rhodococcus sp.* M4 and Ch1-AmDH is a chimeric AmDH.⁷⁵ Rs-AmDH and Ch1-AmDH were able to accept primary amine donors such as methylamine and cyclopropylamine and a range of carbonyl substrates to produce various secondary chiral amines (Figure 31).⁷⁶ Additionally, as a proof-of-concept Rs-AmDH was also able to accept secondary amine donors such as dimethylamine and ethylmethylamine with the ketone substrate 4-phenyl-butan-2-one **23**, for the synthesis of tertiary amines **23a** and **23b** respectively, albeit with poor conversions (Figure 31).⁷⁶ Surprisingly in some cases formation of the primary amine was observed (Figure 31), the primary amines were always enantiomerically pure (*R*)-amines.⁷⁶

Reductive Amination scope of Ch1-AmDH and Rs-AmDH

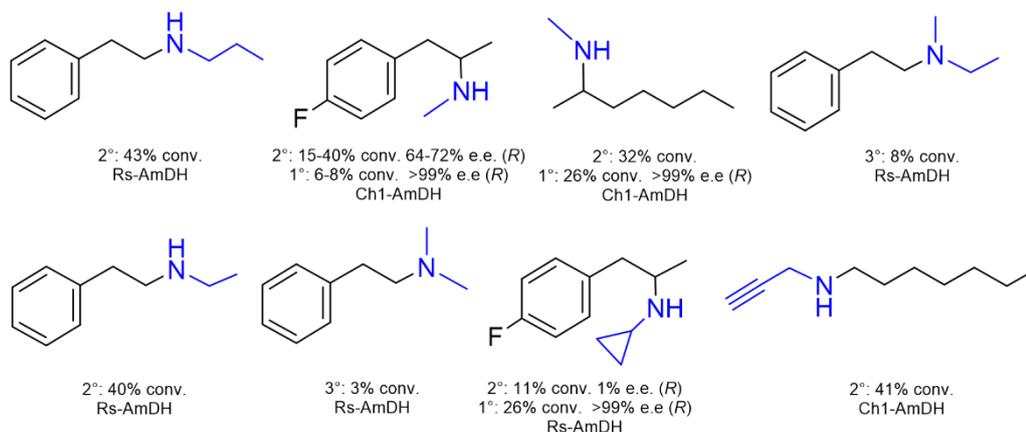


Figure 31: Exemplar reductive amination reactions catalysed by Ch1-AmDH and Rs-AmDH. This schematic displays a few examples of reductive amination reactions of cyclic and straight chained ketones and aldehydes using a range of amine donors and. Amine donors within the product compounds are highlighted and labelled in blue. Conversions and enantiomeric excesses for secondary and tertiary amines are labelled where appropriate. Any primary amine conversions and enantiomeric excesses are labelled where appropriate. Stereochemistry was assigned if determined.

The side mechanism which is responsible for the production of enantiopure primary amines from secondary (or tertiary) amines was derived.⁷⁶ This mechanism involves the NAD⁺ mediated oxidation of the secondary amine to a pro-(*R*)-iminium ion, and

then the hydrolysis of the iminium ion to the primary amine and release of a carbonyl by-product.⁷⁶ This activity closely resembles the canonical transamination activity seen in many PLP dependent enzymes (Figure 15).^{49,51,54} This novel mode of binding for larger amine donors as well as the transaminase like side reaction was crucial in elucidating the prospective promiscuity of AmDHs and is the foundation required for the further rational engineering of these enzymes to improve stereo- and substrate selectivity and scope.

To follow the advancements made by increasing AmDH promiscuity, for the synthesis of secondary and tertiary amines, and uncovering transaminase like activity, efforts turned to engineering AADHs to display both alcohol dehydrogenase (ADH) and AmDH functions.⁷³ L-lysine-dehydrogenase (LysEDH) whose native reaction concerns the oxidative deamination of L-lysine was engineered to produce the first dual enzyme that displays both ADH and AmDH activity (Figure 32).^{73,74} In a one enzyme reaction a variant of LysEDH, named LE-AmDH-v1, could catalyse the alcohol amination of benzyl alcohol **24** to benzylamine **24b** via benzaldehyde formation **24a** (Figure 32). In real time the ADH activity and AmDH activity was separated by the subsequent additions of a nicotinamide adenine dinucleotide oxidase (NOx) and the *Candida boidinii* formate dehydrogenase (Cb-FDH) to act as recycling systems in the alcohol oxidation and reductive amination respectively (Figure 32).⁷³ Interestingly, this variant only possessed one mutation; F173A, which increased space and hydrophobicity in the binding pocket owing to improved activity and enantioselectivity.⁷⁴ Again, this example highlights the promiscuity of these enzymes, making them an attractive scaffold for improving access to a range of biologically and chemically significant functionalities.

Dual ADH and AmDH activity of LE-AmDH-v1

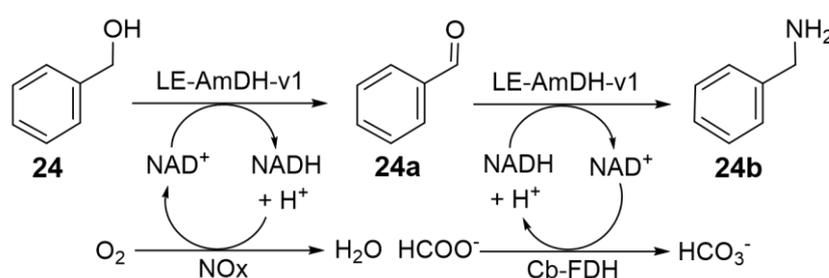


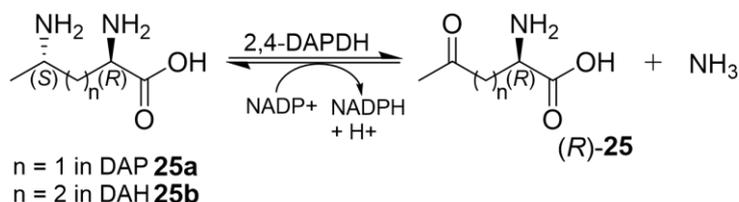
Figure 32: The one-enzyme synthesis of benzaldehyde from benzylic acid using LE-AmDH-v1. This schematic displays the alcohol amination of benzylic acid using just one enzyme, a LysEDH variant. This works though the alcohol oxidation to ketone intermediate **24a** which then undergoes reductive amination catalysed by the same LysEDH variant. Recycling systems provided by NOx and Cb-FDH are also highlighted to regenerate NAD⁺/NADH cofactors for use in the oxidoreductase reaction(s).

1.3 Native Amine Dehydrogenases

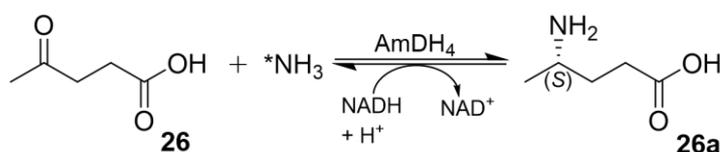
More recently, due to the advances in creating 'artificial' (i.e. engineered from AADHs) AmDHs^{73–83} and the discovery of RedAms⁶⁷, attention has been drawn to finding natural enzymes that are capable of carrying out the asymmetric reductive amination of a range of carbonyl centres using both ammonia and larger amine donors. A sequence driven search⁸⁸ of enzymes capable of transforming carbonyl-based compounds into amines using ammonia and NAD(P)H cofactors was performed.⁸⁹ Those that worked on α -keto acids were specifically removed to search for enzymes capable of accepting more unfunctionalised ketones. An enzyme, from ornithine-fermenting bacteria, was discovered and is namely (2*R*,4*S*)-2,4-diaminopentanoate dehydrogenase (2,4-DAPDH).^{89,90} 2,4-DAPDH is thought to be active towards 2,4-diaminopentanoate (2,4-DAP) **25a** and 2,5-diaminohexanoate (2,5-DAH) **25b** (Figure 33).⁸⁹ This unlocked the possibility of exploiting the reverse reaction for the search of dehydrogenases capable of reductive amination with ammonia. The protein sequence of 2,4-DAPDH as determined from *Clostridium sticklandii* (UniProtKB: E3PY99), was then used as a reference for a more refined sequence driven search to map enzymes of similar activity and function.⁸⁹ Six similar enzymes, uncovered as a result of this search, were found to be active towards 4-oxopentanoic acid and thus led to the description of the first ever native ω -amine dehydrogenase (nat-AmDH) i.e., an enzyme that was able to accept carbonyls that do not possess α/β -keto acid functionalities.⁸⁹ One of these enzymes, namely AmDH4/4OP-AmDH from *Petrotoga mobilis* strain DSM 10674, displayed high activity and high enantioselectivity towards 4-oxopentanoic acid (4OP) **26** to produce (4*S*)-4-aminopentanoic acid⁸⁹ (4AP) **26a** (Figure 33). Additionally, AmDH4 showed slightly better activity towards larger substrates such as 5-oxohexanoic acid **27** over other AmDHs discovered from these rounds of sequence searches and so possibly has improved substrate promiscuity.⁸⁹ Furthermore, AmDH4 is derived from a thermophilic bacterium and so its canonical thermostability was thought to be beneficial for further searches and engineering that would ultimately be useful for industrial or synthetic purposes. Therefore, 4OP-AmDHs were selected as ideal candidates for even further sequence driven approaches to find distant homologues with activities towards unfunctionalised carbonyls.⁹¹

Enzymes discovered from sequence driven searches

Reaction catalysed by 2,4-diaminopentanoate Dehydrogenase (2,4-DAPDH)



Reductive Amination of 4-oxopentanoic acid using AmDH4 (4OP-AmDH)



* typically from NH₄Cl or NH₄HCO₂ buffers

Figure 33: Reactions catalysed by 2,4-diaminopentanoate Dehydrogenase (2,4-DAPDH) and AmDH4 (4OP-AmDH). This schematic displays the reactions catalysed by 2,4-DAPDH to form (4S)-4-aminopentanoic acid and the reaction catalysed by AmDH4 with 4-oxopentanoic acid and ammonia.

The structure of AmDH4 was more recently determined and co-crystallised with NAD⁺.⁹¹ The biological assembly of these AmDHs tend to be dimers (dimer visualisations largely mentioned in *Results and Discussion*) in which each monomer binds one molecule of NAD(P)H/NAD(P)⁺ and one molecule of substrate/product in their active site. Each monomer of AmDH4 is made up of a N-terminal Rossmann domain and a C-terminal domain comprised largely of six beta pleated sheets (Figure 34). One structure of AmDH4 (PDB: 6G1M) was resolved to give two dimers in the asymmetric unit with each dimer displaying both 'open' and 'closed' conformation monomers which is likely the result of dynamic movement on substrate binding (Figure 34).⁹¹ When modelling in the product (4S)-4AP it was clear that it is stabilised by electrostatic interactions from R161, N163 and H264, as well as Q266 and F168 which line the floor and the ceiling of the active site to close the substrate or product into the hydrophobic pocket (Figure 34). Additionally, E102 is situated at the correct proximity of the ketone, or subsequent amine group of 4OP or 4AP respectively, to facilitate the activation of the amine donor for attack at the carbonyl centre (Figure 34).

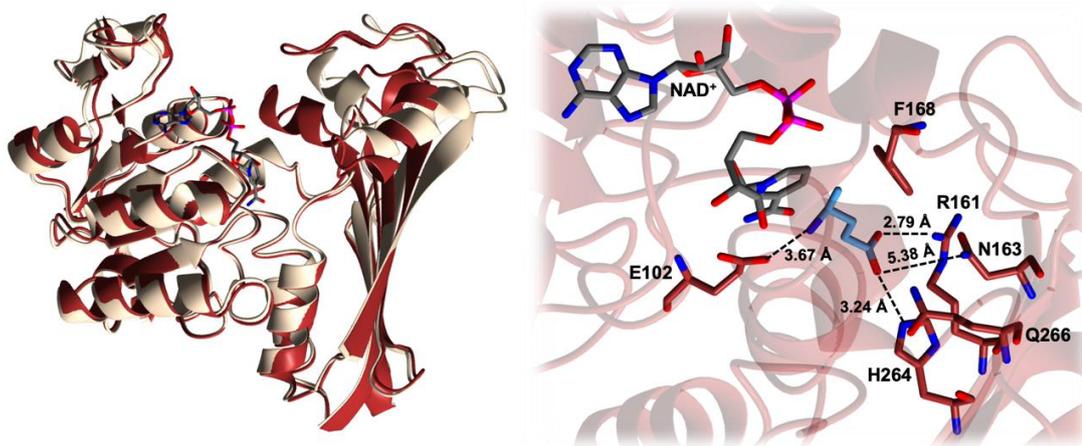


Figure 34: Open and closed forms of AmDH4 and its active site in complex with NAD⁺ and (4S)-4AP. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the monomeric forms of open (beige chain A) and closed (red chain B) conformers of AmDH4 (PDB: 6G1M) in complex with NAD⁺ (grey) and (4S)-4AP (blue). The right-hand panel represents the active site residues of chain B (red) of AmDH4. Dotted lines represent interactions between residues and the (4S)-4AP ligand (blue), all distances are measured in Ångstroms. *N.B.* (4S)-4AP was modelled from a ligand-built coot dictionary and SSM superimpositions with PDB: 6IAU.

1.3.1 *Cfus*AmDH, *Msme*AmDH and *Micro*AmDH

*Cfus*AmDH, *Msme*AmDH and *Micro*AmDH were discovered from the sequence driven search mentioned previously.⁹¹ *Cfus*AmDH was discovered from *Cystobacter fuscus*, *Msme*AmDH from *Mycobacterium smegmatis* and *Micro*AmDH from *Microbacterium* sp. MA1 and were all optimised for *E. coli* expression. Initial preliminary activity screening found that these nat-AmDHs were not active towards substrates such as 2,4-DAP **25a** but active towards other amines of interest such as cyclohexylamine **21b**. This suggests the role of these nat-AmDHs are likely to be involved in bioprocessing of amines that are very distant from the original 4OP-AmDHs. These nat-AmDHs were confirmed to be true AmDHs as their reactions were not dependent on pH suggesting they catalysed both the formation of the iminium ion as well as the reduction to the resulting amine in solution (Figure 35).

Reaction catalysed by native amine dehydrogenases (nat-AmDHs)

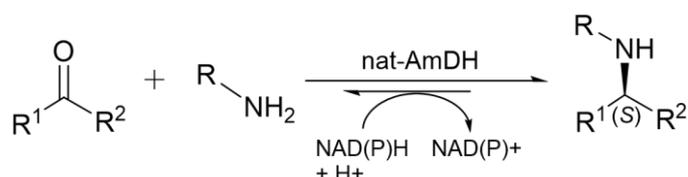


Figure 35: Reaction catalysed by native Amine Dehydrogenases (nat-AmDHs). This schematic displays the reactions catalysed by nat-AmDHs to form the (S)-chiral amines from unfunctionalised carbonyls. R¹ and R² represent a range of substituted groups where R typically represents H or CH₃.

These three enzymes were screened for carbonyl acceptor libraries as well as amine donor libraries (Figure 36). *Msme*AmDH, *Cfus*AmDH and *Micro*AmDH all portrayed similar carbonyl tolerances and (*S*)-enantioselectivities with ammonia as the amine donor, displaying preferences for cycloalkanones and aliphatic aldehydes and little or no activity was monitored with aromatic ketones; dissimilar to previously engineered AmDHs.^{73–83} The two highest specific activity hits cyclohexanone **21** and isobutyraldehyde **28** were tested against a range of amine donors specifically methylamine, ethylamine, benzylamine, cyclopentylamine and 3-pentanamine. In these cases, only methylamine was accepted as an alternative to ammonia. Albeit only accepting a limited variety of amine donors, these nat-AmDHs therefore give rise to novel reactions using primary amine donors as opposed to ammonia, which previously described initial artificial AmDHs could not do.^{77–82}

Carbonyl substrate scope of native Amine Dehydrogenases (nat-AmDHs)

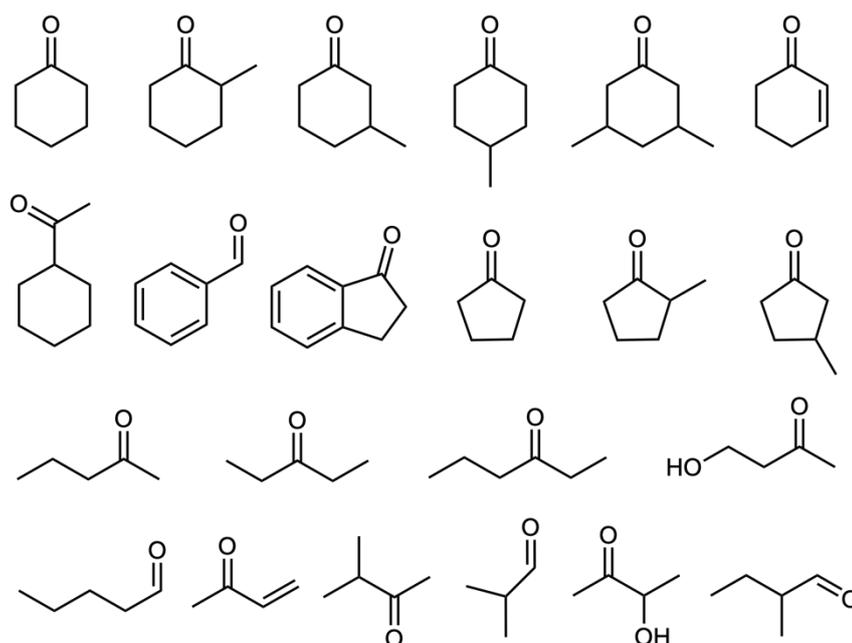


Figure 36: Carbonyl substrate scope of native Amine Dehydrogenases (nat-AmDHs). This schematic displays the selection of carbonyls tested for reductive amination using *Msme*AmDH, *Cfus*AmDH and *Micro*AmDH with ammonia.

The crystal structure of *Cfus*AmDH and *Msme*AmDH were subsequently determined by X-ray crystallography and deposited in the PDB under accession codes 6IAU and 6IAQ respectively.⁹¹ Like AmDH4 the biological assembly of both *Cfus*AmDH and *Msme*AmDH are homodimeric with each monomer being comprised of the large N-terminal Rossman domain and the C-terminal beta pleated sheet domain. Both proteins were crystallised with, and electron density was resolved for, NADP⁺.⁹¹

Unfortunately, no appropriate substrate was trapped in the *Msme*AmDH crystal structure apart from the cryoprotectant ethylene glycol.⁹¹ *Cfus*AmDH crystals managed to be co-crystallised with cyclohexylamine and the ligand bound structure was resolved.⁹¹ However, both proteins were still found in their closed conformations. Similarly, to AmDH4 the substrate cage is closed on one side by aromatic residues Y168 (*Cfus*AmDH) and W164 (*Msme*AmDH) and the NADP⁺ cofactor and closed on the other by Y173 (*Cfus*AmDH) and F169 (*Msme*AmDH) (Figure 37). Additionally, catalytic residues E108 from *Cfus*AmDH and E104 from *Msme*AmDH are thought to play the same role in canonical catalysis as E102 from AmDH2. Both catalytic glutamates are positioned at ideal distances to allow for the amine donor activation ultimately for nucleophilic attack at the carbonyl centre (Figure 37). Very recently *Msme*AmDH and *Cfus*AmDH have been scaled up for the biotransformations of a large range of substrates including (*S*)-MOIPA **36** a crucial chiral chemical in the widespread herbicide Outlook®.⁴ Additionally, the rational and random mutagenesis of *Cfus*AmDH has been conducted, parts of which will be discussed in (*Results and Discussion 4.4*).⁹²

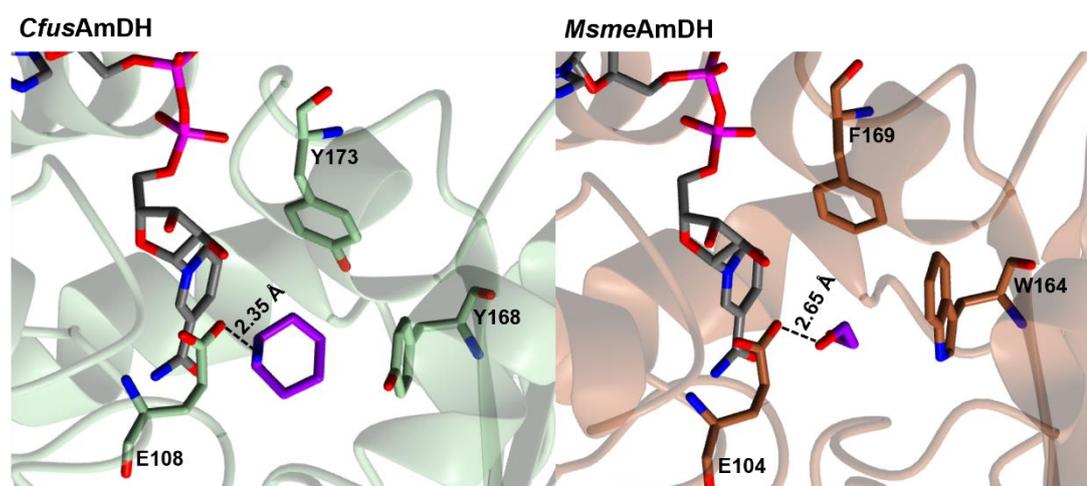


Figure 37: Active sites of native AmDHs *Cfus*AmDH and *Msme*AmDH. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the active site residues of *Cfus*AmDH (green) (PDB: 6IAU) in complex with NADP⁺ (grey) and cyclohexylamine (purple). The right-hand panel represents the active site residues of *Msme*AmDH (orange) (PDB: 6IAQ) in complex with NADP⁺ (grey) and crystallisation additive ethylene glycol (purple). All interactions are displayed by a black dashed line with distances being measured and labelled in Ångstroms. All residues are labelled accordingly.

2. Aims of the project

Thanks to the recent discovery of the nat-AmDHs *CfusAmDH* and *MsmAmDH*⁵⁶ a sequence driven search of undefined genes found in the metagenome was conducted. This metagenomic screening was extended to the Marine Atlas of Tara Oceans Unigenes (MATOUv1) and revealed several more sequences to be annotated as AmDHs. One of these enzymes was named MATOUAmDH2 and shares a high sequence identity to *CfusAmDH* and *MsmAmDH* of 33% and 32% respectively.

The aim of this PhD was to characterise a range of native AmDH's of interest both structurally and biochemically to uncover modes of binding, potential reductive amination capabilities and to use these tools for the rational engineering of a nat-AmDH with improved biocatalytic abilities. During this thesis the structural characteristics of MATOUAmDH2 and briefly other nat-AmDHs, and the rational enzyme engineering approach taken to produce a biocatalyst capable of reducing a range of ketones to produce primary and secondary amines of interest will be discussed. Much of this work is also published and available for accessing.^{92,93} The individual aims of this project are as listed:

- *Obtain the crystal structure of MATOUAmDH2 through X-ray crystallography*

At the time of this PhD start date no crystal structure for the MATOUAmDH2 enzyme was available. The first aim was to express the gene, initially in the pET22b(+) vector (optimised and provided by Genoscope) and purify the resulting protein in order to obtain, at least, the cofactor bound-crystal structure. A ligand bound crystal structure was also an attractive aim to better understand the mode of binding of this enzyme and elucidate any canonical or novel mechanistic and structural features.

- *To screen MATOUAmDH2 for carbonyl and amine donor scopes and activities*

Screening MATOUAmDH2 towards a range of substrates and donors was an additional aim to examine the reductive amination capabilities of this enzyme in relation to other nat-AmDHs of interest such as *CfusAmDH* and *MsmAmDH*. The examination of MATOUAmDH2's kinetic characteristics towards important

substrates acts as a foundation to further rational engineering of this enzyme. Additionally, the use of MATOUAmDH2 in the biotransformations of relevant substrates and donors, allowed for the initial investigation of the activity of this enzyme which could be further improved by rational engineering.

- *To engineer MATOUAmDH2 for improved substrate scope and specificity*

After structurally and biochemically characterising the wild-type enzyme the subsequent aim of this project was to then rationally engineer MATOUAmDH2 using chemically and structurally informed mutagenesis. The rational engineering of MATOUAmDH2 was predominantly to improve the substrate scope of this enzyme towards accepting larger carbonyls as well as improving overall activity. All rationally designed mutants were to be assessed by determining kinetic parameters or assessing their biocatalytic natures by performing a range of biotransformations.

- *To characterise other nat-AmDHs of interest*

A small selection of other nat-AmDHs which showed potential promising characteristics such as improved substrate scopes, improved stabilities or novel activities were to be characterised by structural resolution. Notably *MicroAmDH*, *TtherAmDH* and a mutant of *CfusAmDH* (*CfusW145A*) were to be studied independently and attempts to obtain the crystal structure(s) of these enzymes were to be made. This was to uncover the structural features that are thought to be responsible for the promising qualities that they exhibited.

3. Methods

3.1 Molecular Biology

3.1.1 Primer and Recombinant Plasmid Design

Sub-cloning of MATOUAmDH2 from a pET22b(+) vector optimised by Genescope (codon optimisation carried out by Twist Bioscience) for *E. coli* expression and LIC ligation, was carried out using the pETYSBLIC-3C vector. The pETYSBLIC-3C vector contains a cleavable extended His tag region on the N-terminus; a recombinant plasmid map, from SnapGene Viewer (v.5.0.7), of MATOUAmDH2:pETYSBLIC-3C is provided in Figure 38. All primers were manually designed in SnapGene (v.5.0.7) and Serial Cloner (v.2.6). Typically, gene specific primers were designed with ~15 bp overhangs complimentary to the vector. Vector primers were designed to be independent of gene complementarity and so without overhangs for versatility. Primers were evaluated with the online OligoEvaluator™ calculator provided by Sigma-Aldrich, paying particular attention to GC percentages, any secondary structures and ensuring gene/vector specific T_m values were within 1°C of each other in a primer pair. All primers were provided by Sigma-Aldrich at a scale of 0.025 μmol, purified using a desalt procedure being shipped as dry weight. All primers and protein sequences are detailed in *Appendices* - Table 17 and 18 respectively.

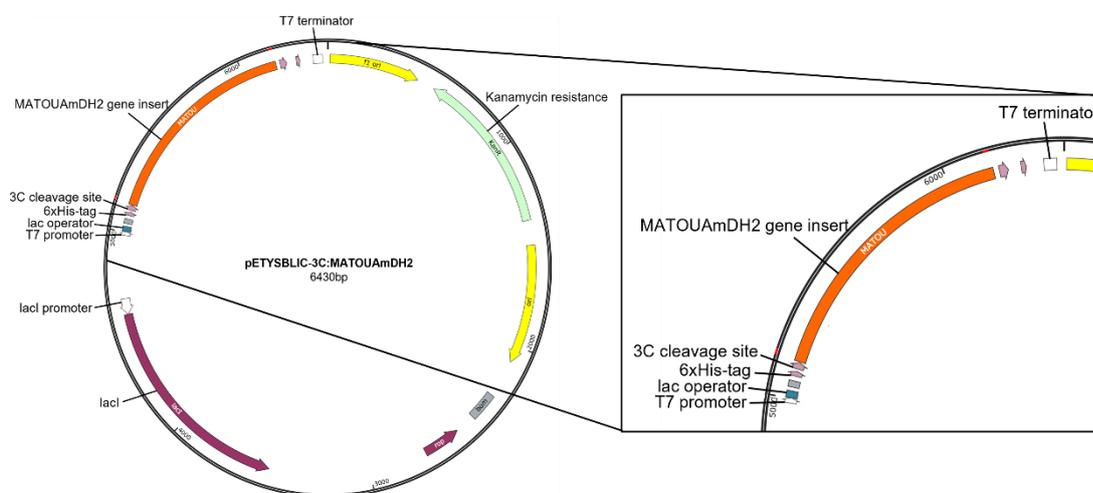


Figure 38: Recombinant plasmid map of pETYSBLIC-3C with the MATOUAmDH2 insert. Plasmid map highlighting regions of interest including the 6xHis tag N-terminal to a 3C protease cleavage site (lilac). The following features of importance: kanamycin resistance (green), lacI operon (purple), T7 promoter and terminator sites (white) and MATOUAmDH2 gene insert (orange).

3.1.1.1 Site Directed Mutagenesis Design

All mutant primers were designed in SnapGene (v.5.0.7) and Serial Cloner (v.2.6) and evaluated as above via the online Sigma-Aldrich OligoEvaluator™ calculator and provided by Sigma-Aldrich as above. Single reverse and forward codon substitutions were typically designed to approximately be in the middle of the primer which ranged from 25-33 bp long. A series of alanine mutants were designed to assess chemical functionality surrounding the ceiling and floor of the MATOUAmDH2 active site. These sites are highlighted in the structure of the binding pocket (Figure 39). Subsequently it was decided that a reduced library of point mutations would be made at two different amino acid positions; L180 and M215. Five mutants, included in the NDT codon set, for each position were designed and were as follows L180G, L180S, L180H, L180N, L180P, M215G, M215L, M215S, M215H and M215N. Mutant primers are detailed in *Appendices - Table 17*.

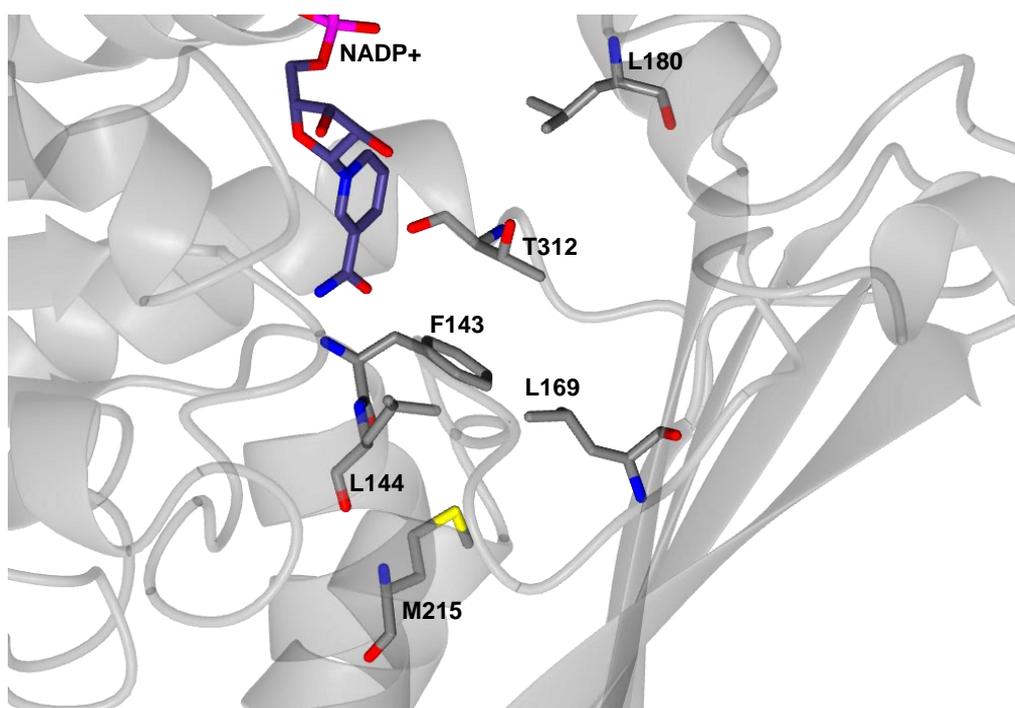


Figure 39: Structurally informed alanine-based mutant design for MATOUAmDH2. X-ray crystallography structure of MATOUAmDH2's binding pocket visualised in CCP4mg (v. 2.10.10). Residues subjected to alanine mutagenesis are labelled and visualised in cylindrical form (grey), NADP⁺ is represented in purple.

3.1.2 Polymerase Chain Reaction (PCR)

Initial amplification of the standard MATOUAmDH2 insert (in the pET22b(+) vector) was carried using the components described in Table 1, in a 50 μL PCR reaction using a KOD Hot Start DNA Polymerase kit provided by Sigma-Aldrich. For the amplification of MATOUAmDH2 insert samples concentrations were used at 20 ng. Each reaction mastermix was split into two reactions and made up to 25 μL by either the addition of just 0.5 μL KOD Hot Start Polymerase or 0.5 μL KOD Hot Start Polymerase and 0.5 μL DMSO; the polymerase typically added at the end of the reaction mixture. pETYSBLIC-3C was amplified using the following components described in Table 2, using 1 μL of pETYSBLIC-3C vector (150 ng) and Q5® High Fidelity DNA Polymerase (NEB) which was added to make up the final volumes of 2 x 25 μL reactions with one reaction having the addition of 0.5 μL DMSO. KOD Hot Start DNA Polymerase (Sigma-Aldrich) was occasionally used to amplify the vector using the same components as for the insert, but with 150 ng of pETYSBLIC-3C. Touch downs were incorporated into the PCR cycles to improve binding specificity to the region of interest. PCR cycling conditions are detailed in Table 3 for the MATOUAmDH2 insert. Cycling conditions for the pETYSBLIC-3C amplification are detailed in Table 4.

Table 1 – PCR components for the amplification of all MATOUAmDH2 based inserts

Reagent	Volume (μL)*	Final Concentrations/Quantities
10X Buffer for KOD Hot Start Polymerase	5	1X
25 mM MgSO_4	3	1.5 mM
2 mM dNTPs	5	0.2 mM
10 μM Forward Primer	1.5	0.3 mM
10 μM Reverse Primer	1.5	0.3 mM
Template DNA (MATOU568AmDH, 20 ng. μL^{-1})	1	20 ng
KOD Hot Start Polymerase (1 U. μL^{-1})	1**	0.02 U. μL^{-1}

*Total volume of 50 μL split into 25 μL volumes for each PCR reaction.

**0.5 μL for each reaction.

Table 2 – PCR components for the amplification and linearisation of the pETYSBLIC-3C vector

Component*	Volume (µL)	Final amount
Q5 Mastermix (containing polymerase)	12.5	
Forward Primer	1.25	0.5 µM
Reverse Primer	1.25	0.5 µM
LIC3C Template DNA	1	150 ng
Nuclease free H ₂ O	9	

*0.5 µL of DMSO added to one 25 µL reaction

Table 3 – PCR cycling conditions for the MATOUAmDH2 insert

	Step	Temperature (°C)	Time (s)
	Initial Activation	95	180
15 cycles	Touch Down	95	20
		86-72*	10
		72	40
20 cycles	Denaturation	95	30
	Annealing	72	10
	Extension	72	20
	Final Extension	72	90

*-1°C each cycle until a final region-specific T_m value.

Table 4 – PCR cycling conditions for the pETYSBLIC-3C vector

	Step	Temperature (°C)	Time (s)
	Initial Activation	95	180
15 cycles	Touch Down	95	20
		85-71*	10
		72	40
20 cycles	Denaturation	95	30
	Annealing	71	10
	Extension	72	150
	Final Extension	72	180

*-1°C each cycle until a final region-specific T_m value.

3.1.2.1 Site Directed Mutagenesis PCR

Initial PCRs were carried out using specific primers for each mutant and gene specific T_m values as annealing temperatures. Initial PCRs were carried out in a similar mode to the MATOUAmDH2 insert; using KOD Hot Start DNA Polymerase (Sigma-Aldrich) and similar component ratios with 20 ng of MATOUAmDH2:pETYSBLIC-3C template DNA. Initial PCR cycling conditions also included the touch down steps. However, these initial PCR conditions were found to be non-optimal for the designed mutants. PCR components were similarly adapted from the standard KOD Hot Start DNA Polymerase (Sigma-Aldrich) protocol as detailed in Table 1. Additionally the PCR cycling conditions were stripped back to remove the touchdown steps and adapted from the standard protocol, conditions being described in Table 5. Annealing temperatures were calculated as the lowest gene specific T_m value in the primer pair minus 5°C.

Table 5 – PCR cycling conditions for MATOUAmDH2 mutants

	Step	Temperature (°C)	Time (s)
	Activation	95	120
30 cycles	Denature	95	20
	Annealing	Lowest primer T_m – 5°C	10
	Extension	70	150

3.1.2.2 DpnI Digestion of PCR Products

Most DpnI digestions were carried out overnight, incubating each 25 μ L PCR reaction with 1 μ L (20 Units) of DpnI (provided by New England Biolabs®) at 37°C for 8 h with a 20 min 65°C heat inactivation step. Occasionally DpnI digestions were carried out using the CutSmart Time-Saver protocol (NEB) with 40-50 Units of DpnI (NEB) added to ~6.5-8.5 ng of DNA and appropriate amounts of CutSmart Buffer (NEB). The Time-Saver protocol digestion reactions were incubated at 37°C for 20 min with a heat inactivation at 65°C for 20 min.

3.1.3 In-Fusion® Cloning and Stellar™ Competent Transformations

In-Fusion cloning relies on the ability to fuse overhangs of around 15bps long with complementary regions using their In-Fusion ligation enzyme. The overhangs in this case were to the pETYSBLIC-3C site just before the 6x His Tag or just before the T7 tag site. This cloning region was amplified and linearised through PCR and overhangs attached to the gene of interest in this case MATOUAmDH2. In the case of mutant design, the 'overhangs' are complementary sequences from the recombinant plasmid containing a point mutation, generated using PCR and DpnI digestion within the gene of interest itself which self-anneal via In-Fusion ligation. Figure 40 shows a schematic of these procedures.

The cloning reaction was carried out using the Takara Bio In-Fusion® HD Cloning Kit and manual. The reaction was carried out after gel extraction and spin column PCR clean-up (*Methods - 3.1.4*). Typically, 50 ng of insert DNA was used and a mole-based ratio (considering the differences in molecular weights) of 2:1 of insert to vector was used to calculate the final amount of vector needed. This varied depending on how much DNA was available after PCR purification but typically fell within the manual guidelines for both the vector and insert sizes. Normally reactions were carried out using 2 µL of 5X In-Fusion HD Enzyme Premix (containing the ligation enzyme) and appropriate volumes of insert and vector made up to 10 µL with dH₂O. The reactions were then incubated at 50°C for 15-20 min and placed on ice until transformation.

Transformation of Stellar™ Competent cells (*E. coli* HST08), provided by Takara Bio, using the In-Fusion generated recombinant plasmids were performed; the procedure was followed as per the In-Fusion® HD Cloning Kit User Manual. 2.5 µL of each In-Fusion reaction were added to 50 µL of thawed cells and left on ice for 30 min for stabilisation of the plasmid and cells. The cells were then heat shocked at 42°C for 45 s to allow uptake of the recombinant plasmid. They were then briefly placed on ice for a few minutes before adding prewarmed media, typically SOC, until a total volume of 500 µL. The cells were then incubated at 37°C for 1 h, with shaking around 180-220 rpm. Plating was performed via two methods, one sample would just involve taking 100 µL of the culture after incubation and plating. For the other method the remaining culture (400 µL) would be spun for 5 min at 6,000 rpm and then the supernatant would be removed. The pellet would be resuspended in 100 µL of fresh SOC media and the remaining resuspension would be plated. Samples were plated on agar plates containing 30 µg.mL⁻¹ kanamycin, and incubated overnight without shaking, at 37°C.

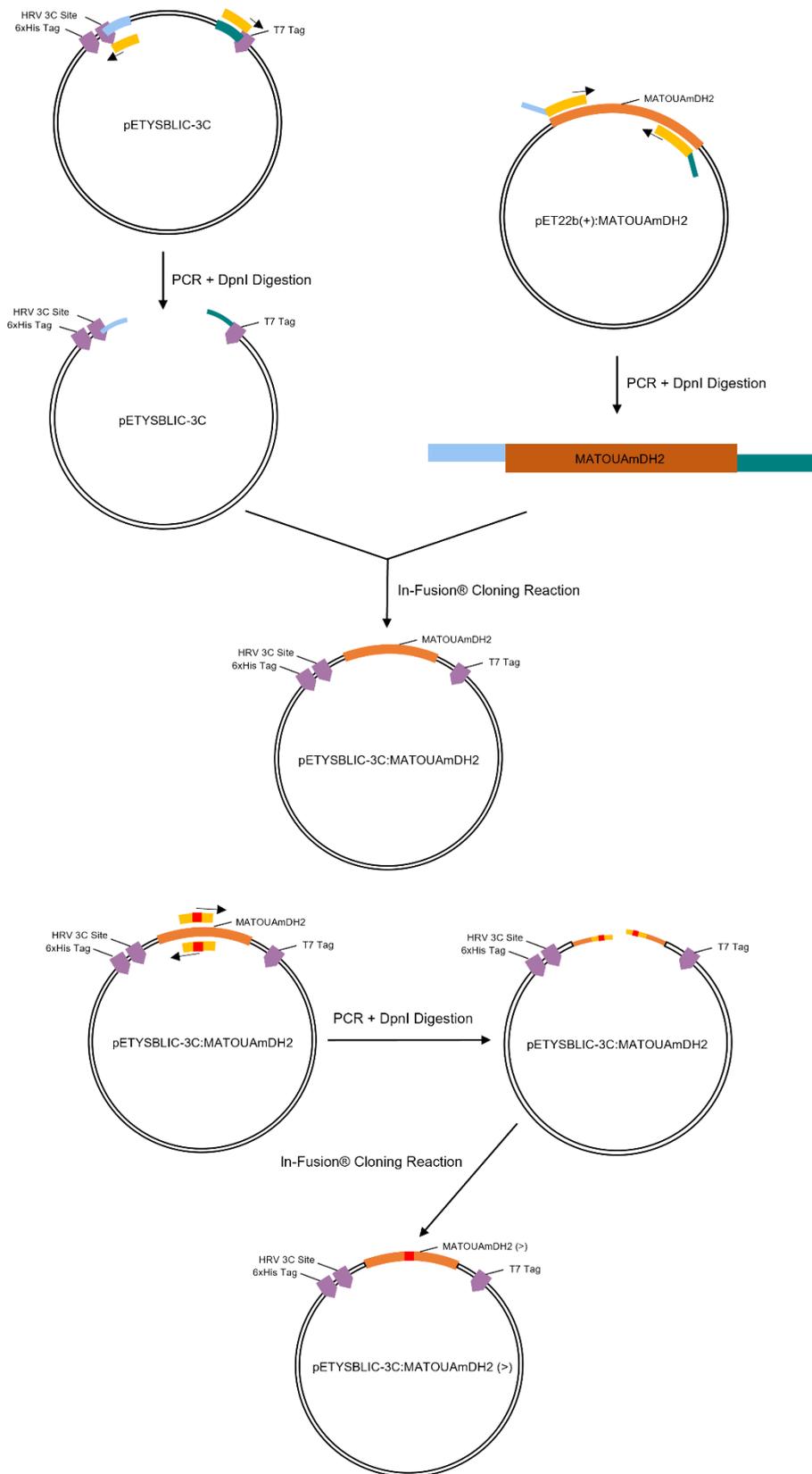


Figure 40: Schematic of the generation of MATOUAmDH2 based recombinant plasmids via PCR, DpnI digestion and In-Fusion® Cloning. Specific regions of interest for highlighting overall positioning of inserts/primers are in purple e.g., 6xHis Tag. Primers are highlighted in yellow. MATOUAmDH2 genes are in orange. Overhangs are highlighted (for the pETYSB LIC-3C vector in green and blue, yellow for MATOUAmDH2 mutants). Mutation sites are highlighted in red. > denotes a genetic change from one amino acid to another, i.e., generation of a point mutant.

3.1.4 Confirmation of Fragments via Gel Electrophoresis

Gel electrophoresis is regularly used to separate and analyse DNA bands based on molecular weights and morphologies. A current is applied across a gel bed and DNA moves away from the negative electrode to the positive electrode due to the negatively charged backbone properties of DNA. The extent of DNA migration through the gel is dependent on molecular weights; with larger bands moving more slowly and migrating less, and smaller bands moving faster and migrating a further distance. Typically, 1% gels were made up with agarose and 1x TAE buffer. A 1:10,000 dilution of SYBR® Safe DNA stain was added to each gel before pouring and allowing to set. After loading samples and DNA ladders, typically 1 kb or 100bp ladders (NEB), agarose gels were then run for approximately 50-70 min at 110V with 1x TAE running buffer. Agarose gels were visualised using a Safe Imager™ transilluminator and imaged, any bands of interest were excised using sterile scalpels. Bands were then spin column purified using the Monarch® Gel Extraction Kit provided by NEB following the standard protocol and eluting with DNA elution buffer (T1016-3) or dH₂O.

3.1.4.1 Double Digests

Double digests were carried out to assess the presence of MATOUAmDH2 inserts within the recombinantly designed plasmids. There are many restriction sites in the pETYSBLIC-3C plasmid, but only restriction sites around the cloning region were considered. WebCutter 2.0 was used to assess any possible overlapping restriction sites within the MATOUAmDH2 insert. Restriction endonucleases NdeI and XbaI were decided on based off their commercial availability and compatibility with each other. A double digest protocol was followed from the NEBcloner® online tool, using the Time-Saver CutSmart buffer protocol. 20 U of each restriction enzyme was added to 1 µg of recombinant plasmid DNA with 5 µL of 10X CutSmart buffer (1X final) and made up to 50 µL with dH₂O. Each reaction was incubated at 37°C for at least 1 h with an additional 20 min 65°C heat inactivation step. The resulting reactions were run on 1% agarose gels, as above.

3.1.4.2 Minipreps and Sequencing

After Stellar™ Competent cell transformations (*Methods 3.1.3*) single colonies were selected, and overnights were produced by following the protocol below in *Methods 3.1.6*. After incubating overnight at 37°C with shaking (180-220 rpm), 2 mL of each

overnight was aliquoted out into Eppendorf tubes and subjected to the Monarch® Plasmid Miniprep Kit to isolate high yields of the recombinant plasmid. Plasmid DNA was eluted with 30 μL of DNA Elution buffer provided by the kit. Purified recombinant plasmids were then sent off to sequencing to Eurofins Genomics for Sanger Sequencing using their SupremeRun Tube service. Primers that were used for sequencing were the T7 (TAATACGACTCACTATAGGG) primer and the pRSET-RP (ATGCTAGTTATTGCTCAGC) primer.

3.1.5 BL21(DE3) Transformations

BL21(DE3) is a common *E. coli* strain regularly used for the expression of recombinant proteins where the cloning site is typically downstream of the T7 promoter. BL21(DE3) is an expression host which contains a λ prophage harbouring a copy of the T7 RNA polymerase gene. When this gene is induced it transcribes and translates to produce T7 RNA polymerase which in turn binds to the T7 promoter of the plasmid and transcribes the gene and of interest at the site of cloning insertion and thus produces the recombinant protein. The T7 RNA polymerase gene is under control of the lac operon, which in turn can be induced by isopropyl β -D-1-thiogalactopyranoside (IPTG). IPTG acts as a molecular mimic that can bind to the lac repressor which releases and triggers the transcription of genes in the lac operon. Therefore, BL21(DE3) was a compatible strain for the production of both pET22b(+) and pETYSBLIC-3C constructs. Typically, transformations were performed after miniprepping where 1-50 ng of plasmid DNA (usually 1 μL) was added to 25 μL of BL21(DE3) cells, which had been thawed on ice, and incubated on ice for 30 min. Cells were then heat shocked at 42°C for 45 s and then immediately placed on ice for approximately 2 min. 1 mL of Luria Bertani (LB) media was added, and the Eppendorf tubes were placed at 37°C for 1 h with shaking (180-220 rpm). Usually, 100 μL of culture were taken and plated, sometimes additional methods such as centrifugation (6,000 rpm for 5 min) of the remaining solution and resuspension of the pellet in fresh LB (100 μL) would be carried out and plated separately. All samples were plated on agar plates with appropriate antibiotics (100 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin for pET22b(+)) or 30 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin for pETYSBLIC-3C constructs) and left overnight at 37°C without shaking.

3.1.6 Expression of Recombinant Genes

Single colonies after plate transformations were selected using sterile pipette tips. These colonies were transferred into falcon tubes containing 10 mL of LB media and appropriate antibiotics (100 $\mu\text{g}\cdot\text{mL}^{-1}$ ampicillin for pET22b(+)) or 30 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin for pETYSBLIC-3C constructs). These cultures were then shaken at 180-220 rpm overnight at 37°C. After overnight incubation, all 10 mL overnights were transferred into 2.5 L shake flasks containing 1 L of LB media and appropriate antibiotics. Cells were shaken at 180-220 at 37°C until their optical density (OD_{600}) was between 0.6-0.8 as recorded by UV spectrometry. This is density range at which the *E. coli* cells are in the exponential (log) phase of growth, which is optimal for stable recombinant protein production. After this optical density was reached the cells were induced with IPTG at a final concentration of 1 mM. The cells were then incubated at 16°C overnight or upwards of 20 hours to allow for sufficient protein expression. In relation to recombinant MATOUAmDH2 mutants, used for cell lysate production (*Methods* 3.5.2), it was ensured that all cultures were induced at roughly the same OD_{600} value; within 0.1 of each other to allow for similar overall cell pellet weights.

3.2 Protein Production and Purification

3.2.1 Protein Production

After gene and protein expression the cells were collected and then harvested using a Lynx6000 centrifuge (Thermo Scientific), centrifuging at 9760 g for 20 min at 4°C. The cell pellets were collected and resuspended in 25-100 mL, depending on cell pellet size, using a 50 mM Tris-HCl (pH 7.1) and 300 mM NaCl containing buffer. For cell pellets concerning the production of mutant cell lysates (*Methods* 3.5.2) the cell pellets were weighed and 10 mL of buffer per 1 g of cell pellet was added prior to lysis. Usually, the last 10 mL of buffer was reserved for use in the cell disruptor. Cell pellets were homogenised in the appropriate amount of buffer by stirring on ice or vortexing. Once a homogenous solution was achieved cells were lysed and disrupted using a cell disruptor homogenizer (Constant Systems Ltd) at a pressure of 26-27 kPsi. 10 mL of buffer was typically used at the end of lysis to flow through the system, increasing the final yield of lysate. Lysed cells were then centrifuged at 38,571 g for 40 min at 4°C and the supernatant retained for enzyme purification. In the case of the

mutants which just concerned the use of cell lysate, the lysate was filtered through a 0.2 μM or 0.45 μM filter and then aliquoted into 2 mL volumes and flash cooled using liquid nitrogen.

3.2.2 Protein Purification

3.2.2.1 Nickel Affinity Chromatography

Nickel Affinity chromatography is a commonly used method to purify His-tagged proteins. The resin matrix within these columns consists of a nickel ion (Ni^{2+}) which has been coordinated to nitrilotriacetic acid (NTA). The side chains of histidine (imidazole) within the polyhistidine tag can chelate to these Ni^{2+} ions. So, when a column resin is charged with nickel, for example from NiCl_2 , a His-tagged protein may bind to this resin. Using this type of chromatography, one can control where a specific protein is eluted. Using imidazole, which competitively binds to the resin, along with the protein of interest, protein can be selectively eluted from the column at different imidazole concentrations.

Cell lysates were filtered through a 0.45 μM filter before being loaded onto a 5 mL FF HisTrap™ column (GE Healthcare) preloaded with nickel chloride. An initial 7 column volumes (CV) of buffer containing 50 mM Tris-HCl (pH 7.1) and 300 mM NaCl (Buffer A) was passed through the ÄKTA (GE Healthcare) system equipped with the HisTrap™ column. These first steps can be sent to waste, but they were usually collected in falcon fractions. After this step an initial wash with a buffer containing 50 mM Tris-HCl (pH 7.1), 300 mM NaCl and 500 mM low UV imidazole (Buffer B) at 5% (25 mM) was employed for 5 CV to wash off any unspecific bound proteins. For the next 10 CV an imidazole gradient was employed from 25 mM to 500 mM (100% Buffer B) to elute any bound protein of interest. The last 8 CV employed 100% Buffer B (500 mM imidazole) to wash off any excess bound protein. A schematic in Figure 41 below explains typical the steps employed on the ÄKTA system to elute protein. These steps were typically collected in 96 deep well fraction collectors equipped to the ÄKTA system. During the imidazole gradient a peak should appear corresponding to the A_{280} absorbance, from this UV absorbance protein concentrations can be estimated.

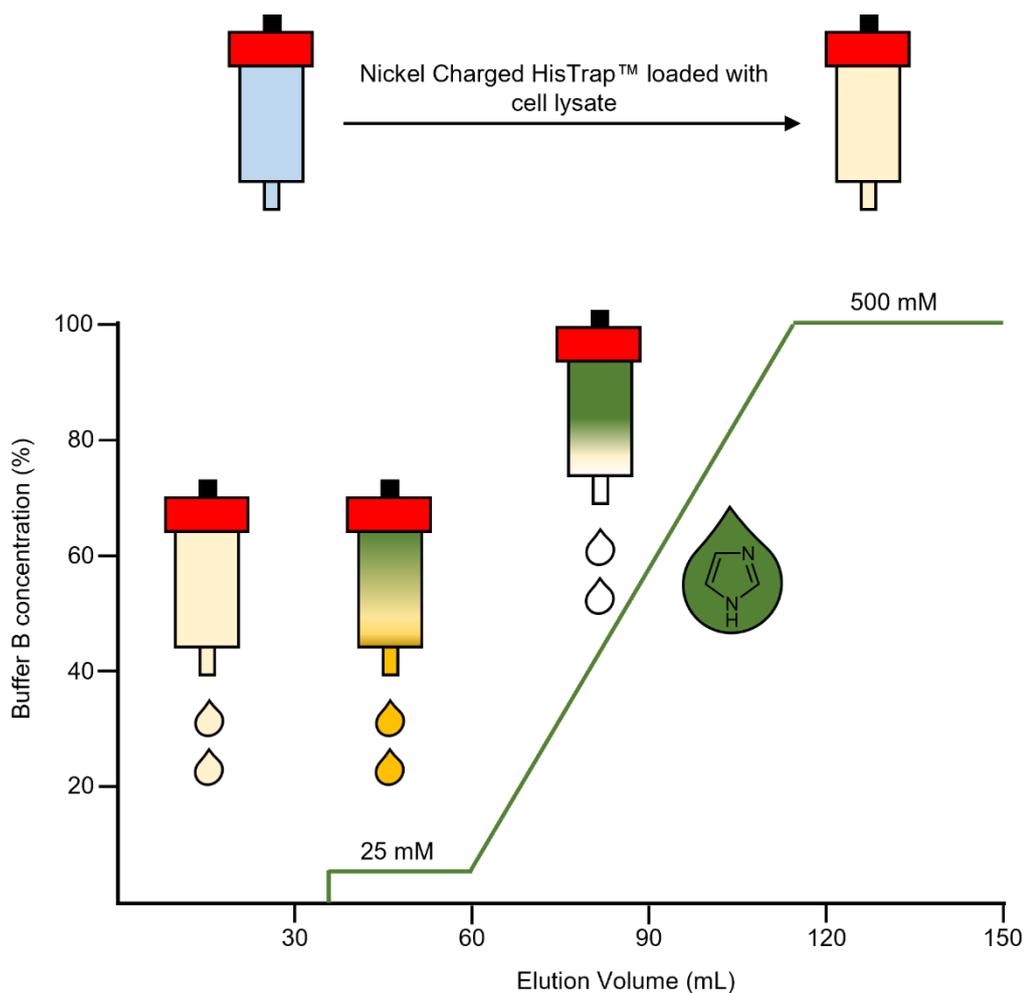


Figure 41: Scheme of the steps involved in Nickel Affinity Chromatography. The x-axis displays the elution volume (mL) of sample being collected, the y-axis displays the Buffer B concentration (%). The green line represents the imidazole gradient employed with labelled features highlighting specific imidazole concentrations at each step (mM). Green is a reference for the imidazole containing buffer where yellow represents bound proteins with darker colours in reference to more unspecific bound proteins. N.B. that colours are for explanatory purposes only.

3.2.2.2 Size Exclusion Chromatography (SEC)

Size exclusion or gel filtration (GF) chromatography is another commonly used method to purify proteins utilising differences in molecular sizes. The resin matrix in these columns contain beads which contain different pore sizes. As proteins move through the matrix smaller proteins or oligomers can occupy both the larger space between the beads and the pores within the beads. Larger proteins or oligomers can only pass-through gaps big enough, usually the space between the beads. As the larger moieties can only occupy a smaller volume of space, they elute the quickest with smaller moieties taking more time to pass through the column and all the possible routes they can occupy. A simplified diagram of how SEC works is shown in Figure 42 below.

After analysing protein fractions via SDS-PAGE (*Methods 3.2.3*) and selecting the fractions containing the protein of interest the fractions were combined and concentrated to a volume of 2-4 mL through a PES 10,000 MWCO concentrator (Vivaspin®). The resulting protein sample was filtered through a 0.2 µm filter before being loaded onto a Superdex® High Load® 16/600 gel filtration column (120 mL) equipped on an ÄKTA (GE Healthcare) system. Samples were run for 1-2 CV depending on the elution time of the protein. Eluted fractions were collected in a 96 deep well fraction collector equipped to the ÄKTA system.

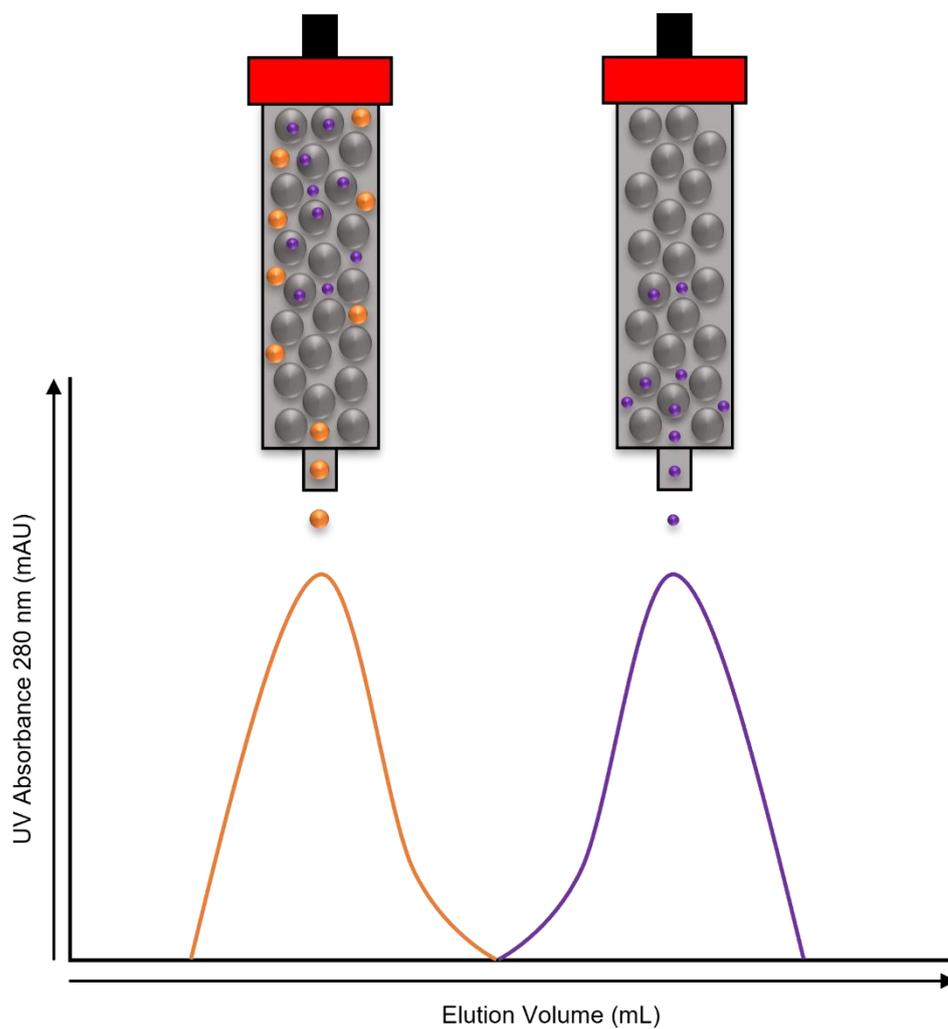


Figure 42: Scheme of the steps involved in Size Exclusion Chromatography (SEC). The x-axis represents the increasing elution volume (mL) with the y-axis indicating the increasing UV A_{280} absorbance (mAU). Columns and bead matrixes are represented in grey. Smaller moieties are represented in purple and larger ones in orange, similarly with the peaks on the chromatogram.

3.2.3 SDS-PAGE Analysis

SDS-PAGE is a technique routinely used to separate proteins based off molecular weight solely without any influence from charge or structural properties. The sodium dodecyl sulfate (SDS) is responsible for concealing the proteins native charge by surrounding the polypeptide backbone, giving all proteins in the sample a uniform charge-mass ratio. Proteins can then move through stacking and separating gels where they are separated based on how fast they move through the polyacrylamide matrix; with smaller proteins moving faster and migrating further than larger ones.

For the sample preparation for use in SDS-PAGE analysis, protein fractions were collected and diluted to appropriate concentrations and a reducing SDS loading dye was added so its final concentration was 1X. Reducing SDS loading dye not only contains SDS for mitigating charges but also thiols such as β -mercaptoethanol which reduces and cleaves disulfide bonds between cysteine residues in the secondary and tertiary structures. Once the dye was added the samples were then incubated at 95°C for 5 minutes to disrupt and denature bonding such as hydrogen bonding between the secondary and tertiary structures. Samples were typically loaded in 10-15 μ L volumes and a pre-stained, broad range, (250-10 kDa) ladder, provided by ThermoFisher, was loaded (5 μ L). Samples were then run for approximately 50-80 min at 200 V on a 12% gel with running buffer. Tables 6, 7 and 8 show the components of the loading dye, running buffer and 12% SDS gel matrix.

Table 6 – Ingredients for 10 mL of 2X SDS Loading Dye

Component	Volume (mL)
Deionised H ₂ O	4.8
0.5 M TRIS HCl pH 6.8	1.2
Glycerol	1.0
SDS (10%)	2.0
Bromophenol blue (1%)	0.5
β -mercaptoethanol	0.5

Table 7 – Ingredients for 1 L of ready to use Running Buffer

Component	Volume (mL)
Deionised H ₂ O	250
4 x Running Buffer (0.1 M TRIS Base, 0.75 M Glycine)	740
SDS (10%)	10

Table 8 – Ingredients to make 2x 12% SDS gels

	Component	Volume (mL)
Resolving gel layer	Deionised H ₂ O	3.2
	Resolving gel buffer (1.5 M TRIS pH 8.8, 0.4% SDS)	2.5
	Acrylamide stock	4.2
	APS (10%)	50 µL
	TEMED	8-10 µL
Stacking gel layer	Deionised H ₂ O	3.2
	Stacking gel buffer (0.5 M TRIS pH 6.8, 0.4% SDS)	1.3
	Acrylamide stock	0.5
	Bromophenol blue*	10 µL
	APS (10%)	25 µL
	TEMED	8-10 µL

3.3 Crystallisation and structural studies

3.3.1 Generating Crystals

Protein bands were analysed via SDS-PAGE and fractions containing the correct bands were pooled and concentrated using Vivaspin® concentrators with a 10,000 MWCO PES filter. Concentrations were determined using a NanoPhotometer (Implen). This records the absorbance at 280 nm which corresponds to the absorbance of tryptophan and tyrosine residues; the concentration output accounts for the average occurrence of these amino acids in proteins. Typically, protein was

then incubated with 4-10 mM NAD(P)⁺/NAD(P)H (final concentration is 2 mM in sub-wells) and placed on ice for 30 min. Prior to dispensing, the protein:cofactor complexes were spun in a table top centrifuge at 6,000 -13,000 rpm for 60 – 90 s to sediment any aggregates. All crystal trays were dispensed in a sitting-drop format all dispensed by a Mosquito robot (TTP LabTech).

3.3.1.1 Initial Screening

To assess crystallisation compatibility protein samples were subject to pre-dispensed screens. Screens were in MRC 2 drop 96 well format and pre-dispensed solutions were provided by Hampton Research or Molecular Dimensions. Namely the Index™ screen and the Clear Strategy™ screen (CSS) were routinely used. In the case of the CSS screen one buffer was selected for wells 1-6 and another for wells 7-12, usually varying in pH. The pH values of the buffers were selected to fit below and above the proteins isoelectric point (pI); calculated by using the ExpASy online tool. In the case of the CSS screen 6 µL of appropriate buffer (usually 1 M) was added to 54 µL of pre-dispensed solution and mixed via pipetting prior to dispensing. Protein complexes were dispensed in a 1:1 ratio with the reservoir mother liquor, usually 150 nL:150 nL. Usually sub well 1 contained a different protein concentration from subwell 2.

3.3.1.2 Optimisations

Optimisations were designed to improve crystal quality from initial screening. Depending on each sub-well drop condition which provided adequate crystal forms screens were then designed to fit either a 48-well (MRC) or in some cases a 96-well format (MRC 2-drop). Typically, optimisations consisted of varying a condition in one direction across from wells 1-6/12 and varying another condition across A-H. Some conditions would then be kept consistent throughout the whole plate and made up to appropriate final volumes with water (150 µL for 48-well format and 80 µL for 96 well format). See Figure 43 for a diagram depicting a typical crystal screen optimisation set up and design. Protein complexes were dispensed in a 1:1 ratio of protein sample to mother liquor (800 nL:800 nL for 48-wells and 150 nL:150 nL for 96-wells). Often protein and ligand concentrations would be varied to also improve crystal quality during optimisations. Specific conditions for all crystals obtained are aptly listed in *Results and Discussion 4.1-4.4*.

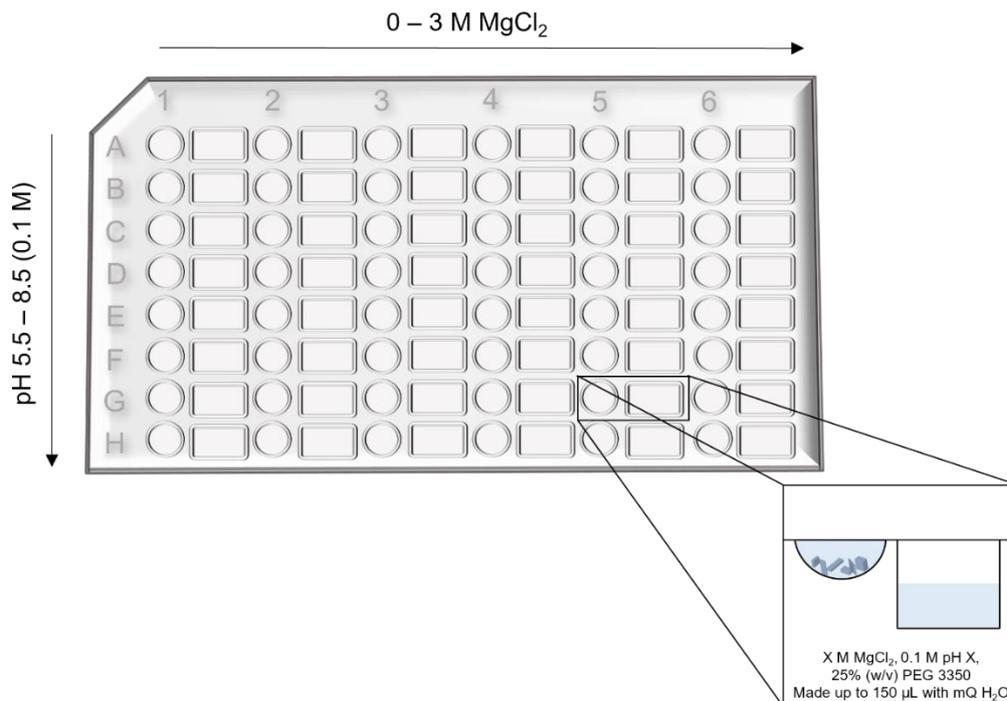


Figure 43: Diagram depicting a typical 48-well optimisation screen design. This is a cartoon representation of a 48-well plate (MRC) with one sub well for each reservoir. Columns are numbered through 1-6 and rows numbered A-H. G5 has been magnified to show the sitting-drop vapour diffusion set up with crystals depicted in the sub-well drop. N.B. This is typical design; precipitant components, concentrations and pH values were all subject to change as well as overall reservoir volumes.

3.3.1.3 Fishing

Crystals of interest were fished with synthetic loops varying in diameter from 0.025 – 1 mm equipped to magnetic rod. Usually, 0.5 μ L of mother liquor from the drop's reservoir was added to the sub-well prior to fishing to get rid of any film or for better separation of stacked crystals. Single crystals were selected and fished without the addition of cryoprotectant due to the presence of PEG 3350 in most samples. Crystals were then quickly vitrified using liquid nitrogen. Crystals were stored under liquid nitrogen in dewars before being sent to the Diamond Light Source synchrotron.

3.3.2 Co-crystallisation

Co-crystallisation i.e., incubating the apo protein, in solution, with ligand prior to dispensing is an ideal way to attempt to get ligand bound protein structures. Particularly, with AmDHs, which undergo dynamic conformational changes on ligand binding, this reduces the possibility of the crystal dissolving, which may happen during ligand soaking. Prior to dispensing, where applicable, the protein:cofactor complexes were also incubated with ligand on ice for the 30 minute duration. In the case of most ligands, they were added at a concentration of 10 mM, for most ligands they were in

a solution of DMSO usually at a stock concentration of 1 M. In the case of 2-aminonorbonane **29**, the salt (HCl) was dissolved in water, some stock was also quenched and extracted into ethyl acetate. Both the water dissolved, and ethyl acetate dissolved stocks were trialled during crystallisation.

3.3.2.1 Ligand Soaking

Ligand soaking is an alternative method to co-crystallisation where you can bind a ligand in situ in the preformed crystal. This was particularly useful when the apo protein formed large single crystals during initial screening trials. Ligand soaking is routinely done by adding some ligand stock to the specific mother liquor condition in the reservoir of the crystal drop. Typically, this solution would also need to contain cryoprotectant before being added to the crystal in the sub-well, to prevent damage when flash freezing with liquid nitrogen. In the case of the *CfusAmDH* W145A crystals these were soaked with a solution of the specific mother liquor, which already contained 25% PEG 3350, with an added 10 mM of *n*-pentylamine **30** and left to incubate for 30 min prior to fishing.

3.3.3 Data Collection and Processing

Some crystals were initially tested in-house on a Rigaku Micromax-007HF fitted with a Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. Crystals that diffracted in-house to a resolution of $< 3 \text{ \AA}$ were sent to the Diamond Light Source synchrotron. Often, crystals were not tested for diffraction in-house and sent to the Diamond Light Source regardless. Data sets were typically collected on the i03 or i04 beamlines. Data was integrated and scaled at the Diamond Light Source using their Xia2 pipeline, namely the DIALS,⁹⁴ XDS⁹⁵ and SCALA⁹⁶ frameworks. Normally the merged free MTZ file generated was imported into the CCP4i2 (v.7.1 and v. 8.0) interface. Data reduction was then carried out on this imported MTZ file using the Aimless⁹⁷ program and the resolution of the data set was manually cut off at an appropriate value. The specific protein sequence was also imported into CCP4i2 and the number of copies in the unit cell and Matthews probability was calculated.

3.3.4 Molecular Replacement and Refinement

All data and refinement statistics are tabulated and detailed in *Results and Discussion* 4.1-4.4. All the structural resolution, building and refinement was normally carried out in the CCP4i2 (v.7.1 and v.8.0) interface. In X-ray crystallography molecular replacement is a useful tool to solve the phase problem. Molecular replacement uses a known structure as a model to best fit the experimental intensities and densities and generates a structural solution. For molecular replacement to generate a reliable solution the model used must be structurally similar to of the protein of interest. The software used for molecular replacement in this case was MOLREP.^{98,99} Models used for the MOLREP software were usually monomeric proteins without any water or ligands/cofactors present generated manually from known PDB coordinates. In the instance of the apo MATOUAmDH2 structure and *Cfus*AmDH W145A structures monomeric wild type *Cfus*AmDH (PDB: 6IAU) was used. In the case of MATOUAmDH2 structures co-crystallised with cyclohexylamine a monomeric form of the apo MATOUAmDH2 structure served as the model. In the case of the MATOUAmDH2 M215S structure co-crystallised with 2-aminonorbonane **29**, the cyclohexylamine **21b** bound structure served as the model. Once MOLREP had generated an appropriate structural solution the structure was subjected to an initial refinement cycle using REFMAC5.^{100,101} After the initial refinement initial building of chains were completed manually in COOT.¹⁰² The structure was then exposed to iterative cycles of REFMAC5, usually 20-30 cycles, using local non-crystallographic symmetry (NCS) restraints. Building was carried out in COOT sometimes using assistance from the SSM superimposition tool for appropriate conformers. Once the protein backbone was built, residual density in the omit maps ($F_o - F_c$) were present for cofactors/and ligands. NAD(P)⁺/NAD(P)H was built by fetching the molecule from the known libraries in COOT. Typically, ligands were built using the SMILES format and the cif file was fitted and modelled into the omit electron density where applicable. Waters were modelled using the inbuilt Find Waters function in COOT. Waters, cofactors and ligands were refined within the structure using REFMAC5.

3.3.5 Deposition and Visualisation

For some of the structures discussed here coordinates and structure factor files were deposited in the Protein Data Bank (PDB) under accession codes as follows; apo-MATOUAmDH2 (NADP⁺); 7ZBO, MATOUAmDH2:NADP⁺:cyclohexylamine; 7R09,

*Cfus*AmDH W145A NAD⁺; 7QZN, *Cfus*AmDH W145A:NADP⁺:*n*-pentylamine; 7QZL.
All visualisations and structural analyses were done using CCP4MG (v.2.10.11).

3.4 Kinetics

3.4.1 Spectroscopically Monitoring NAD(P)H Oxidation

NAD(P)H and NAD(P)⁺ possess different spectroscopic characteristics. Both structures are detailed in *Appendices* 119. Due to the additional lone pair of electrons in NAD(P)H the reduced form is in a much lower energy state and therefore absorbs lower energy than the oxidised form, NAD(P)⁺. This is evident by an absorbance spectrum of NAD(P)H where there are two absorbance maxima, one at 250 nm and one at 340 nm, NAD(P)⁺ only has one absorbance maxima at 250 nm. During the accumulation of product during the reductive amination NAD(P)H is also oxidised and NADP⁺ accumulates. We can take advantage of the differences in spectroscopic properties of NAD(P)H/NAD(P)⁺ and monitor the progression of reductive amination at a wavelength of 340 nm. As the reaction progresses the absorbance in the 340 nm range should decrease overtime which we can record using a UV-Vis spectrometer. Figure 44 shows a graphical representation of the basis of this experimental design.

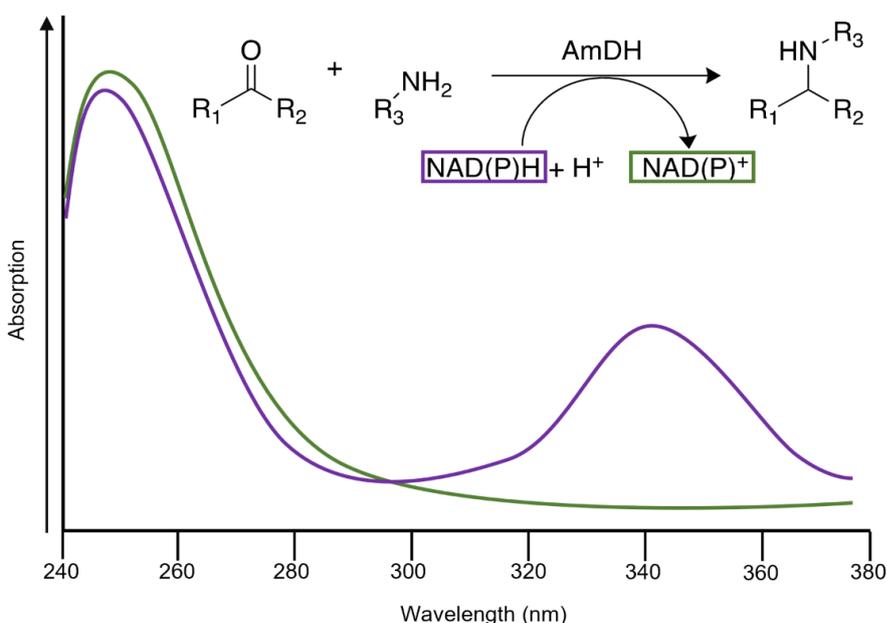


Figure 44: Experimental design for monitoring reductive amination using NAD(P)H-dependent UV-vis spectrometry. This is a graphical representation of the differences in NAD(P)H (purple) and NAD(P)⁺ (green) spectra. A second absorbance maxima peak is present for NAD(P)H at around 340 nm. A schematic of the general reaction catalysed by AmDHs also features with reduced and oxidised cofactor labelled by corresponding colours.

3.4.2 pET22b(+) MATOUAmDH2 Kinetics

Kinetic experiments were largely based on methods used to obtain data as reported previously.¹⁰³ The NADPH tetrasodium salt discussed here was supplied by Roche, all other chemicals supplied by Sigma Aldrich. Protein concentrations were confirmed using the use of a NanoPhotometer (Implen). Kinetic monitoring of NADPH oxidation was recorded at range of cyclohexanone **21** concentrations. The total volume of the reaction in the cuvette was 1 mL made up to the final volume with 2 M Ammonium Formate buffer pH 8.0. The UV-Vis spectrometer was blanked with 1 mL of 2 M Ammonium Formate buffer pH 8.0. Cyclohexanone concentrations were varied from 0-20 mM (the following final millimolar concentrations were selected at 0.25, 0.5, 0.75, 1, 2, 2.5, 5, 10 and 20, 0 mM reactions were also recorded without the addition of cyclohexanone). It was made sure that the percentage DMSO was kept constant throughout the reaction at 10% by making up cyclohexanone stocks with 100% DMSO and keeping the volume of cyclohexanone stock added to the cuvette constant, at 10 μ L each. After the addition of cyclohexanone to buffer, the cuvette was incubated in the block holder at 50°C for a few minutes and then NADPH added and incubated again. NADPH concentrations were kept constant at 0.3 mM (3 μ L). 0.025 mg.mL⁻¹ of MATOUAmDH2 (25 μ L) was added to initiate each reaction and then briefly mixed. The reactions were recorded at a block temperature of 50°C scanning for approximately 4-10 min based on the observed plateau time. Spectra were monitored at 340 nm. Initial gradients used to calculate velocity are graphed in *Appendices 120*.

3.4.3 pETYSBLIC-3C MATOUAmDH2 and Mutant Kinetics

The set up for MATOUAmDH2 expressed in the pETYSBLIC-3C vector and mutants and F143A, L144A, L169A, L180A, M215A and T312A were the same as above. The only differences were the block temperature, which in this case was set to 25°C and the protein concentrations were increased ten-fold for mutants to 0.25 mg.mL⁻¹ and to 0.20 mg.mL⁻¹ for wild type enzyme. Unfortunately, F143A, L144A and L169A all had precipitation and aggregation problems during protein purification and when isolated, non-aggregated, protein was used it still precipitated in the UV-vis cuvette rendering data collection non-viable. Initial gradients used to calculate velocity are graphed in *Appendices 121*.

3.4.4 Data Processing

For data processing straight line graphs were cut off usually around 2-5 min, again based on the observed plateau. The gradients were calculated and normalised to 0 mM gradients. The inverse of the gradients were then used to calculate the velocity from the molar extinction coefficient of NADPH. The final velocity values (y) were plotted against concentration of cyclohexanone (x) (in mM) and Michaelis Menten parameters were calculated and fitted using the inbuilt function in OriginPro® (OriginLab, v.2021) (Equation 1). Standard errors were also calculated in OriginPro® from an average of three repeats for each reaction and y-error bars were plotted. The k_{cat} was calculated from the V_{max} and total enzyme used in the reactions.

Equation 1 – Michaelis Menten Parameter Fitting

$$a) y = \frac{V_{\text{max}} x}{K_m + x} \quad b) V_{\text{max}} = (\text{max } y) \quad c) K_m = x \text{ at } \frac{V_{\text{max}}}{2}$$

Where x represents substrate concentration and y represents velocity.

3.5 Biotransformations

3.5.1 pET22b(+) MATOUAmDH2 Biotransformations

All chemicals and enzyme recycling systems were provided by Sigma-Aldrich with the except of NADPH which was provided by Roche. Purified pET22b(+) MATOUAmDH2 was used, concentrations were confirmed via the use of a NanoPhotometer (Implen). Standards were made up by adding either 1-2 μL (for cyclohexanone or the amine product) to 400 μL ethyl acetate and NaOH where applicable. Methylamine-HCl stocks were made up in mQ H_2O with cyclohexanone stocks being made up in DMSO. All reactions were incubated at either 30°C or 50°C with shaking at >200 rpm. Reactions were initially screened at 0.1 $\text{mg}\cdot\text{mL}^{-1}$ MATOUAmDH2 with 0.2 mM NADPH but no reaction took place. Initial reactions for both ammonia and methylamine amine donors were carried out with a glucose-6-phosphate dehydrogenase (3 $\text{U}\cdot\text{mL}^{-1}$): glucose-6-phosphate (1.1 equiv.) recycling system (G-6P DH/G-6P). Subsequent reaction conditions where ammonia was the amine donor are as follows (3 mL total reaction volume): 10 mM cyclohexanone **21**, 12 mM glucose (1.2 equiv.), 3 $\text{U}\cdot\text{mL}^{-1}$ (9

U) glucose dehydrogenase (G-DH), 0.5 mM NADPH and 1 mg.mL⁻¹ MATOUAmDH2 made up with 2 M ammonium formate buffer pH 8.2. Subsequent preliminary reactions with methylamine as the amine donor were also carried out. Reaction conditions were as follows (1 mL total reaction volume): 10 mM cyclohexanone **21**, 12 mM (1.2 equiv.), 3 U.mL⁻¹ (9 U) G-DH, 0.5 mM NADPH, 1 mg.mL⁻¹ MATOUAmDH2 and 250 mM methylamine (25 equiv.) in 50 mM TRIS-HCl buffer pH 8.5. Aliquots of 200-100 µL were taken every 1-2 h typically between t=0-8 and t=16-26 h. With t=0 time points being taken directly after the addition of enzyme. Aliquots were worked up through the addition of NaOH and ethyl acetate; the organic layer was dried using MgSO₄ and then collected.

Samples were analysed using GC-FID using a gas chromatography trace equipped with a Flame Ionization Detector (FID). 1 µL samples were loaded onto an Agilent HP-5ms Ultra Inert column (30m x 250µm x 0.25µm) using a split ratio of 1:50 (for ammonia-based reactions) or 1:10 (for methylamine-based reactions). For ammonia-based reactions the initial oven temperature was 80°C where the ramp was set to 5°C min⁻¹ until 90°C and then 30°C min⁻¹ until a maximum temperature of 150°C (Figure 45). For methylamine-based reactions the ramp was set at 10°C min⁻¹ until a maximum temperature of 170°C (Figure 45). The needle was washed with ethyl acetate between each sample injection.

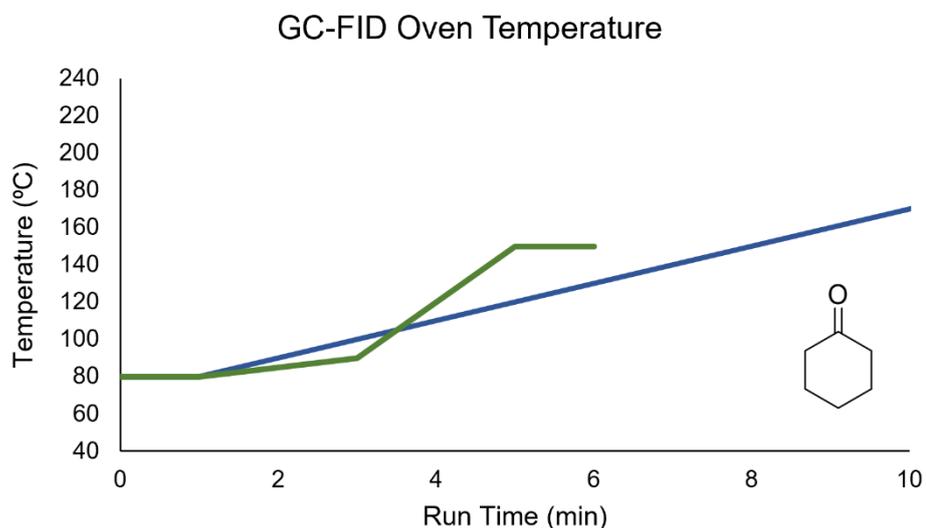


Figure 45: Programmed GC-FID oven temperatures for analysing cyclohexanone products and reactants. This is a graphical representation of the oven temperatures that were varied for the analysis of reactions where cyclohexanone is the substrate for both methylamine and ammonia runs using pET22b(+) constructs. Run times (x-axis) are displayed in minutes and temperature (y-axis) is displayed in degrees Celsius.

3.5.2 pETYSBLIC-3C MATOUAmDH2 and Mutant Biotransformations

For the alanine screens (L180A, M215A, T312A and wild type enzyme) purified protein was used with concentrations being confirmed via the use of a NanoPhotometer (Implen). Standards and chemicals were made the same as above, except where specified otherwise (*Methods 3.5.1*). All reactions were incubated at 25°C with shaking at 150-300 rpm. 200 μL aliquots were taken every 1-2 h for $t=0-8$ h and then a $t=24$ h time point also being taken. Aliquots were worked up via the addition of 20 μL 10 M NaOH and 600 μL ethyl acetate; the organic layer was dried using MgSO_4 and then collected. Samples were analysed using GC-FID using a gas chromatography trace equipped with a Flame Ionization Detector (FID). 1 μL samples were loaded onto an Agilent HP-5ms Ultra Inert column (30m x 250 μm x 0.25 μm) using a split ratio of 1:50. The needle was washed with ethyl acetate between each sample injection.

Reaction conditions for purified protein where ammonia was the amine donor are as follows (3 mL total reaction volume): 10 mM cyclohexanone **21**, 12 mM glucose (1.2 equiv.), 3 $\text{U}\cdot\text{mL}^{-1}$ (9 U) glucose dehydrogenase (G-DH), 0.5 mM NADPH and 1 $\text{mg}\cdot\text{mL}^{-1}$ MATOUAmDH2 made up with 2 M ammonium formate buffer pH 8.0. Reaction conditions for purified protein where methylamine was the amine donor are as follows (3 mL total reaction volume): 10 mM cyclohexanone **21**, 12 mM glucose (1.2 equiv.), 3 $\text{U}\cdot\text{mL}^{-1}$ (9 U) glucose dehydrogenase (G-DH), 0.5 mM NADPH, 250 mM methylamine (25 equiv.) and 1 $\text{mg}\cdot\text{mL}^{-1}$ MATOUAmDH2 made up with 50 mM Tris-HCl buffer pH 8.5. Three repeats were collected for each reaction of interest.

For reactions concerning a reduced library of point mutations at positions L180 and M215, lysates were produced, and concentrations approximated from cell pellet weight and SDS-PAGE analysis (*Methods 3.1.6, 3.2.1 and 3.2.3*). Reaction conditions for purified protein where ammonia was the amine donor are as follows (2-3 mL total reaction volume): 10 mM ligand, 12 mM glucose (1.2 equiv.), 3 $\text{U}\cdot\text{mL}^{-1}$ (9 U) glucose dehydrogenase (G-DH), 0.5 mM NADPH and 20 $\text{mg}\cdot\text{mL}^{-1}$ cell lysate made up with 2 M ammonium formate buffer pH 8.0. Reaction conditions for purified protein where methylamine was the amine donor are as follows (2-3 mL total reaction volume): 10 mM ligand, 12 mM glucose (1.2 equiv.), 3 $\text{U}\cdot\text{mL}^{-1}$ (9 U) glucose dehydrogenase (G-DH), 0.5 mM NADPH, 250 mM methylamine (25 equiv.) and 20 $\text{mg}\cdot\text{mL}^{-1}$ cell lysate made up with 50 mM Tris-HCl buffer pH 8.5. Some reactions were scaled down to 500 μL for initial testing purposes. Two repeats were collected for each reaction of interest.

3.5.2.1 Cyclic Substrates Based Reactions

For both ammonia and methylamine reactions cyclohexanone **21** was dissolved in DMSO. Norcamphor **29a** and 1,2-cyclohexanedione **31** were both dissolved in ethyl acetate. The initial oven temperature was 80°C where the ramp was set to 15°C min⁻¹ until a maximum temperature of 200°C, as shown in Figure 46 below.

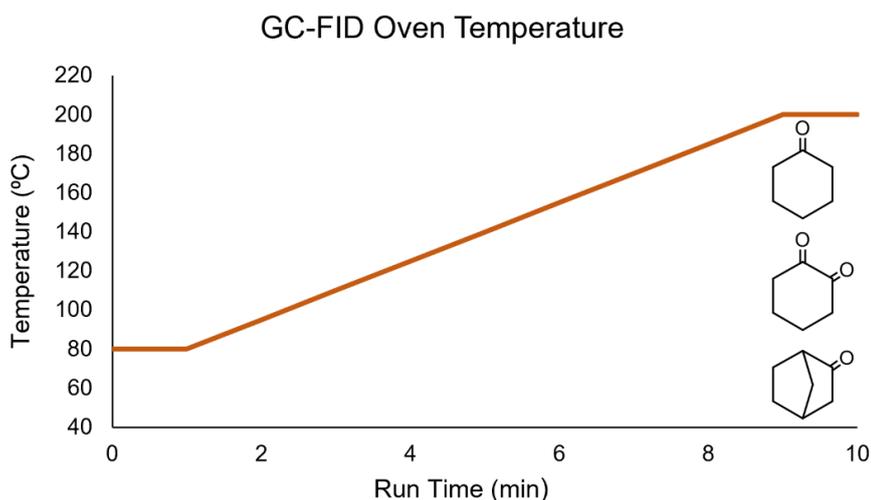


Figure 46: Programmed GC-FID oven temperatures for analysing cyclic products and reactants. This is a graphical representation of the oven temperatures that were varied for the analysis of reactions where cyclohexanone, 1,2-cyclohexanedione and norcamphor were the substrates. Run times (x-axis) are displayed in minutes and temperature (y-axis) is displayed in degrees Celsius.

3.5.2.2 Aliphatic Substrates and Ammonia Reactions

For ammonia reactions 2-heptanone **32** and 2-pentanone **20** were dissolved in ethyl acetate. The initial oven temperature was 60°C where the ramp was set to 5°C /min until a maximum temperature of 110°C, as shown in Figure 47 below.

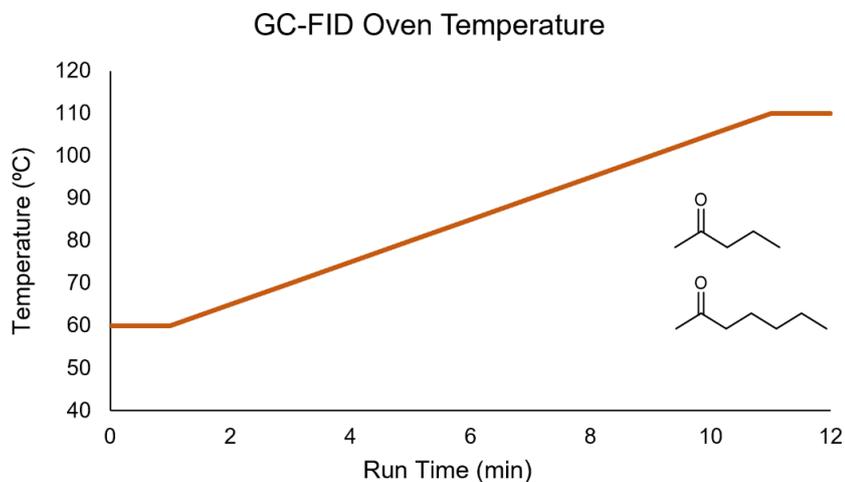


Figure 47: Programmed GC-FID oven temperatures for analysing aliphatic products and reactants. This is a graphical representation of the oven temperatures that were varied for the analysis of reactions where 2-pentanone and 2-heptanone were the substrates. Run times (x-axis) are displayed in minutes and temperature (y-axis) is displayed in degrees Celsius.

3.5.2.3 β -Tetralone and Ammonia Reactions

For ammonia reactions β -tetralone (2-tetralone) **15** was dissolved in DMSO. For tetralone an isothermal oven temperature was employed where initial temperature was set to 140°C and held for ten minutes until a ramp of 15°C /min until 180°C, this is displayed in Figure 48 below.

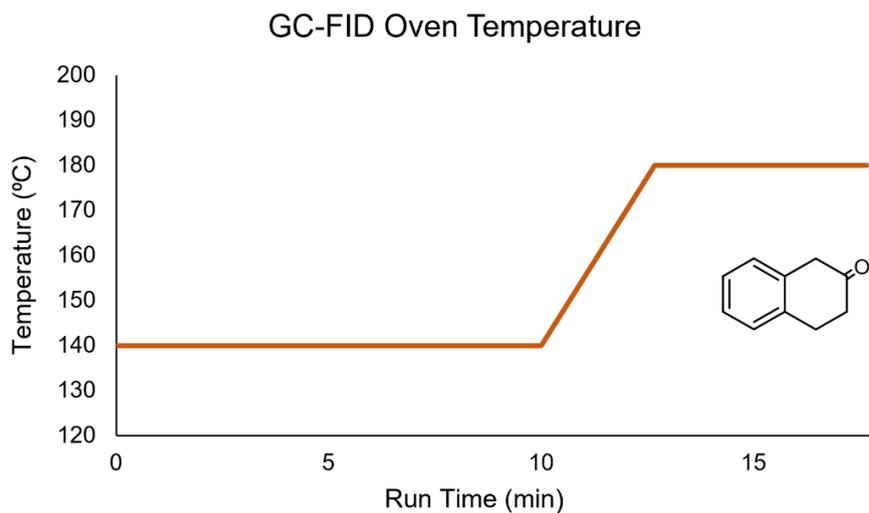


Figure 48: Programmed GC-FID oven temperatures for analysing 2-tetralone products and reactants. This is a graphical representation of the oven temperatures that were varied for the analysis of reactions where 2-tetralone was the substrate. Run times (x-axis) are displayed in minutes and temperature (y-axis) is displayed in degrees Celsius.

3.5.3 Peak Identification and Data Processing

Retention times for all standards are reported in *Appendices*: Figure 122-130 and Table 19. Peaks were processed by the data analysis interface of Aligent 7890B to give the percentage areas in question. Graphs were plotted using OriginPro® (OriginLab, v.2021). Standard errors were calculated in OriginPro® from an average of two or three repeats for each reaction and y-error bars were plotted.

4. Results and Discussion

4.1 Characterisation of pET22b(+) MATOUAmDH2

MATOUAmDH2, a native amine dehydrogenase (nat-AmDH), was discovered by sequence driven searches of the Marine Atlas of Tara Oceans Unigenes (MATOUv1) metagenomic database.¹⁰³ MATOUAmDH2 was provided by Genoscope in a pET22b(+) vector which was optimised for *E. coli* expression. MATOUAmDH2 shows sequence similarity to both *CfusAmDH* (34%) and *MsmAmDH* (34%) (Figure 49). The protein BLAST alignments are shown in *Appendices: Figure 131 and 132*.

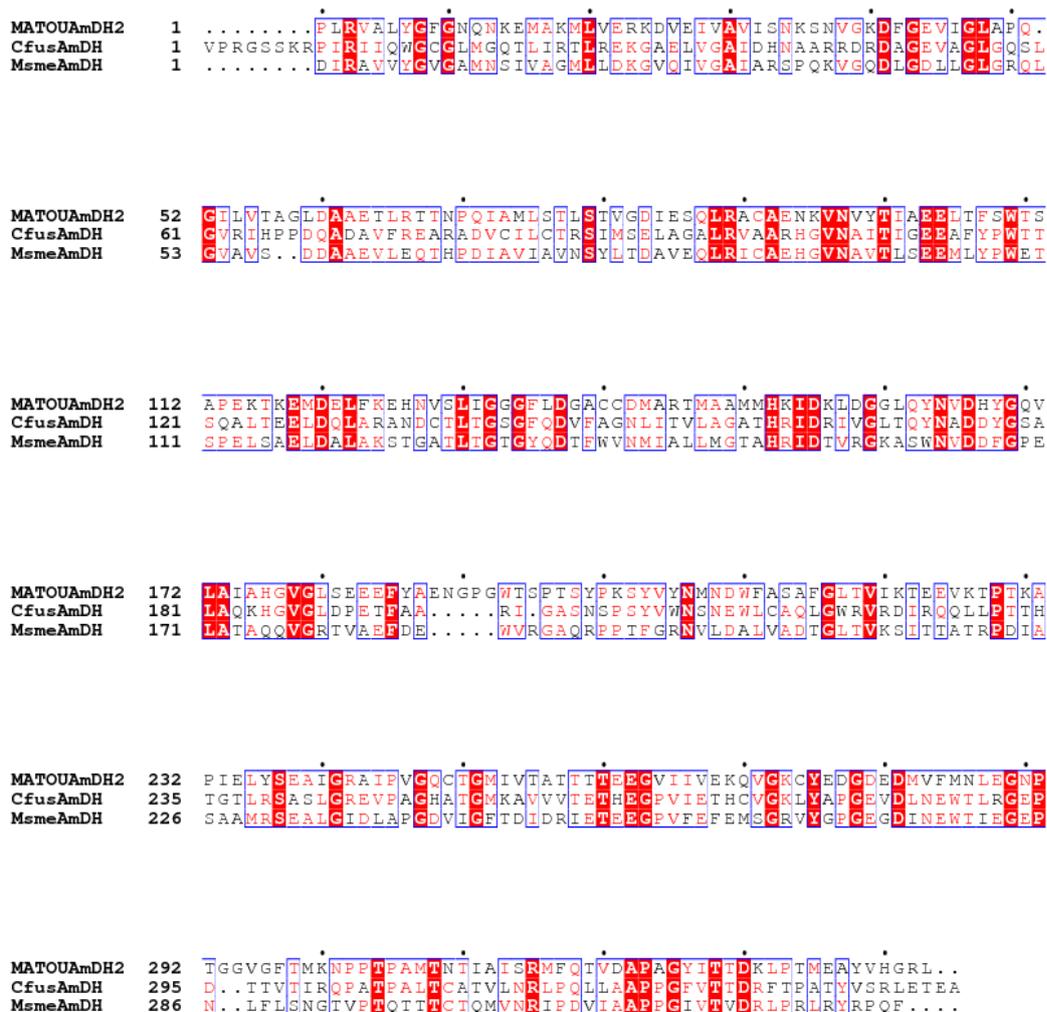


Figure 49: Amino acid sequence alignment of MATOUAmDH2, *CfusAmDH* and *MsmAmDH*. Sequence alignment generated using MultAlin and ESPrnt using native sequences of MATOUAmDH2, *CfusAmDH* and *MsmAmDH*. Red highlights either partial consensus (red text only) or high consensus (block red and white text).

MATOUAmDH2 was successfully transformed into BL21(DE3) cells and the gene was expressed, and protein purified using Nickel Affinity (NiNTA) and Size Exclusion Chromatography (SEC) (*Methods* 3.2). Typical yields after purification were between 50-100 mg of protein per 2 L of culture. Typically, protein eluted at a low-UV imidazole concentration of around 160 mM (Figure 50). During SEC protein eluted at around 55 mL (Figure 51). Usually, the protein was relatively pure with little to no co-eluted protein, shown by the presence of singular A_{280} peaks on both chromatograms (Figure 50 and 51) and SDS-PAGE analysis (Figure 52). SDS-PAGE analysis also revealed a relatively pure over expressed band at approximately the correct molecular weight (38.2 kDa), shown in Figure 52.

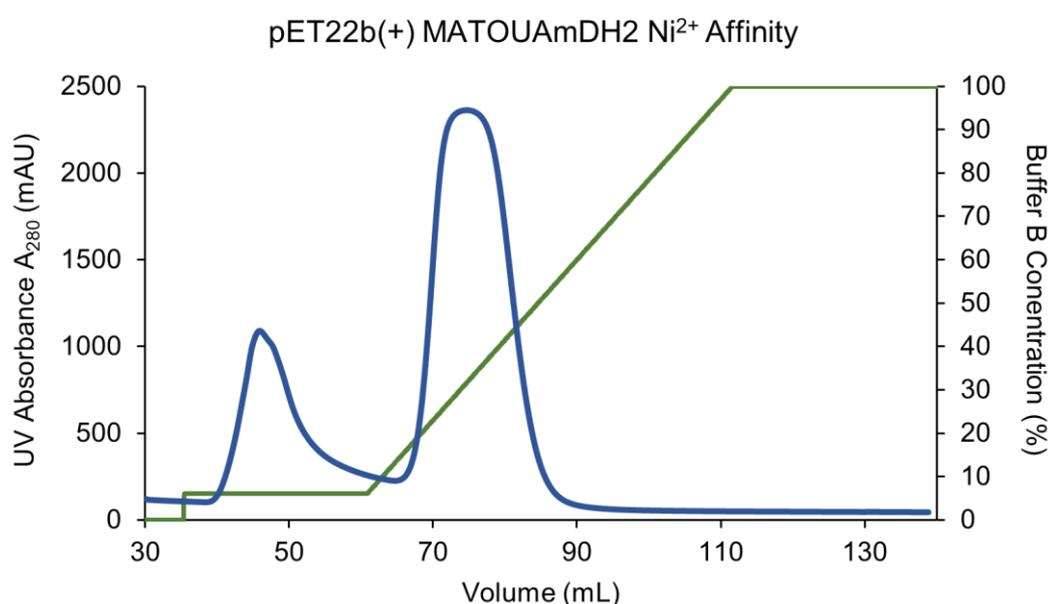


Figure 50: Nickel Affinity Chromatogram for MATOUAmDH2 expressed in the pET22b(+) vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line.

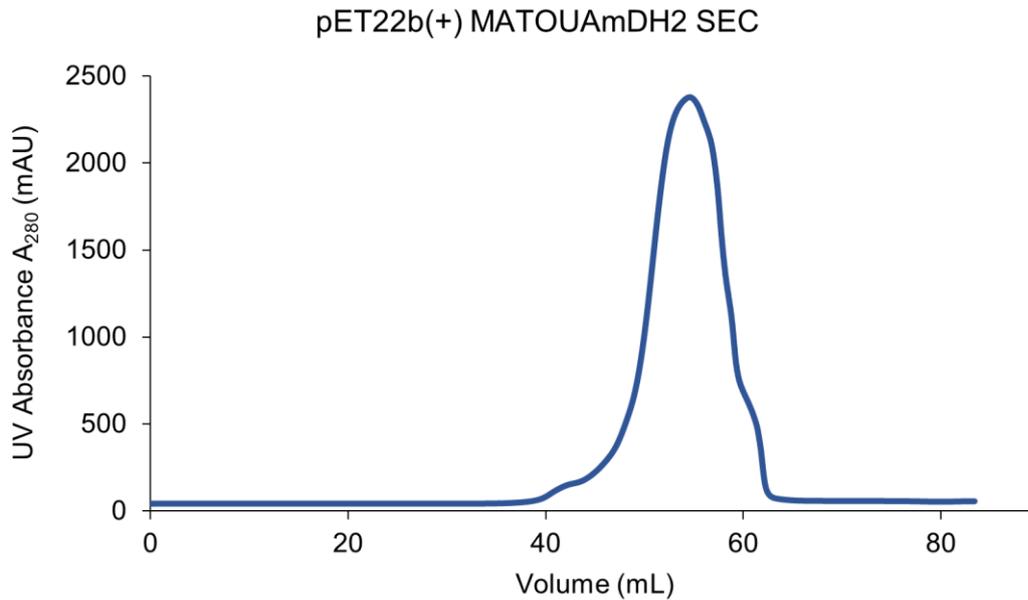


Figure 51: Size Exclusion Chromatogram for MATOUAmDH2 expressed in the pET22b(+) vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue).

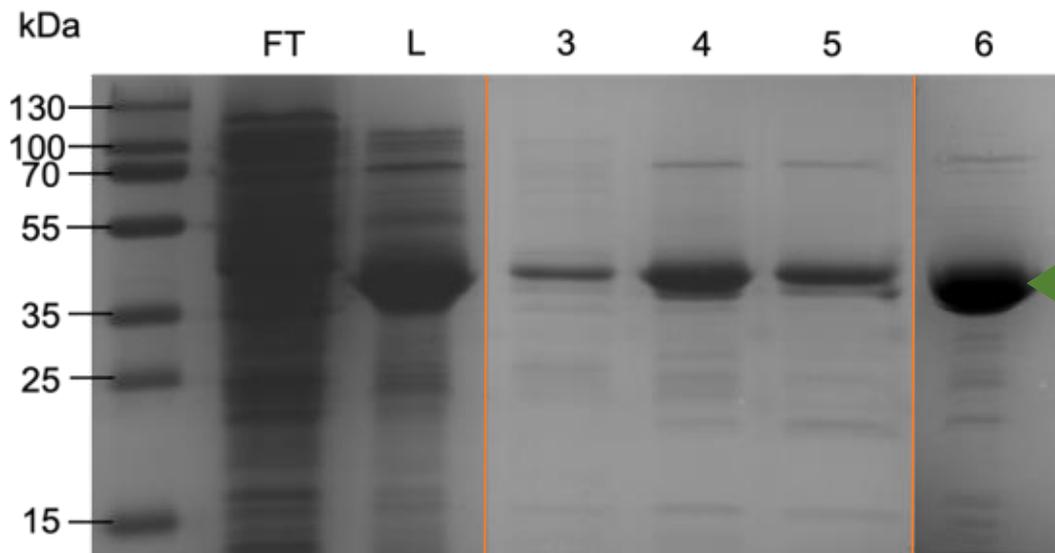


Figure 52: SDS-PAGE analysis of protein fractions from pET22b(+) MATOUAmDH2 protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane FT*: 10 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane L*: 10x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lanes 3, 4 and 5*: 10x diluted pre-determined samples from the ÄKTA fraction collector after Nickel Affinity Chromatography. *Lane 6*: 20x diluted pooled and concentrated pure protein after Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. Orange lines indicate spliced sections of the gel. The green triangle indicates the protein bands of interest.

MATOUAmDH2 protein that was confirmed to be relatively pure after SEC was then usually pooled and concentrated to around 8-20 mg.mL⁻¹. Protein aliquots were flash frozen and stored at -80°C until needed. Kinetic studies were performed on MATOUAmDH2 using a model substrate, cyclohexanone **21**, which has been shown to be an ideal binding partner for not only MATOUAmDH2 but also other previously described nat-AmDH.⁹¹ UV-Vis spectrometry was performed on MATOUAmDH2 at cyclohexanone concentrations ranging from 0-20 mM using the oxidation of NADPH, and thus reduction in A₃₄₀, as a way to calculate velocity (*Methods 3.4.2*). Velocity was plotted against cyclohexanone concentrations and error bars fitted for two repeats (Figure 53). Kinetic parameters were calculated using the Michaelis Menten non-linear curve fitting interface within OriginPro® (OriginLab v.2021) (Table 9).

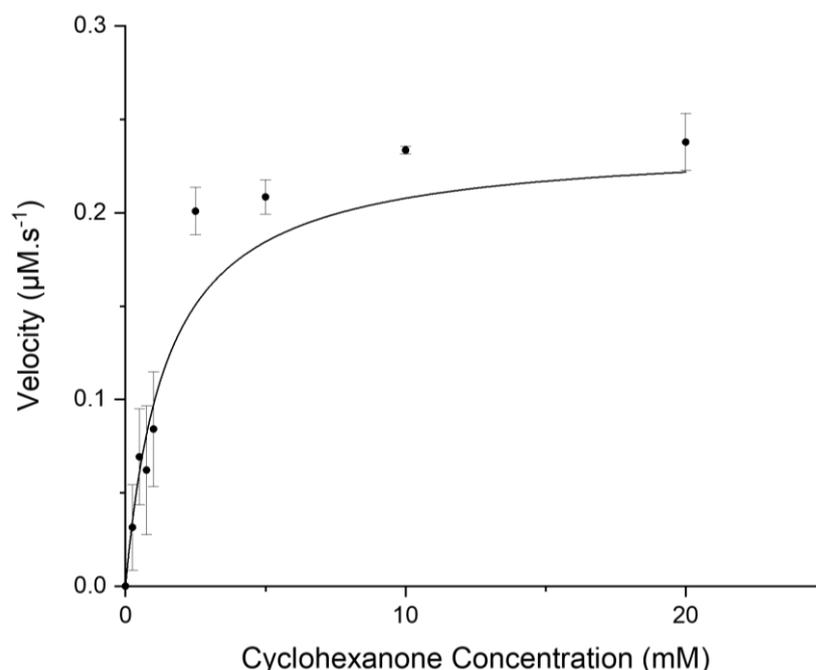


Figure 53: Kinetic plot for pET22b(+) MATOUAmDH2 with cyclohexanone at 50°C. X-axis represents cyclohexanone concentrations in mM. Y-axis represents the velocity in µM.s⁻¹. Michaelis Menten curve fit is represented in black and Y-error bars and caps are displayed at each concentration/velocity point.

Table 9 – Kinetic parameters for pET22b(+) MATOUAmDH2 with cyclohexanone

Variant	V_{max} (µM.s ⁻¹)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ .mM ⁻¹)
pET22b(+) MATOUAmDH2	0.238 ± 0.0195	1.45 ± 0.395	0.364	0.251

Preliminary biotransformations were conducted on pET22b(+) MATOUAmDH2 using cyclohexanone and both ammonia and methylamine as amine donors. Reactions were carried out at 30°C and 50°C for ammonia reactions and 30°C for methylamine reactions. Retention times, when using the preliminary ramp conditions (*Methods 3.5.1*), for cyclohexanone **21** standards and substrate peaks, were approximately 3.10 min - for cyclohexylamine **21b** the retention time was around 2.67 min (Figure 54). Retention times for methylamine-based reactions, when using a different temperature ramp (*Methods 3.5.1*) were as follows; cyclohexanone **21**: 2.39 min and *N*-methylcyclohexylamine **21c**: 2.54 min (Figure 55). Both temperatures yielded good conversions with ammonia, converting around 100.0% and 87.9% of cyclohexanone to cyclohexylamine, after 25 h, for 30°C and 50°C respectively (Figure 56). For preliminary reactions with methylamine adequate conversion, in line with expected values, was seen after 24 h, yielding a conversion of 17.3% (Figure 57). The reaction also continued to progress after 24 h, where a conversion of 26.6% was seen after 42 h (Figure 57). Peak standards and retention times are reported in *Appendices: Figure 120 and 121*.

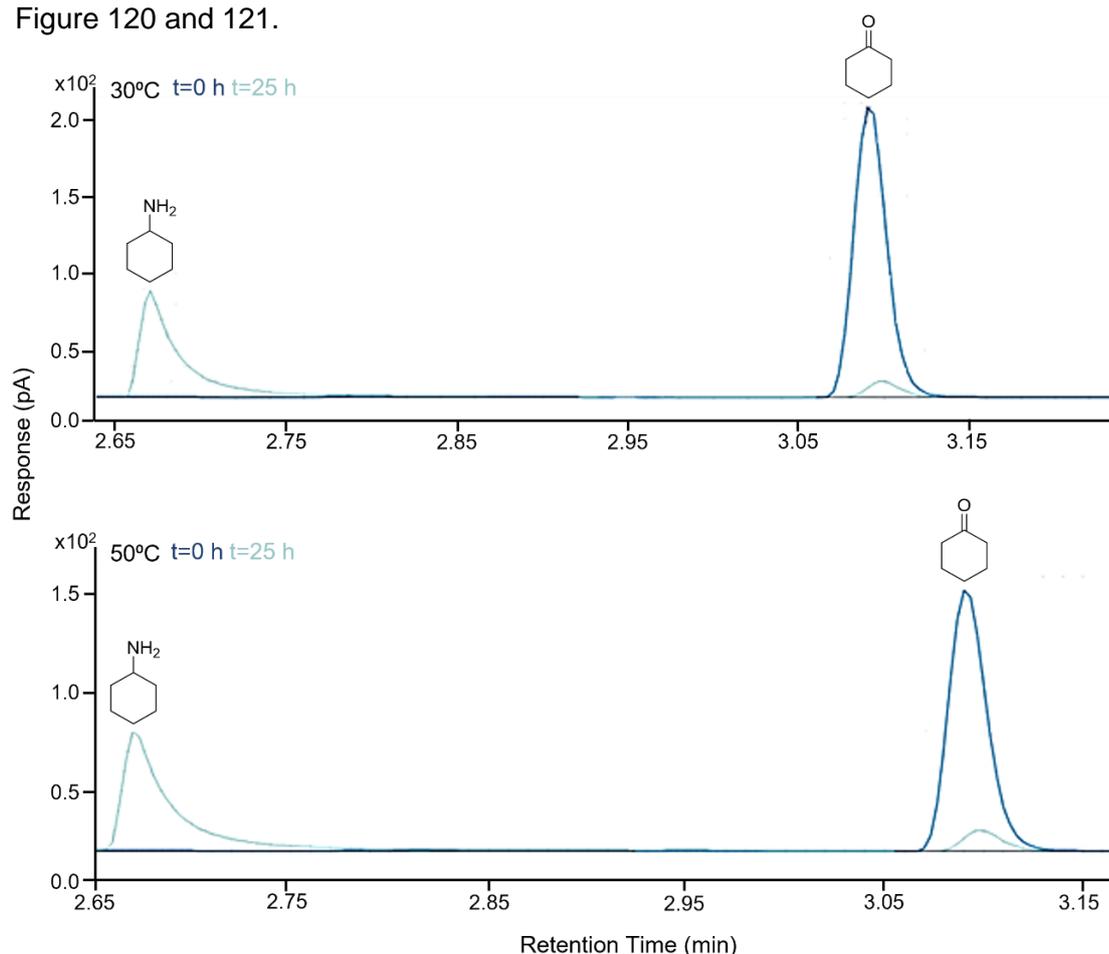
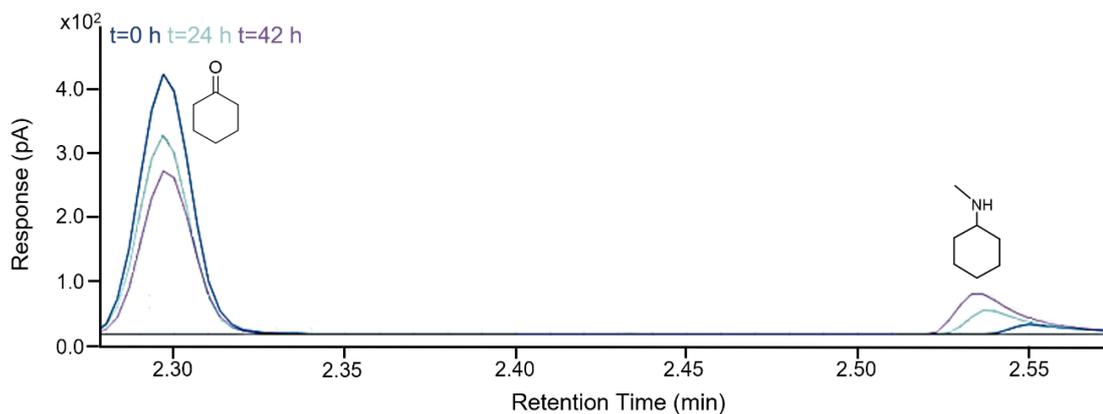
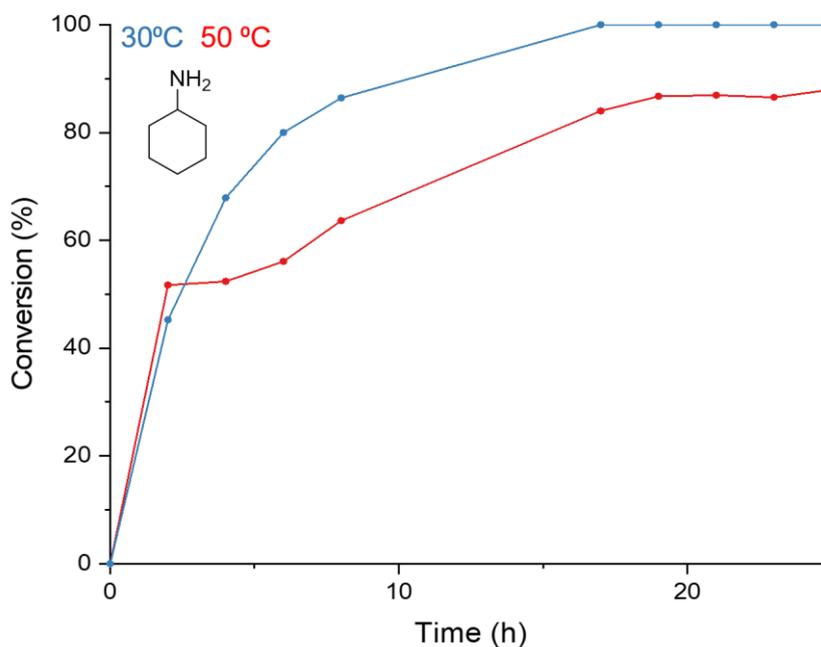


Figure 54: GC-FID traces for pET22b(+) MATOUAmDH2 reactions with ammonia and cyclohexanone. The x-axis represents retention times in min. The y-axis represents the response current in pA. The top panel displays the reaction carried out at 30°C. The bottom panel represents the reaction carried out at 50°C. Time points at 0 h are represented in dark blue at t=25 h time points represented in light blue. Retention times linked to chemical standards are labelled accordingly.



**Figure 55: GC-FID traces for pET22b(+)
MATOUAmDH2 reactions with methylamine and cyclohexanone.** The x-axis represents retention times in min. The y-axis represents the response current in pA. All reactions were carried out at 30°C. Time points at 0 h are represented in dark blue at t=24 h time points represented in light blue and t=42 h points are in purple. Retention times linked to chemical standards are labelled accordingly.



**Figure 56: Reaction progressions of biotransformations using pET22b(+)
MATOUAmDH2 with ammonia and cyclohexanone.** The x-axis represents time (in h) where each time point of the reaction was collected. The y-axis represents the percentage conversion to product (cyclohexylamine). Conversion percentages are plotted for appropriate time points for both 30°C (blue) and 50°C (red).

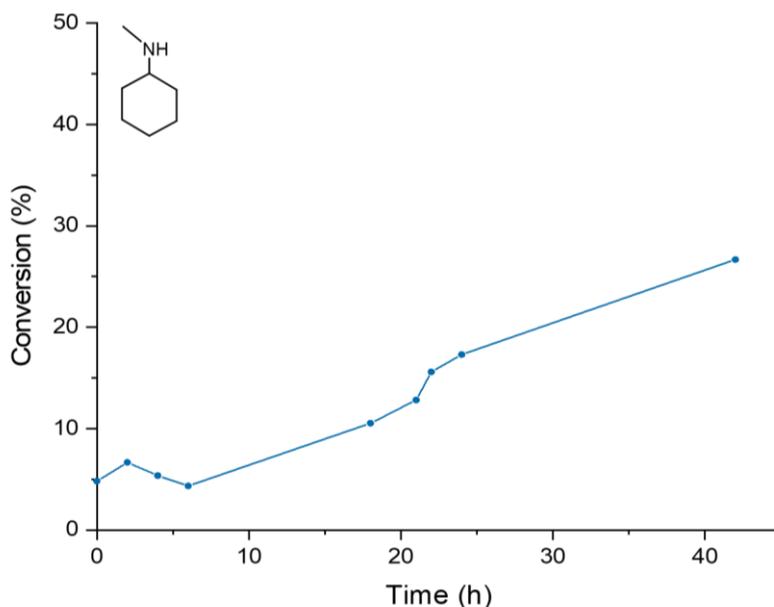


Figure 57: Reaction progression of a biotransformation using pET22b(+) MATOUAmDH2 with methylamine and cyclohexanone. The x-axis represents time (in h) where each time point of the reaction was collected. The y-axis represents the percentage conversion to product (N-methylcyclohexylamine). All data was collected at 30°C.

Purified pET22b(+) MATOUAmDH2 also underwent initial crystallisation trials using commercially available pre-dispensed screens. Namely the Clear Strategy™ Screen (CSS) and Index screens. The CSS screens yielded a range of needles using buffers close to, and above, the proteins isoelectric point, the theoretical pI was calculated at 5.05. The buffers included in most screens were either HEPES or Tris based buffers at pH 5.5-8.5. However, the best crystal forms resulting from initial screening trials were from Index screens. After ~3 d of growth, groups of clustered plates started to form. The best hits were from drops containing 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Bis-Tris pH 5.5 and 25% (w/v) PEG 3350 with drop concentrations (150 nL protein complex:150 nL mother liquor) at 15 mg.mL⁻¹ MATOUAmDH2 and 5 mM NADP⁺ (Figure 58).



Figure 58: Crystal forms from initial screening of pET22b(+) MATOUAmDH2 in complex with NADP⁺. Microscopy images of plate clusters from a pre-dispensed Index (MRC-96 well) screen.

From initial screening trials optimisation screens were designed in 48 well formats optimising $(\text{NH}_4)_2\text{SO}_4$ concentrations, buffer pH and PEG 3350 percentages. In addition to these optimisation screens some additive screens were also set up, also in 48 well formats using 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Bis-Tris pH 5.5 and 25% (w/v) PEG 3350 supplemented with 0-25% of either ethylene glycol, glycerol, butan-1,4-diol or methylpentanediol (MPD). The best crystals were obtained from drops containing 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Bis-Tris pH 5.5, 25% (w/v) PEG 3350 and 4-10% MPD (Figure 59). These drops contained final drop concentrations of $10 \text{ mg}\cdot\text{mL}^{-1}$ MATOUAmDH2 in complex with 5 mM NADP^+ . These singular plates were fished after ~5 d of growth, and flash cooled with liquid nitrogen and no additional cryoprotectant. Any crystals which showed a resolution greater than 3 Å and ordered diffraction patterns, during in-house diffraction testing (Figure 59), were sent to the Diamond synchrotron for data collection.

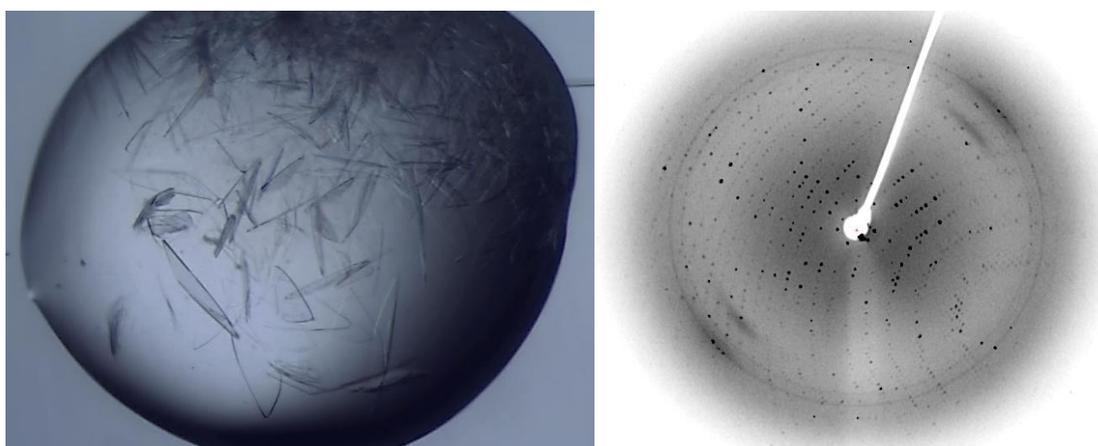


Figure 59: Crystals and diffraction patterns from optimisations of pET22b(+) MATOUAmDH2 in complex with NADP^+ . The left panel represents microscopy images of singular plates and clusters from an optimisation screen (MAXI-48 well). The right panel represents the diffraction pattern collected in house using a Rigaku Micromax-007HF fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector visualised using Advx (v.1.9.14).

Initially a dataset from Diamond beamline I03 was used to build and refine a structure to a resolution of 2.5 Å. There were four molecules in the asymmetric unit and the space group was $P2_12_12_1$. The unit cell dimensions for this dataset are as follows; $a = 91.41$; $b = 114.39$; $c = 154.38$ $\alpha = \beta = \gamma = 90^\circ$. However, a better dataset was obtained and resolved to a resolution of 2.3 Å with the same approximate unit cell, same space group also with four molecules in the asymmetric unit. The data and refinement statistics for the 2.3 Å structure are reported in Table 10. This data set was deposited in the Protein Data Bank (PDB) under accession code 7ZBO.

Table 10 – Data collection and refinement statistics for the 2.3 Å pET22b(+) MATOUAmDH2 in complex with NADP⁺.

2.3 Å pET22b(+) MATOUAmDH2:NADP⁺

Beamline	Diamond I04
Wavelength (Å)	0.97950
Resolution (Å)	58.62-2.32 (2.37-2.32)
Space Group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
Unit cell (Å)	a = 91.10 b = 112.19; c = 153.15 α = β = γ = 90°
No. of molecules in the asymmetric unit	4
Unique reflections	68672 (4581)
Completeness (%)	100.0 (100.0)
R _{merge} (%)	0.09 (0.49)
R _{p.i.m.}	0.09 (0.49)
Multiplicity	1.9 (1.9)
<I/σ(I)>	3.6 (1.1)
Overall B from Wilson plot (Å ²)	26
CC _{1/2}	0.99 (0.67)
R _{cryst} / R _{free} (%)	24.0/28.4
r.m.s.d 1-2 bonds (Å)	0.01
r.m.s.d 1-3 angles (°)	1.68
Avge main chain B (Å ²)	39
Avge side chain B (Å ²)	40
Avge water B (Å ²)	26
Avge NADP ⁺ B (Å ²)	38

Brackets refer to data in the highest resolution shell.

The final structure was refined to R values of 24.0% and 28.4% for R_{crystal} and R_{free} values respectively. Four chains were built with chains A:D and B:C forming dimers associating at a series of beta pleated sheet structures (Figure 60). Looking at all the chains from the pET22b(+) MATOUAmDH2:NADP⁺ structure showed that chains B and D seemed to be more ‘open’ conformers whereas chains A and C appeared to be more of a ‘closed’ form as seen in Figure 60. Residues 8-158 and 309-351 make up the large N-terminal Rossman fold domain with residues 159-308 making up a tertiary beta pleated sheet comprised of 8 strands (Figure 60 and 61). Each monomer binds one molecule of NADP⁺ at the interface between the two domains, for which there was clear omit F_o-F_c electron density in each binding pocket (Figure 62). Each closed monomer displayed a binding pocket width of approximately 13 Å at the largest opening which is highlighted in Figure 61. DALI analysis¹⁰⁴ revealed that the two most related structural homologues were *CfusAmDH* (PDB: 6IAU) showing 33% sequence identity and an RMSD value of 2.4 Å over 340 C α atoms, and *MsmAmDH* (PDB: 6IAQ) with a sequence identity of 33% and also a RMSD value of 2.4 Å for 340 C α atoms; both of which were previously crystallised and reported.⁹¹ Sequence alignments are reported in *Appendices*: Figure 131 and 132.

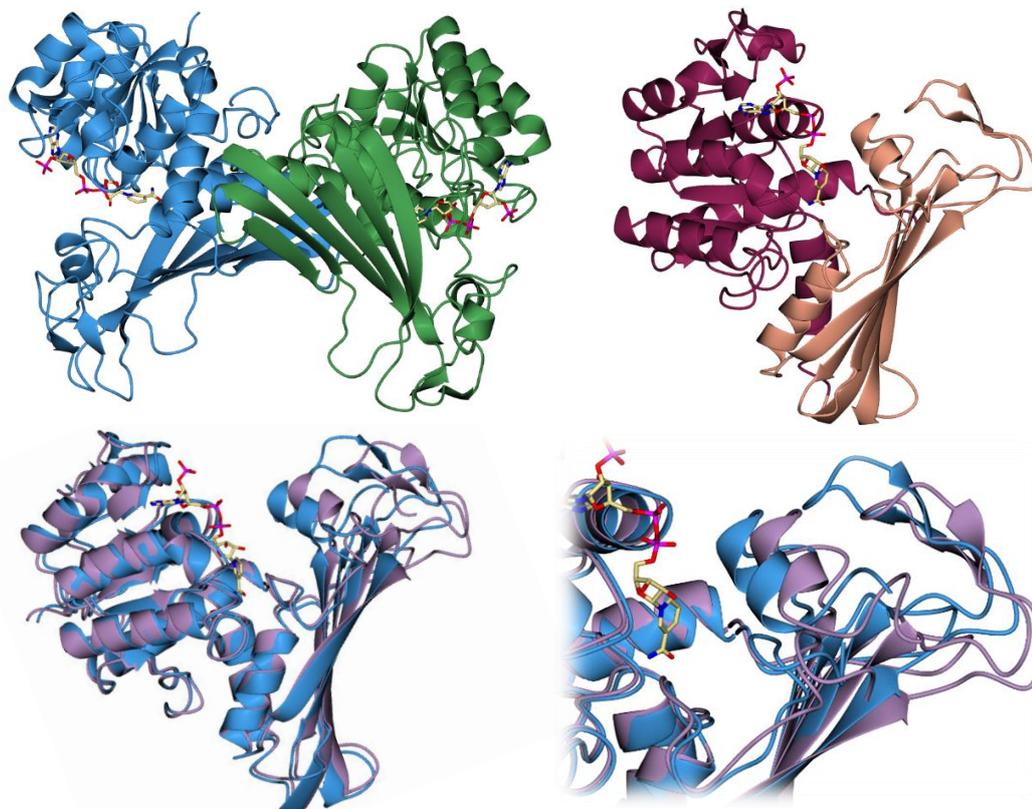


Figure 60: Structural properties of pET22b(+) MATOUAmDH2 in complex with NADP⁺. NADP⁺ is displayed in cylindrical form in yellow in all cases. All images are visualised in CCP4mg (v.2.10.11). The top left-hand panel shows an associated dimer in its biological assembly made up of chains B (green) and C (blue). The top right-hand panel displays the different tertiary structures within one monomer; the N-terminal Rossman domain (wine red) and C terminal beta sheet (orange). The bottom panels display both open and closed forms seen in the apo structures from chains C (blue) and D (lilac).

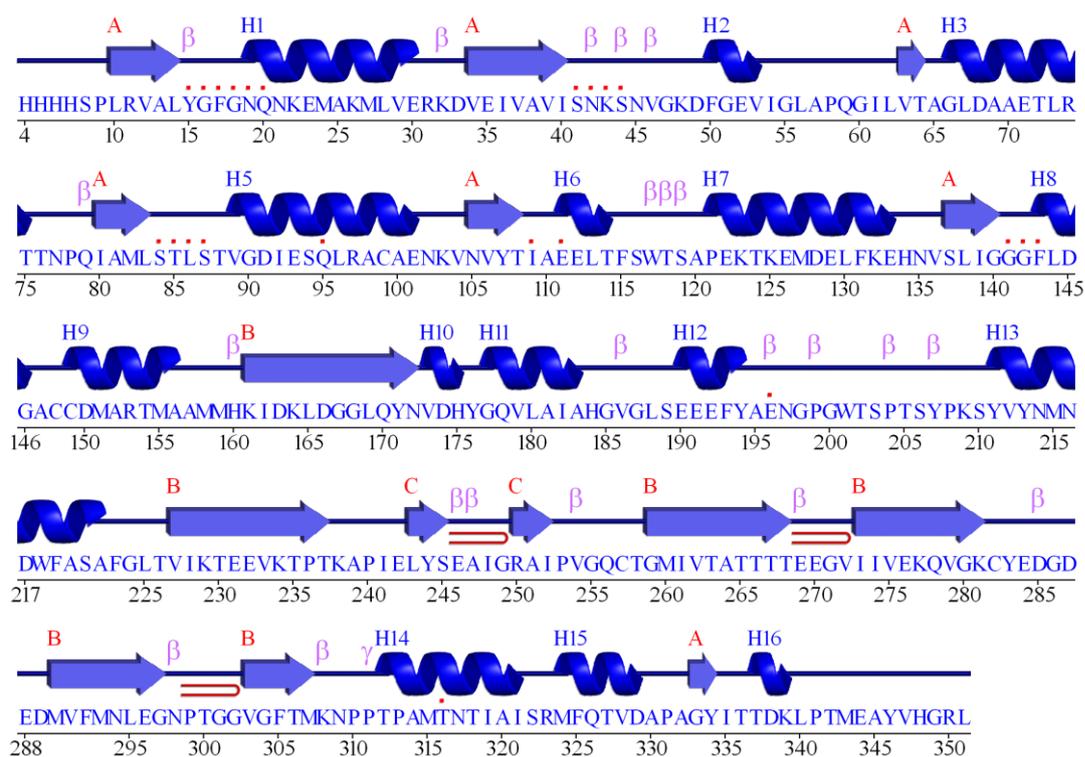


Figure 61: Secondary structure visualisation of pET22b(+) MATOUAmDH2. Visualisation was carried out using PDBsum. Secondary structures are separated into alpha helices or beta strands, beta hairpins are indicated in red turns. The sequence and sequence numbers are displayed as reference points.

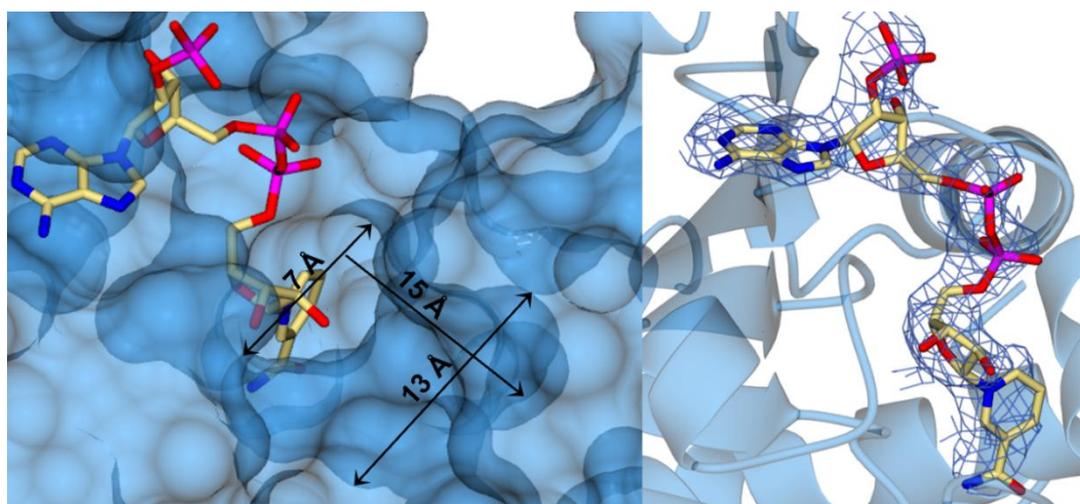


Figure 62: Visualisation of the binding pocket of pET22b(+) MATOUAmDH2 and associated NADP⁺ omit map. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel displays the binding pocket of chain C in complex with NADP⁺ (yellow). Arrows display the width of the narrowest and widest opening of the binding pocket as well as the depth. All measurements are in Ångstroms, approximated from the inbuilt scale. The right-hand panel displays the F_0-F_c electron density associated with NADP⁺ (yellow) from chain C, which is represented in blue cylindrical wire and set at a map contour level of 3σ .

Attempts were made to produce a ligand bound structure of MATOUAmDH2 expressed in the pET22b(+) vector. A range of ligands were screened for co-crystallisation including cyclohexylamine **21b**, 2-aminocyclohexanone **31a**, hexylamine **33** and pentylamine **30**. Unfortunately, initial screening trials and optimisations both yielded poor crystal forms on co-crystallisation ligand binding. Often crystal forms tended to represent needle or fan like structures most of which were unrecoverable. Figure 63 displays an attempt at obtaining a co-crystallised ligand bound structure using 10 mM pentylamine in complex with drop concentrations of 10 mg.mL⁻¹ MATOUAmDH2 and 5 mM NADP⁺. The fan like crystal forms were imaged from drops containing 0.2 M (NH₄)₂SO₄, 0.1M Bis-Tris pH 5.5, 25% (w/v) PEG 3350 and 6% MPD. Any that were fishable displayed poor electron diffraction and no real ordered phasing when tested in-house (Figure 63).

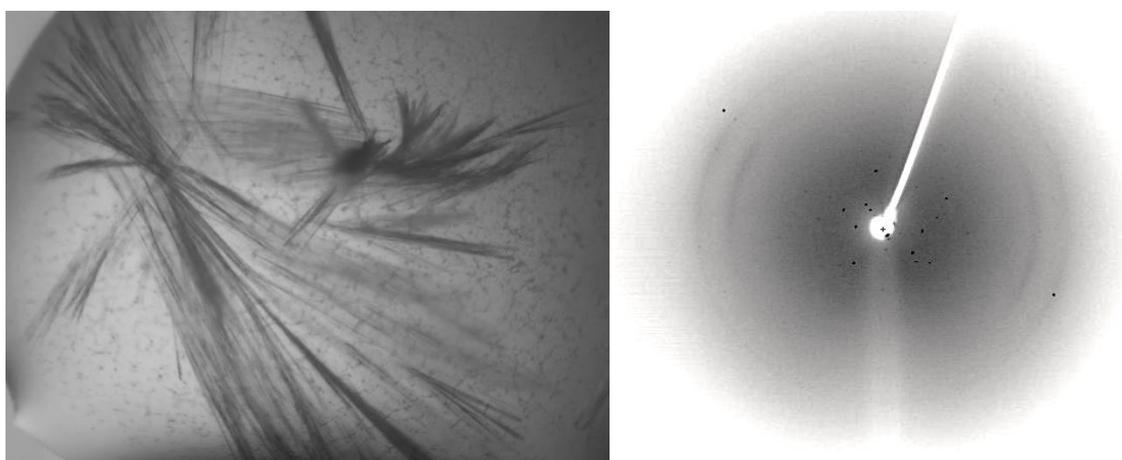


Figure 63: Crystals and diffraction patterns from trials of pET22b(+) MATOUAmDH2 in complex with NADP⁺ and pentylamine. The left panel represents microscopy images of fan like structures from an optimisation screen (MAXI-48 well). The right panel represents the diffraction pattern collected in house using a Rigaku Micromax-007HF fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector visualised using Advx (v.1.9.14).

4.1.1 Discussion

MATOUAmDH2 expressed very well using the pET22b(+) vector in which it was provided. The purification often gave high protein yields with pure products. This may be attributed to the very good stability seen with this wild-type protein, which supports previous data suggesting MATOUAmDH2's relative activity is highest at around 50°C and still retains around 75% of its activity when stored at 25°C for a week.¹⁰³ Kinetic data obtained gave a k_{cat} value for MATOUAmDH2 of 0.36 s⁻¹ which is slightly lower than previous data obtained (0.52 s⁻¹) but is well within the realms of activity of this enzyme.¹⁰³ Furthermore, this is a much-improved k_{cat} value for cyclohexanone-based

reactions when compared to previously reported nat-AmDHs.¹⁰³ The K_m value in this case is slightly higher than expected, whilst still outperforming previous nat-AmDHs⁹¹; this is possibly due to human error. At lower cyclohexanone concentrations the reaction is progressing the fastest and the data obtained can differ the most depending on how quickly the cuvette is inserted and scanning has started as there is no autosampler. This may also be linked to the relatively high specific activity at 50°C, which is particularly noticeable by relatively high standard errors for cyclohexanone concentrations between 0 and 1 mM and thus poor non-linear curve fitting overall. As a result of this, and for other reasons detailed in (*Results and Discussion 4.3*), the remaining kinetics measurements were carried out at 25°C. This would avoid any over- or underestimation of kinetics parameters which were solely dictated by differences in repeats due to possible human error.

Preliminary biotransformations using pET22b(+) MATOUAmDH2 worked extremely well with ammonia and cyclohexanone **21** as a model system. At 30°C the enzyme gave 100.0% conversion after only 17 h. This is a particularly nice result as it appears to be the highest conversion for time reported for any substrate with this enzyme, with most aliphatic based biotransformations only reaching around <70.0% conversion after 24 h.¹⁰³ At 50°C good overall conversions were still observed, with 87.9% conversion after 25 h, but the enzyme seems to slow down after around 2 h and not reach its full potential conversion. This may be because initially at 50°C MATOUAmDH2's specific activity is at its highest, but it denatures over the course of the biotransformation reaction and the accumulation of aggregation slows down the reaction progression. At 50°C the total reaction conversion is most likely being overestimated, and just like with the UV-vis spectrometry experiments it was decided that any further biotransformations were to be carried out at 30-25°C. Preliminary biotransformations were also carried out using methylamine as the amine donor. The reaction progression was typical, having a much lower overall conversion when compared to ammonia-based reactions, at around 17.3% conversion to *N*-methylcyclohexylamine **21c** after 24 h. However, the reaction keeps progressing after this time and 26.6% conversion was observed after 42 h, which still doesn't plateau meaning that a bit more conversion after this time might be observed. Additionally, the data obtained for methylamine-based reactions showed that the experimental procedure need to be slightly altered. The GC oven temperature ramp was adjusted in later experiments (*Results and Discussion 4.3*) so that peaks were easier to see and analyse at lower concentrations (at the beginning of the reactions: t=2-6 h) to produce more reliable and accurate data.

The structure of MATOUAmDH2 in complex with its cofactor NADP⁺ was relatively easy to obtain only requiring a few optimisation screens. The presence of MPD seemed to be crucial in producing single plates optimised from plate clusters formed in previous drops. In comparison, obtaining a ligand bound structure using co-crystallisation proved much more difficult. When screening a range of ligands and drop conditions most of the crystals formed resembled fan or needle like structures making them nearly impossible to fish. Any that were fished showed poor diffraction and no ordered phasing as mentioned above. For this reason, it was decided to sub-clone MATOUAmDH2 into an in-house designed vector; the YSBLIC-3C vector. This vector has a cleavable N-terminal His tag which is connected via an extended linker of amino acids. Due to rigidity of the non-cleavable His tag provided on the pET22b(+) vector, which was seen in the electron density of the apo datasets, it was believed that this extended linker region would confer a degree of flexibility, which may improve crystal packing when binding ligand and inducing dynamic movement i.e., closing around the binding pocket. In addition, if this also did not confer better crystal packing there was the option to cleave the His tag and try to crystallise this protein, the outcome of which is mentioned in *Results and Discussion 4.2. Results and Discussion 4.2* explores in detail the effect of sub-cloning into the YSBLIC3C vector in an attempt to generate a ligand bound structure.

Most amino acid backbones could be fitted into the resolved 2.3 Å dataset, except for 189–193 and 190–193 in chains C and D respectively. These regions could not be modelled due to the lack of sufficient or sparse electron density in these regions possibly due to extra mobility in these sections. These regions fall within the flexible loop hinge regions which control the opening and closing of the more ordered active site. Furthermore, the structure as a whole was fairly difficult to resolve when compared to most others in this thesis. This is largely due to two factors; the nature of the model used during molecular replacement and due the two different conformers presented in the unit cell. The model used in molecular replacement was the published structure of *CfusAmDH* (PDB: 6IAU), as it showed high sequence identity (33%) it was an ideal model for this dataset. However, this structure was a ligand bound structure; its active site being in a closed conformer. This meant it resolved well for ‘closed’ chains A and C but not so well for chains B and D which were more ‘open’ conformers. RMSD values for *CfusAmDH* (PDB: 6IAU) chain A to chains in the 2.3 Å data set are as follows: A: 2.23 Å for 311 residues, B: 2.60 Å for 298 residues, C: 1.66 Å for 313 residues D: 2.45 Å for 295 residues. Overall, the structure of MATOUAmDH2, even in its ‘closed’ state was still more open than previous nat-

AmDHs including the *Cfus*AmDH structure used for molecular replacement. This made it hard to not only fit the whole chains for B and D but also the dynamic regions for chains A and C as well. To fit these regions averages of both chains in each conformer were typically overlaid into the electron density and the backbones and sidechains built manually in COOT. DynDom was used for further characterisation of movement between conformers using chain C ('open') and D ('closed') as examples. The RMSD values between 'open' and 'closed' overall monomer conformations was calculated as 1.75 Å across 340 residues. The 'fixed' N-Rossman domain between 'open' and 'closed' conformers gave an RMSD value of 0.27 Å. The biggest distance between both conformers, as expected, was seen between the 'flexible' domain (residues 159-308), responsible for the hinging movement, with an RMSD value of 2.78 Å and rotation angle of 17°.

The two different conformers described may have been difficult to solve structurally but also provided information on how the mechanism of this enzyme may work. Residues in regions between Y171-E190 and P240-G255 are responsible for making up this loop hinge region which controls the how open and closed the states are (Figure 64 and 65). This flexible hinge region is probably responsible for bringing conserved, active site residues closer on ligand binding. Active site residues are mentioned in detail in *Results and Discussion 4.2*. Taking a closer look at the active site reveals that when the protein is in its 'closed' form residues V173, Y176 and L180 make significant movements to close the ceiling of the active site (Figure 64). Interestingly, as mentioned above, even the 'closed' forms showed a more 'openness' when compared to previous nat-AmDHs (Figure 65). This is not just an artifact of this being a substrate-free structure but rather due to substitutions in the floor of the active site. F144 and W145 from *Cfus*AmDH are substituted for A147 and C148 respectively, in MATOUAmDH2 (Figure 65). This extra space affords a larger binding pocket and subsequently the possibility to bind larger substrates. Therefore, MATOUAmDH2 is an ideal target for exploring enzyme engineering, to increase its substrate scope. Not only this but both the preliminary kinetic and biotransformation data seems to highlight how stable this enzyme is and its potential for high catalytic specificity. This means MATOUAmDH2 is an ideal candidate to be used in enzyme cascades where the reaction progression is over a few days. Not only this but this is an excellent starting point for engineering this wild-type enzyme to become an even better biocatalyst. This will be highlighted later in (*Results and Discussion 4.3*).

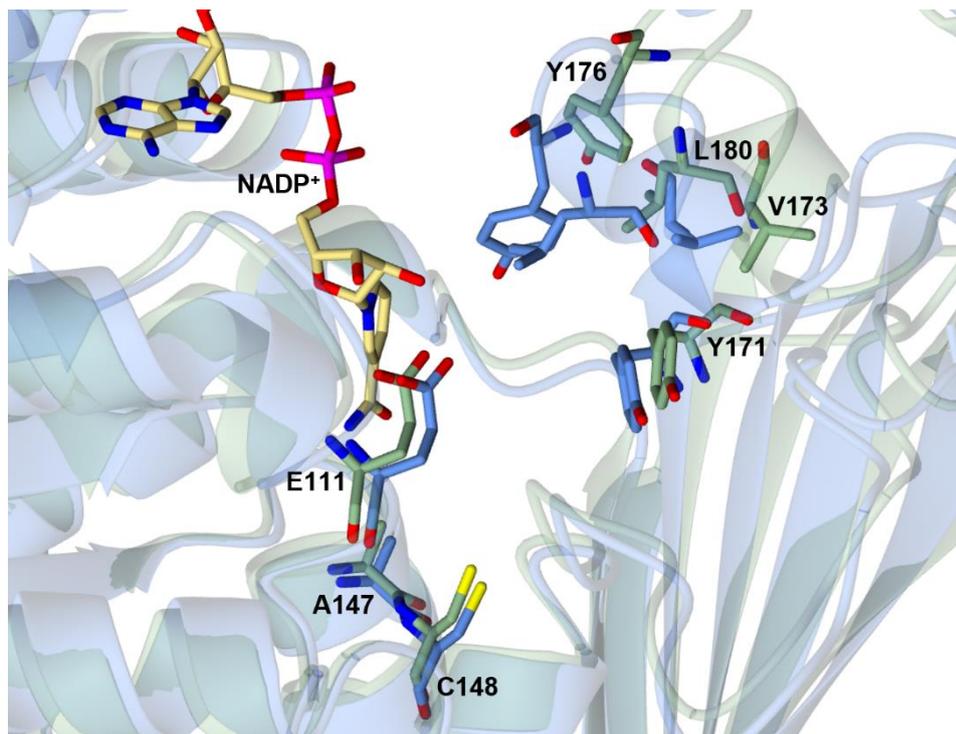


Figure 64: Open and closed forms of pET22b(+) MATOUAmDH2 in complex with NADP⁺. NADP⁺ is displayed in cylindrical form in yellow in all cases. All images are visualised in CCP4mg (v.2.10.11). Open conformer (green) from chain B and closed conformer (blue) from chain C are SSM superimposed. Residues are numbered accordingly.

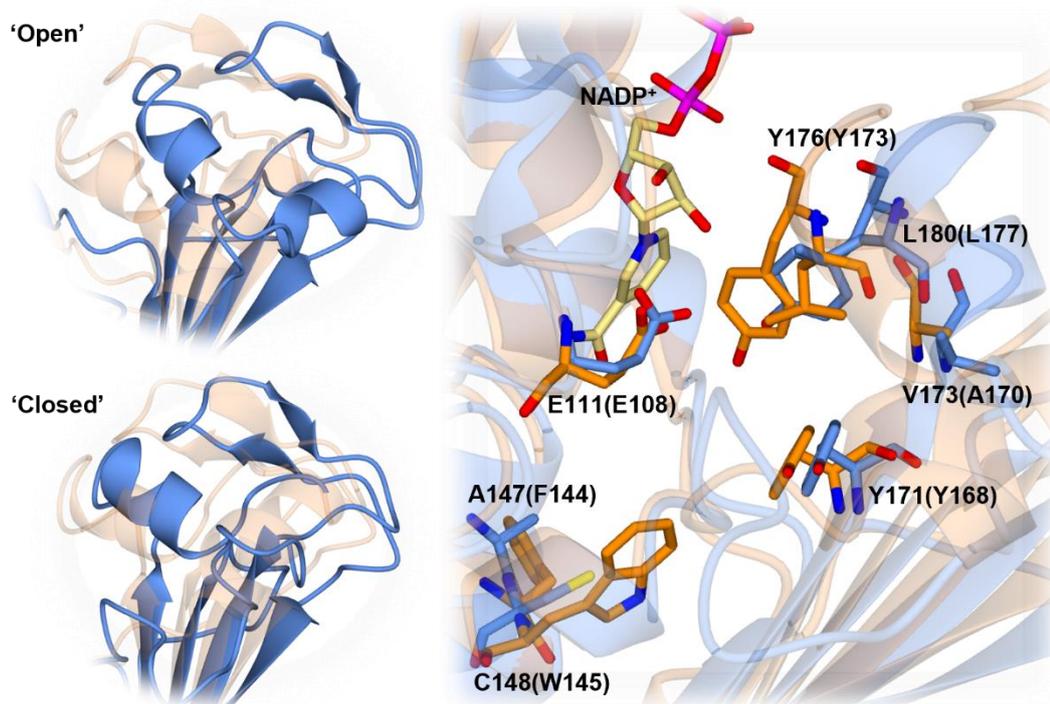


Figure 65: Comparison of ligand free-pET22b(+) MATOUAmDH2 and *CfusAmDH* active sites. All images are visualised in CCP4mg (v.2.10.11). The left-hand panel represents open and closed conformers (blue) from chains B and C respectively SSM superimposed with chain A from *CfusAmDH* (PBD: 6IAU) (orange). The right-hand panel represents a SSM superimposition of key residues from chain C of MATOUAmDH2 (blue) with chain A from *CfusAmDH* (orange). Residues are labelled accordingly with *CfusAmDH* residues in brackets.

4.2 Co-crystallisation of MATOUAmDH2 with cyclohexylamine

As mentioned in *Results and Discussion 4.1* obtaining a ligand bound structure of MATOUAmDH2 via co-crystallisation was not previously successful. As no dataset was obtained for pET22b(+) MATOUAmDH2 bound to a ligand it was subcloned into the in-house pETYSBLIC-3C vector, for reasons specified in *Results and Discussion 4.1.1*. Primers for the MATOUAmDH2 insert and pETYSBLIC-3C vector were designed manually so that the vector primers were independent of insert and that the insert contained the overhangs of approximately 10-20 bp which were complementary to the YSBLIC-3C vector. All primers are tabulated in *Appendices – Table 17*. PCR was successful for amplifying the MATOUAmDH2 insert, both with (2%) and without the addition of DMSO, using standard KOD Hot Start Polymerase protocols and touch down PCR cycling (*Methods 3.1.2*). Bands of interest at around ~1032 bp, as seen in Figure 66, were then excised, and gel extracted from the 1% agarose gel.

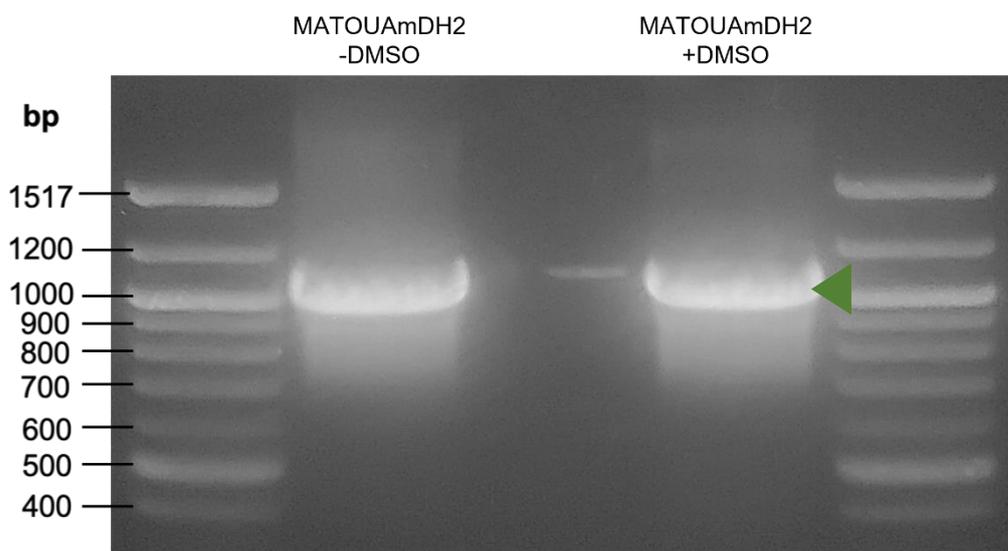


Figure 66: PCR product bands from MATOUAmDH2 insert amplification analysed via gel electrophoresis. The left-hand and right-hand lanes represent a 100 bp QuickLoad NEB marker and associated bp lengths. *Lane MATOUAmDH2 -DMSO*: PCR product from PCR reaction using 20 ng template DNA, KOD Hot Start Polymerase Components without the addition of DMSO. *Lane MATOUAmDH2 +DMSO*: PCR product from PCR reaction using 20 ng template DNA, KOD Hot Start Polymerase Components with the addition of 2% DMSO. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangle represents the bands of interest.

The pETYSBLIC-3C vector was also amplified and linearised using touch down PCR, and standard protocols (*Methods 3.1.2*). The vector also underwent DpnI digestion to remove methylated DNA regions. The amplified, linearised and digested DNA was run on a 1% agarose gel and bands of interest around 6430 bp (Figure 67) were excised and gel extracted to give pure pETYSBLIC-3C vector for use in cloning.

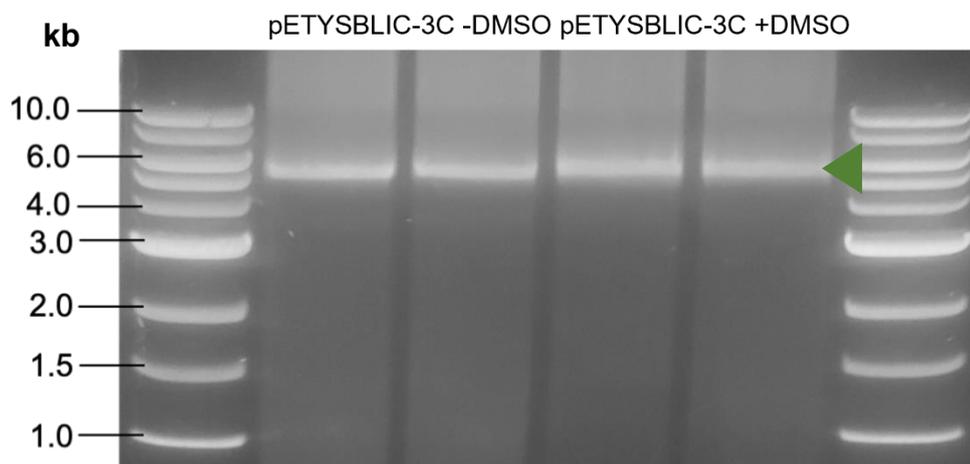


Figure 67: PCR product bands from pETYSBLIC-3C vector amplification, linearisation and digestion, analysed via gel electrophoresis. The left-hand and right-hand lanes represent a 1 kb QuickLoad NEB marker and associated lengths in kilobases (kb). Lane *pETYSBLIC-3C -DMSO*: PCR product from PCR reaction using 150 ng template DNA, Q5 Polymerase Components without the addition of DMSO. Lane *pETYSBLIC-3C +DMSO*: PCR product from PCR reaction using 150 ng template DNA, Q5 Polymerase Components with the addition of 2% DMSO. All samples were subject to overnight DpnI digestion before being loaded. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangle represents the bands of interest.

After both the insert and vector DNA was purified after being extracted from the gel(s) they were subjected to InFusion® cloning to ligate the insert with complementary overhangs to the linearised pETYSBLIC-3C vector. Recombinant plasmids were generated using a 2:1 molar ratio of MATOUAmDH2 inserts to pETYSBLIC-3C vector (*Methods 3.1.3*). Initially a double restriction digest (*Methods 3.1.4.1*) was carried out using NdeI and XbaI. The double digest was not as efficient as expected due to a loss in endonuclease activity, so no presence of a lower molecular weight insert band was seen. However, a higher molecular weight band, above the empty vector band, was observed, indicating the presence of a recombinant plasmid (Figure 68). Successful recombinant plasmids were confirmed using Sanger sequencing.

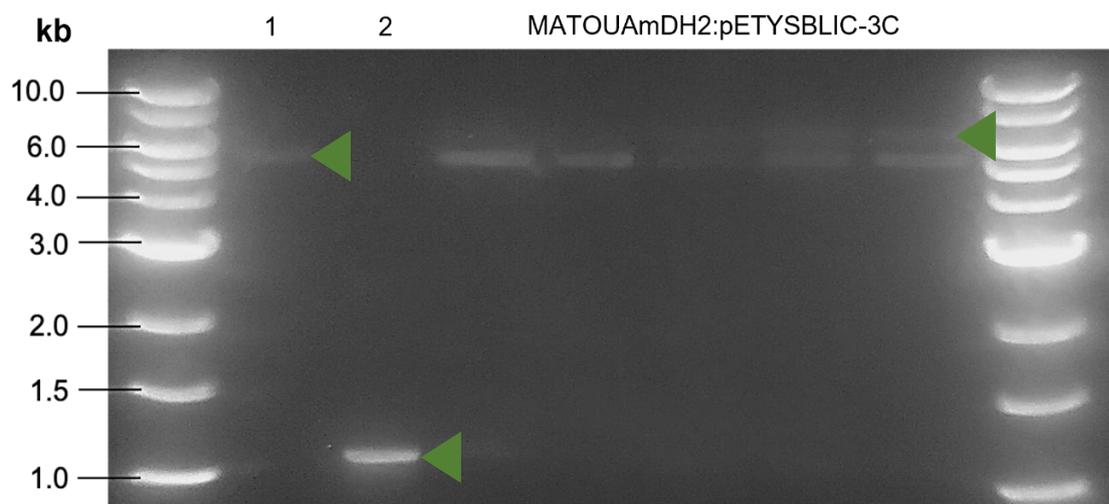


Figure 68: Product bands from the double restriction digest of the recombinant MATOUAmDH2:pETYSBLIC-3C cloning plasmid. The left-hand and right-hand lanes represent a 1 kb QuickLoad NEB marker and associated lengths in kilobases (kb). *Lane 1:* PCR product of the empty pETYSBLIC-3C vector after linearisation, DpnI digestion and PCR purification. *Lane 2:* PCR product of the MATOUAmDH2 insert after PCR purification. *Lanes 3-7 (MATOUAmDH2 :pETYSBLIC-3C):* Isolated and purified recombinant plasmids after InFusion® cloning which were subjected to double restriction digests. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangles represent the bands of interest.

After the generation of the correct recombinant plasmid, it was used to transform *E.coli* BL21(DE3) cell lines, as per *Methods 3.1.5*. The recombinant MATOUAmDH2 pETYSBLIC-3C protein expressed well and was purified using NiNTA and SEC following standard protocol as described in *Methods 3.2*. Yields were slightly lower than when expressed in the pET22b(+) vector, but still yielded sufficient protein for later use at around 20-50 mg per 2 L of LB culture. NiNTA yielded one specific A_{280} protein peak at around 220 mM of low UV imidazole, as evident in Figure 68. Quite a lot of unspecific bound protein was present, eluting at 30 mM imidazole, as evident in the chromatogram (Figure 69). Slight shouldering on the peak was also confirmed to be the correct MATOUAmDH2 protein. The protein was also subjected to SEC which also yielded relatively pure protein with one A_{280} protein peak seen on the chromatogram at around 55 mL; the same as when expressed in the pET22b(+) vector with near identical molecular weights and morphologies (Figure 70). SDS-PAGE analysis also revealed one prominent band present at the correct molecular weight (38.2 kDa) (Figure 71).

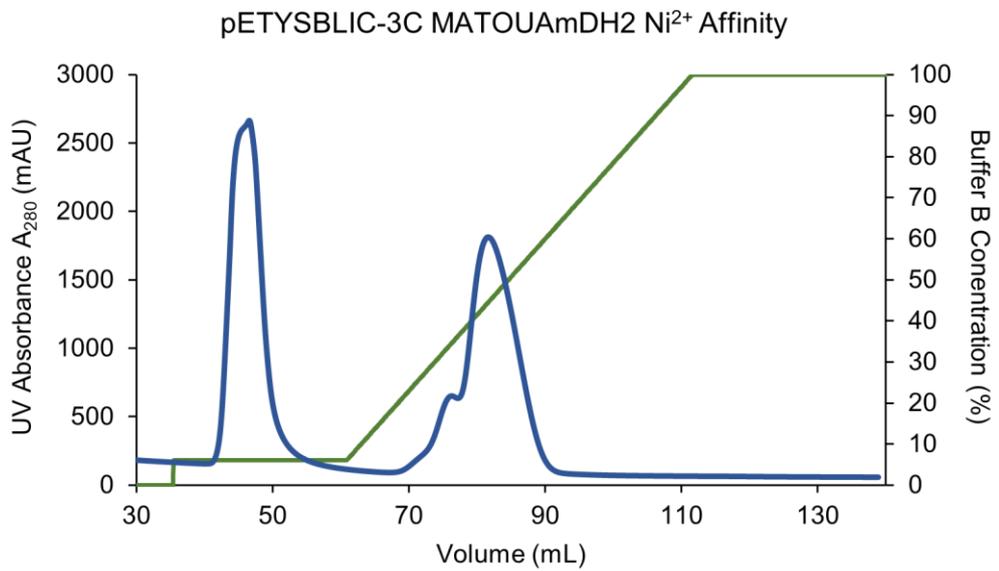


Figure 69: Nickel Affinity Chromatogram for MATOUAmDH2 expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line.

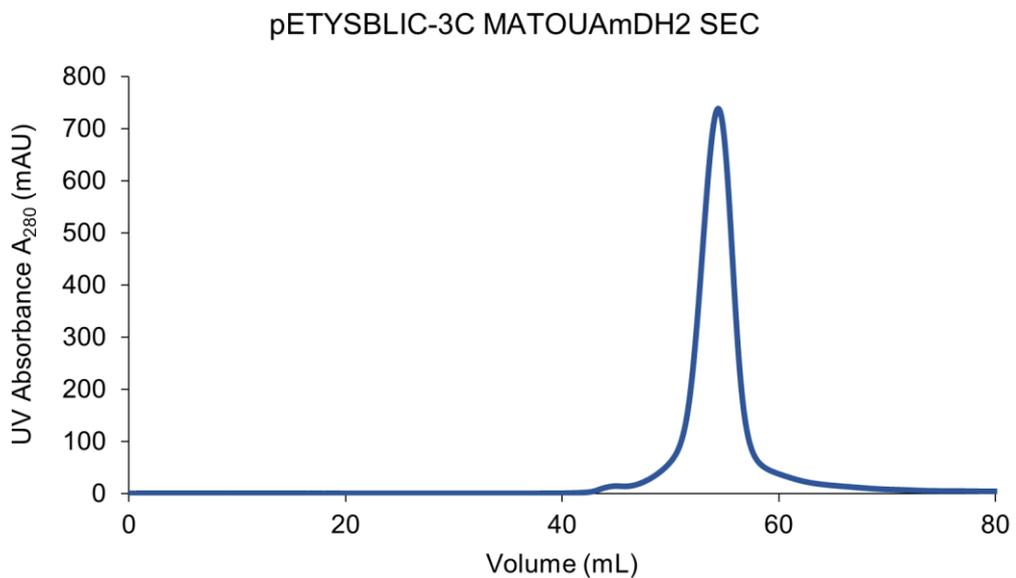


Figure 70: Size Exclusion Chromatogram for MATOUAmDH2 expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue).

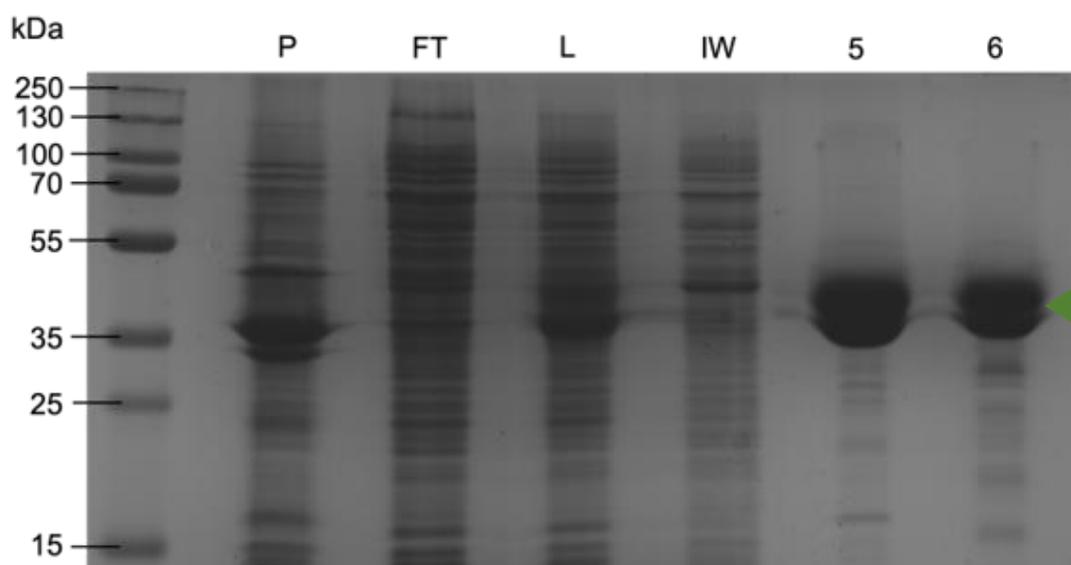


Figure 71: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C MATOUAmDH2 protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane P*: 10 x diluted sample of cell pellet after lysis, resuspended in 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane FT*: 10 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane L*: 10x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lane IW*: 10 x diluted sample of the initial wash, before the imidazole gradient, on the ÄKTA using 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane 5*: 10x diluted combined, pre-determined, samples from the ÄKTA fraction collector after Nickel Affinity Chromatography. *Lane 6*: 20x diluted pooled and concentrated pure protein after Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

MATOUAmDH2 was subcloned into the pETYSBLIC-3C mainly to obtain a construct that could produce more reliably fishable, better diffracting crystals, with the ultimate goal being to attain a ligand bound crystal structure (*Results and Discussion 4.1.1*). Protein was concentrated until either 30 mg.mL⁻¹ or 15 mg.mL⁻¹ (final drop concentrations at 15 mg.mL⁻¹ and 7.5 mg.mL⁻¹ respectively). Protein was incubated with 10 mM NADP⁺ (final drop concentration at 5 mM) and 10 mM cyclohexylamine **21b** or 2-aminocyclohexanone **31a** as per *Methods 3.3.1 & 3.3.2*. The purified MATOUAmDH2 protein expressed in the pETYSBLIC-3C vector with cofactor and ligand was subjected to a range of commercially available pre-dispensed crystallisation screens, namely Index and CSS screens using a 1:1 ratio of mother liquor to protein complex. Initial screening yielded promising crystal forms with small cuboid morphologies (Figure 72). These crystals were obtained in drops containing 0.2 M salt (either NaCl or MgCl₂), 0.1 M buffer pH 5.5-7.5 (either Bis-Tris or HEPES) and 25% (w/v) PEG 3350 at 7.5 mg.mL⁻¹ and 15 mg.mL⁻¹ protein. In an attempt to obtain larger crystals these conditions were optimised into MAXI 48-well designed screens using 15 mg.mL⁻¹ protein with 5 mM NADP⁺ (final drop concentrations) and 10 mM of either cyclohexylamine **21b** or 2-aminocyclohexanone **31a** in a 1:1 ratio of

mother liquor to protein complex. Unfortunately, these drops yielded needle-like or thin clustered plates crystal forms as seen in Figure 73.

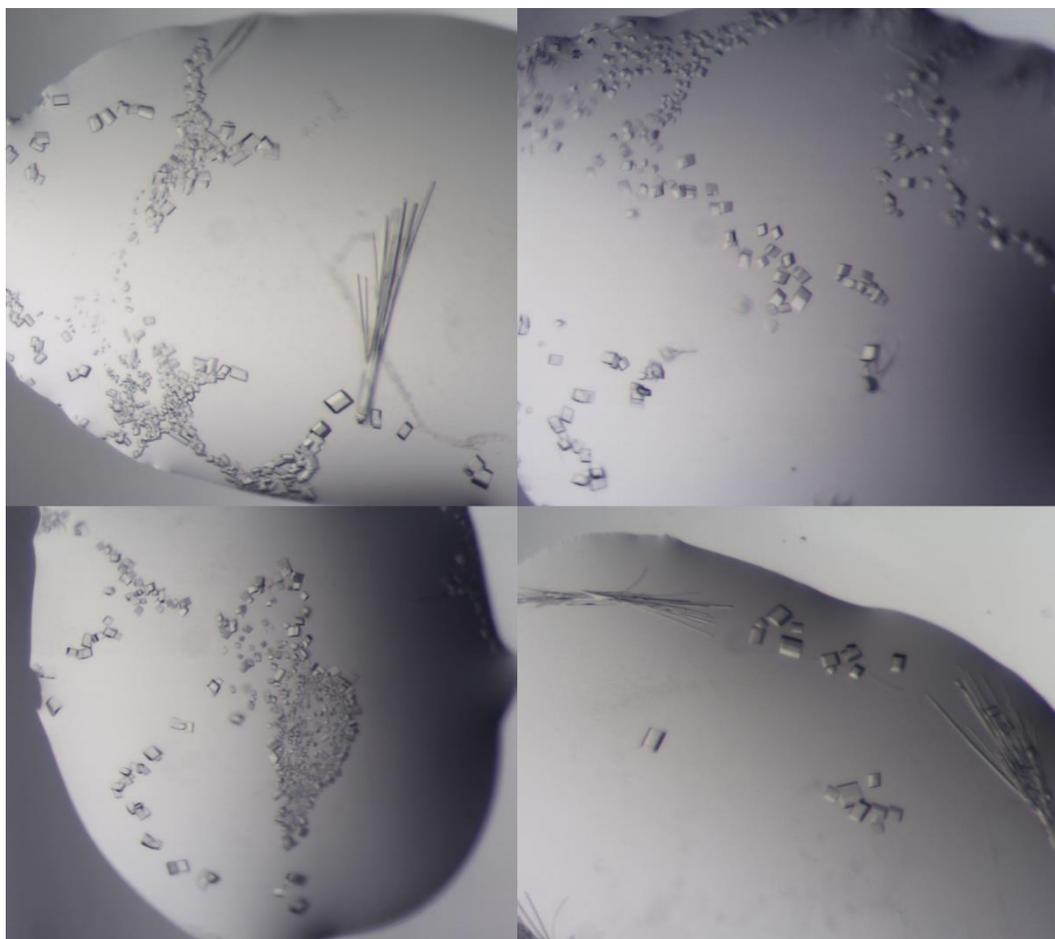


Figure 72: Crystal forms from initial screening of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and either cyclohexylamine or 2-aminocyclohexanone. Microscopy images of small cuboid crystals from a pre-dispensed Index (MRC-96 well) screen.

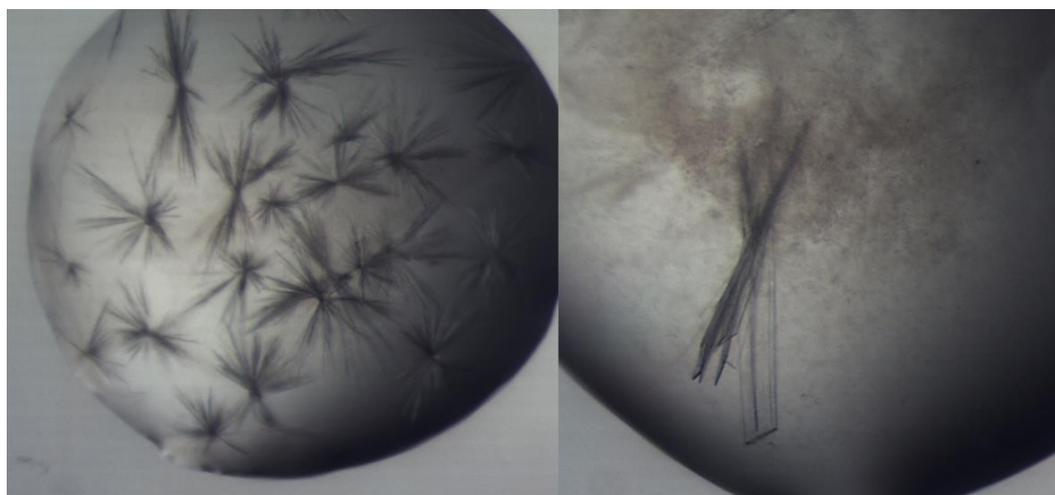


Figure 73: Crystal forms from 48-well optimisations of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and either cyclohexylamine or 2-aminocyclohexanone. Microscopy images of crystal forms from a pre-designed optimisation (MAXI-48 well) screen.

Due to the nature of the 48-well optimisation crystals, compared to the 96-well Index screen crystals, further optimisations were designed and manually dispensed into 96-MRC screens using 1:1 ratios as above. These optimisations set up in smaller wells yielded successful results. In these experiments not only were larger crystals obtained, but they also resembled either cuboid or single thick plates which were easily accessible when fishing (Figure 74). The best hits were in drops containing 0.2 M MgCl_2 , 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350 with $10 \text{ mg}\cdot\text{mL}^{-1}$ MATOUAmDH2 in complex with 5 mM NADP^+ (final drop concentrations) and 10 mM cyclohexylamine **21b**. Most of the crystals seen in Figure 74 were obtained after only ~24 h of crystal growth where they were flash cooled using liquid nitrogen without any additional cryoprotectant added and sent to the Diamond synchrotron for data collection.

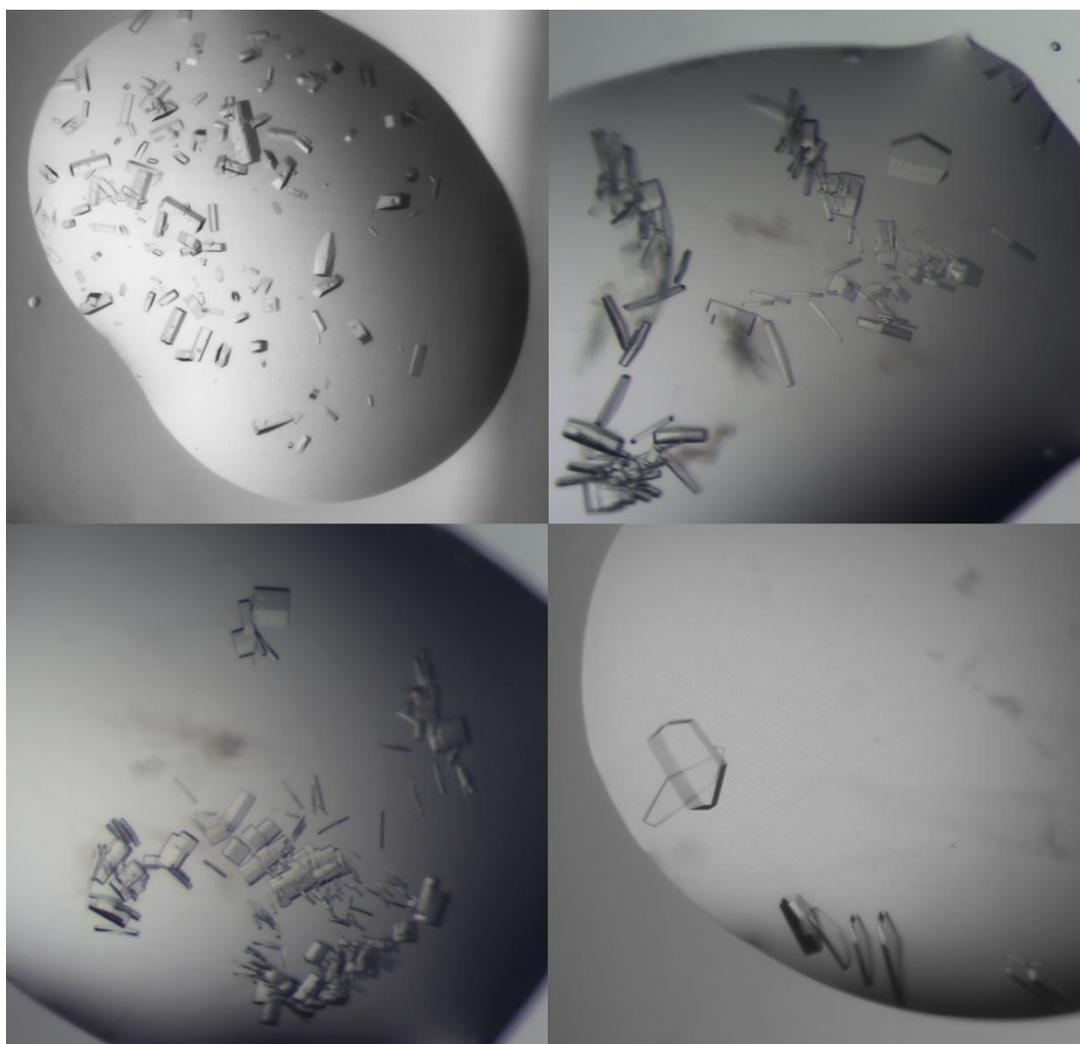


Figure 74: Crystal forms from 96-well optimisations of pETYSBLIC-3C MATOUAmDH2 in complex with NADP^+ and either cyclohexylamine or 2-aminocyclohexanone. Microscopy images of crystal forms from a pre-designed optimisation (MRC-96 well) screen.

A dataset from Diamond beamline I03 was obtained and a new structure obtained and refined to a resolution of 2.08 Å. There was only one molecule in the asymmetric unit occupying a $P3_21$ space group. The unit cell dimensions for this dataset were = 93.51 b = 93.51; c = 75.79 $\alpha = \beta = 90^\circ \gamma = 120^\circ$. The data and refinement statistics are reported in Table 11. This data set was deposited in the PDB under accession code 7R09.

Table 11 – Data collection and refinement statistics for the 2.08 Å pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and cyclohexylamine.

2.08 Å pETYSBLIC-3C MATOUAmDH2:NADP⁺:cyclohexylamine

Beamline	Diamond I03
Wavelength (Å)	0.97625
Resolution (Å)	55.40-2.08 (2.13-2.08)
Space Group	$P3_21$
Unit cell (Å)	a = 93.51 b = 93.51; c = 75.79 $\alpha = \beta = 90^\circ \gamma = 120^\circ$
No. of molecules in the asymmetric unit	1
Unique reflections	23365 (1768)
Completeness (%)	99.9 (99.9)
R _{merge} (%)	0.07 (1.28)
R _{p.i.m.}	0.02 (0.41)
Multiplicity	20.2 (20.7)
$\langle I/\sigma(I) \rangle$	24.4 (2.7)
Overall B from Wilson plot (Å ²)	48
CC _{1/2}	1.00 (0.94)
R _{cryst} / R _{free} (%)	23.3/29.8
r.m.s.d 1-2 bonds (Å)	0.007
r.m.s.d 1-3 angles (°)	1.42
Avg main chain B (Å ²)	58
Avg side chain B (Å ²)	61
Avg water B (Å ²)	57
Avg NADP ⁺ B (Å ²)	54
Avg CHA B (Å ²)	30

Brackets refer to data in the highest resolution shell.

Overall, the structure of this co-crystallised MATOUAmDH2 is very similar to the apo structure (PDB: 7ZBO, *Results and Discussion 4.1*). The same residues make up a structure largely comprised of a N-terminal Rossman domain, and a more C-terminal beta pleated sheet morphology comprised of again 8 beta strands, both seen in the secondary structure visualization (Figure 75).

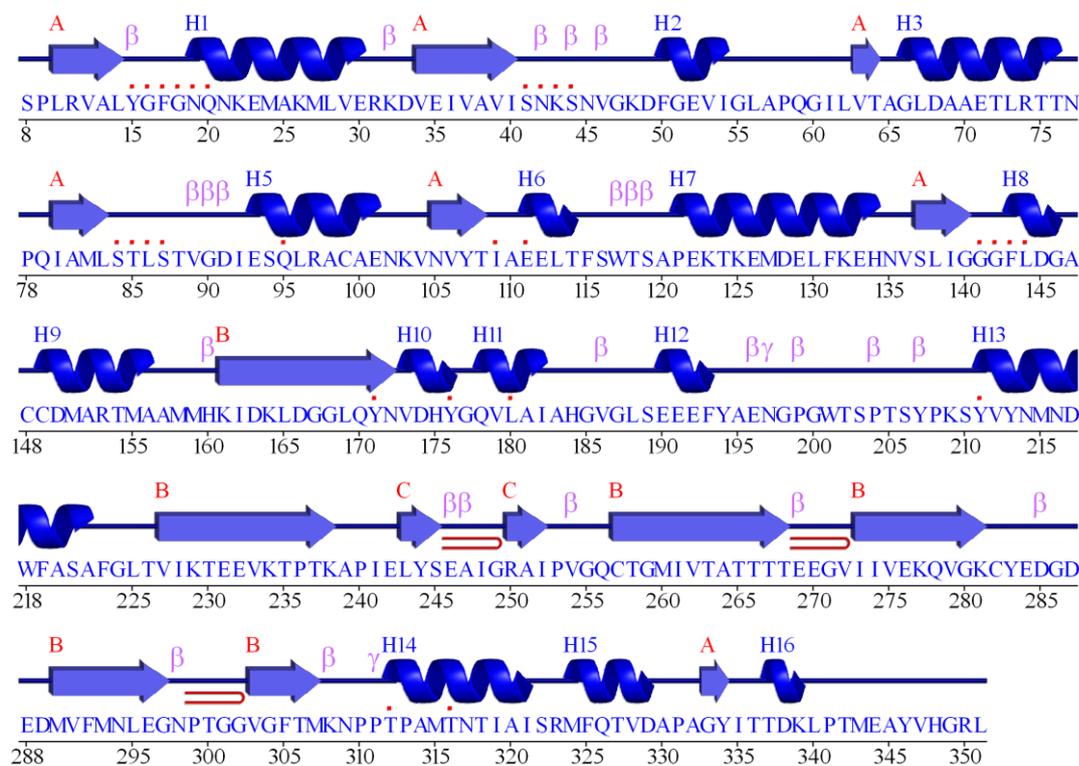


Figure 75: Secondary structure visualisation of cyclohexylamine bound pETYSBLIC-3C MATOUAmDH2. Visualisation was carried out using PDBsum. Secondary structures are separated into alpha helices or beta strands, beta hairpins are indicated in red turns. The sequence and sequence numbers are displayed as reference points.

The same flexible loop hinge region was observed, which was seen in both the apo structure and other nat-AmDHs⁹¹, which controls the opening and closing of the binding pocket. The one molecule in the asymmetric unit is most similar to the ‘closed’ conformers corresponding to chains A and C in the apo structure (Figure 76) with an RMSD values of 0.40 Å for chain C in the apo structure across 340 C α atoms but only 1.52 Å for 332 C α atoms in chain B. The binding pocket of this one ligand bound molecule also has an approximate diameter of 12 Å at the widest point; similar to the ‘closed’ apo chains, which again is also larger than closed, ligand bound, forms of previous nat-AmDHs (*Results and Discussion 4.1.1*).⁹¹ In this ligand bound chain there is clear omit $F_o - F_c$ electron density for both NADP⁺ and cyclohexylamine (Figure 77).

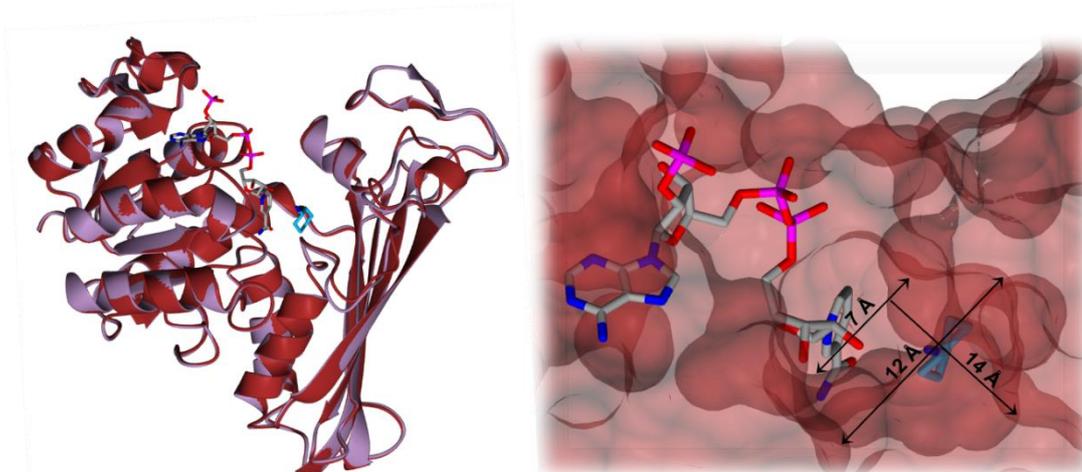


Figure 76: Structure and binding pocket of pETYSBLIC-3C MATOUAmDH2 in complex with NADP⁺ and cyclohexylamine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel displays the overall structure of the one chain of MATOUAmDH2 (red) in complex with NADP⁺ (grey) and cyclohexylamine (blue) SSM superimposed with chain C of the apo structure of MATOUAmDH2 (lilac). The right-hand panel displays the binding pocket of the ligand bound structure in complex with both NADP⁺ (grey) and cyclohexylamine (blue). Arrows display the width of the narrowest and widest opening of the binding pocket as well as the depth. All measurements are in Ångstroms, approximated from the inbuilt scale.

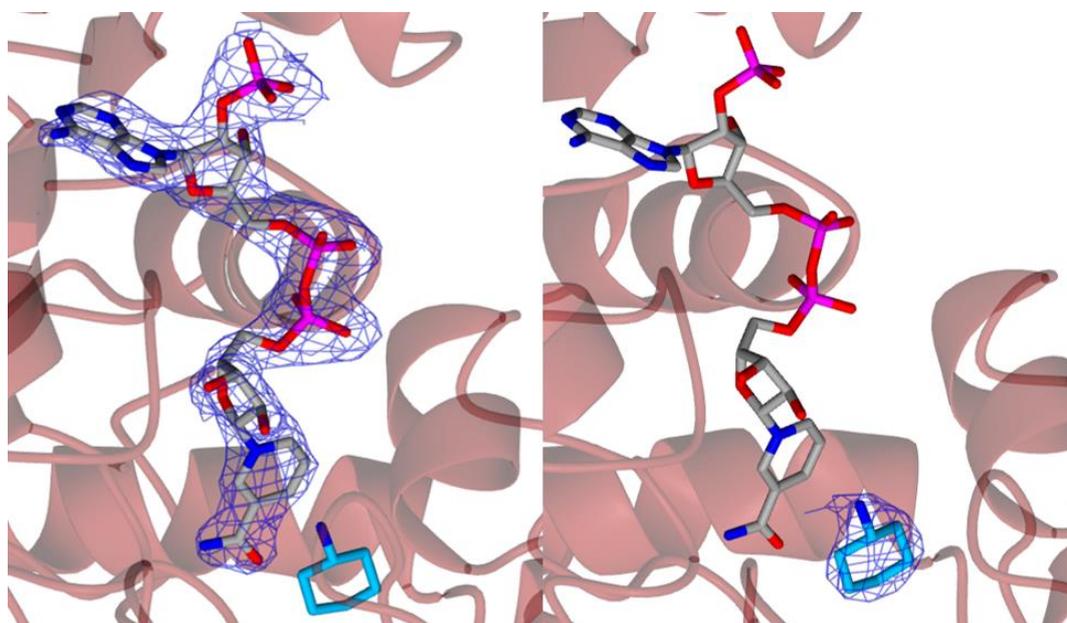


Figure 77: Omit maps for both NADP⁺ and cyclohexylamine in respect to the pETYSBLIC-3C MATOUAmDH2 structure. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel displays the omit map for NADP⁺ (grey) displayed in cylindrical wire form in blue. The right-hand panel displays the omit map for cyclohexylamine (blue) where the $F_o - F_c$ electron density is displayed in cylindrical blue wire at a map level of 3σ .

4.2.1 Discussion

The gene encoding MATOUAmDH2 was sub-cloned into the in-house designed pETYSBLIC-3C vector in an attempt to improve crystal morphologies when binding ligands due to problems experienced when attempting to co-crystallise MATOUAmDH2 with a range of ligands when using protein generated from the initial pET22b(+) vector (*Results and Discussion 4.1*). The gene encoding MATOUAmDH2 was successfully sub cloned into the pETYSBLIC-3C vector where both amplifications of the inserts and vectors were relatively straight forward with manually designed primers. Standard protocols for both the KOD Hot Start and Q5 Polymerases were barely altered with only the addition of touch down steps. In-Fusion® cloning was also successful following the standard protocol closely and using a recommended 2:1 mole-based ratio of insert to vector; mole-based ratios, i.e., considering differences in molecular weights, seemed to be crucial when compared with the use of molar ratios which gave insufficient ligation. Ultimately the successful generation of recombinant plasmid was confirmed by a mixture of a double restriction digest and sequencing results provided by Eurofins Genomics.

Protein expression in a BL21(DE3) system and purification using NiNTA and SEC was also successful in yielding relatively pure protein. The typical yields from the pETYSBLIC-3C preps were slightly lower than the pET22b(+) based preps (*Results and Discussion 4.1*). This may largely be due to non-optimal conditions during expression. In the pETYSBLIC-3C preps there seems to be a rather large amount of non-specific bound protein which elutes during the initial imidazole wash (30 mM) during NiNTA chromatography, which is removed after this purification step. However, this may be due to co-expressed background proteins due to factors such as time of induction, amount of IPTG used and expression temperatures. If obtaining higher yields were necessary this could be optimised via expression tests, however the yields that were obtained were high enough for the further experiments needed and the protein produced was relatively pure after Size Exclusion Chromatography.

When producing a crystal bound structure of MATOUAmDH2 it appeared that well size seemed to be crucial. On scaling up optimisations from 96-well Index screens to 48 well plates crystal forms changed to needle like or clustered forms; proving hard to fish. The only different variable that made a change in this case was to scale back down optimisations into 96-well plates. This provided better crystal morphologies often representing large thick single plates or cubes. The best dataset from these fished crystals showed one molecule in the asymmetric unit with clear F_o-F_c electron

density for both NADP⁺ and cyclohexylamine **21b**. Using the monomeric structure of the apo pET22b(+) MATOUAmDH2 as a model for molecular replacement gave us a very good initial structure solution which was then ultimately refined to R factors of 23.3% and 29.8% for R_{crystal} and R_{free} respectively. This one chain clearly represented a 'closed' conformer, as expected of a ligand bound structure, which represented the 'closed' conformers from the apo-structure, namely chains A and C. This is also seen when comparing the dimensions of the binding pockets where both are approximately 12-13 Å at the widest point. Additionally, this confirmed that this dynamic movement seen between the chains in the apo structure (*Results and Discussion 4.1*) was indeed not an artifact of being an unbound structure but instead truly represents the opening and closing of the binding pocket in a claw like mechanism.

Obtaining a ligand bound crystal structure was crucial in understanding how key active site residues are involved in the mechanism of action. As previously reported the catalytic glutamate residue within the active site is responsible for activating the amine donor for nucleophilic attack at the carbonyl centre. In addition, this glutamate residue is also responsible for coordinating and stabilising the iminium ion during hydride transfer from NAD(P)H. In the case of the MATOUAmDH2 active site E111 is responsible for this role. Looking at the active site of the pETYSBLIC-3C MATOUAmDH2 structure E111 is positioned 2.53 Å from the amine group of cyclohexylamine in the active site (Figure 78). This positions E111 in the ideal position for both nucleophilic attack at the carbonyl centre as well as stabilisation of generated iminium intermediates during the reaction progression. Additionally aromatic active site residues such as F143, Y171, Y176 and Y211 close the active site in a hydrophobic cage along with the NADP⁺ cofactor on the opposite face (Figure 78). As previously mentioned in *Results and Discussion 4.1.1*, the active site floor is closed by smaller substituted residues when compared to other nat-AmDHs such as C148 and A147 in MATOUAmDH2 (Figure 78). This, along with, the extended space afforded at the ceiling of the active site, made MATOUAmDH2 a promising target for binding larger substrates and improving substrate scope through enzyme engineering.

Site directed mutations were designed at various positions within this active site cage, initially just using alanine substitutions to assess functionality of each of these targeted active site residues. Mutants F143A, L144A, L169A, L180A, M215A and T312A were all designed, produced, and assessed, as discussed in *Results and Discussion 4.3*. F143A was designed on the principle that altering the steric effect at this site could lead to the binding of bulkier substrates, as is demonstrated by the

presence of A147 of MATOUAmDH2 in replace of typical larger residues at this position such as F144 from *CfusAmDH2*.⁹¹ L144A and L169A were designed since normally present in these positions are more hydrophilic, smaller, residues such as Q141 and T166 from *CfusAmDH* respectively. Therefore, investigating whether a hydrophobic rather large residue i.e., leucine, was crucial in this position was the target. L180A and M215A were designed to specifically target the sterics and functionalities of the ceiling and floor of the active site where the most dynamic space provided to the substrate cage is observed. T312A was designed with the idea not only of investigating functionality at the back of the substrate cage but also to introduce some more hydrophobicity to the binding pocket and to test the consequent effects of this change.

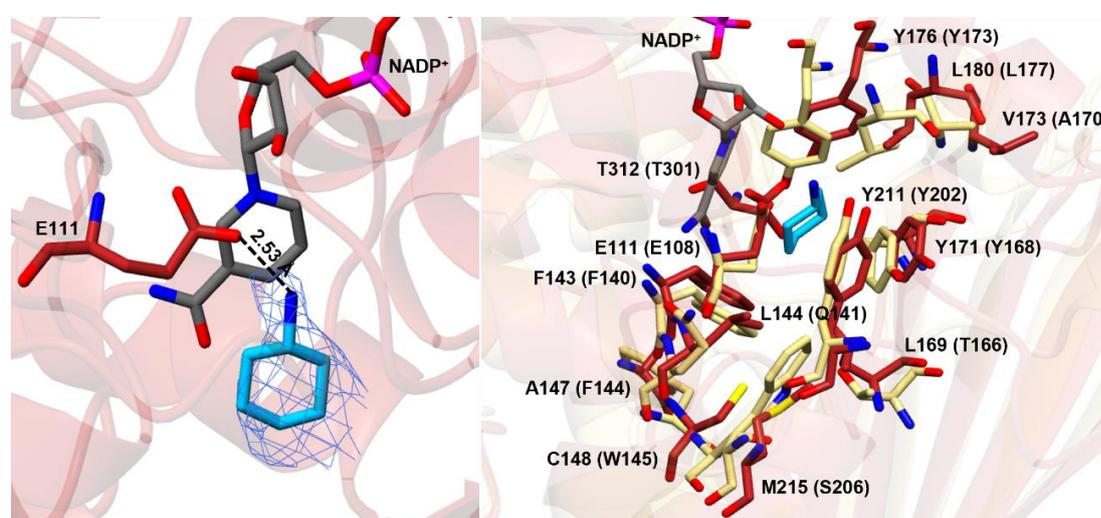


Figure 78: In-depth analysis of the substrate cage and active site residues in MATOUAmDH2. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel displays the positing of Glu111 from MATOUAmDH2 (red) in complex with NADP⁺ (grey) and cyclohexylamine (blue). Hydrogen bonding interactions are displayed by a dash line and distances are indicated in Ångstroms. The right-hand panel displays the binding pocket and substrate cage of the ligand bound structure in complex with both NADP⁺ (grey) and cyclohexylamine (blue). MATOUAmDH2 (red) has been SSM imposed with *CfusAmDH* (yellow) (PDB: 6IAU), and residues are labelled accordingly with *CfusAmDH* residues bracketed.

4.3 Engineering MATOUAmDH2 for biocatalysis

For reasons explained in Results and Discussion 4.2.1, six mutants were isolated and designed for an alanine-based site directed mutagenesis screen. Primers for F143A, L144A, L169A, L180A, M215A and T312A, in the pETYSLIC-3C vector, were all manually designed as in *Methods 3.1.1.1* and primers are detailed in *Appendices – Table 17*. PCR was carried out as before, using standard KOD Hot Start Polymerase protocols and incorporation touch down steps as had previously worked during the sub-cloning of MATOUAmDH2 (*Results and Discussion 4.2*). However, this did not produce any bands of interest. PCRs were optimised varying annealing temperatures, type of polymerase, altering concentrations of both template DNA and primers. New primers were also designed using Takara Bio® primer designer. Unfortunately, none of these optimisations yielded any successful amplifications. Subsequently, touch downs were removed from the PCR cycling conditions and successful amplification of regions of interest was seen for all resulting mutants when using both types of primers. Manually designed primers gave the best results and were used in almost all further cases. (Figure 79 and 80).

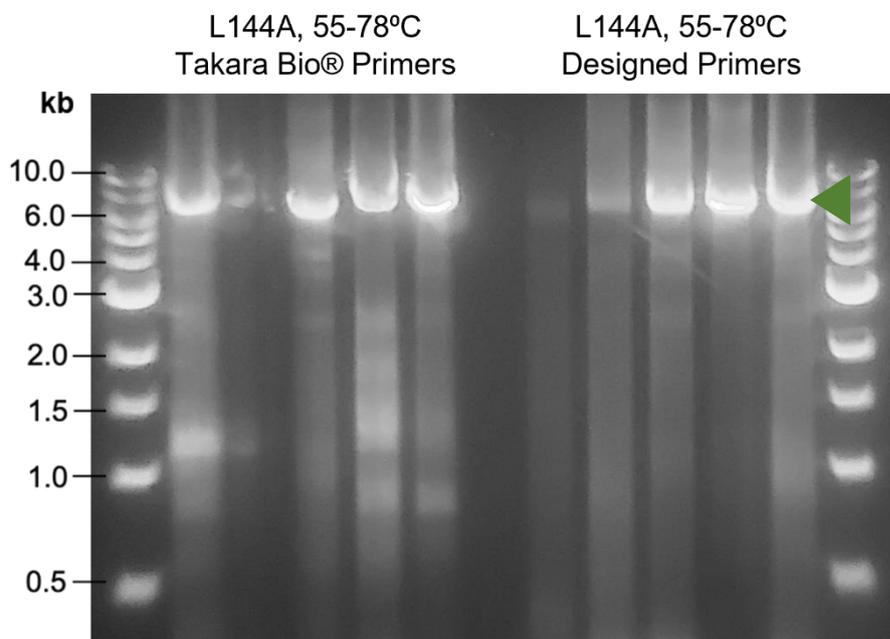


Figure 79: PCR product bands from an exemplar L144A thermogradient trial using both types of primers. The left-hand and right-hand lanes represent a 1 kb QuickLoad NEB marker and associated kb lengths. The left-hand collection of bands represent PCR reactions run at a range of annealing temperatures (55-78°C) using L144A Takara Bio® primers. The right-hand collection of bands represent PCR reactions run at a range of annealing temperatures (55-78°C) using L144A manually designed primers. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangle represents the bands of interest.

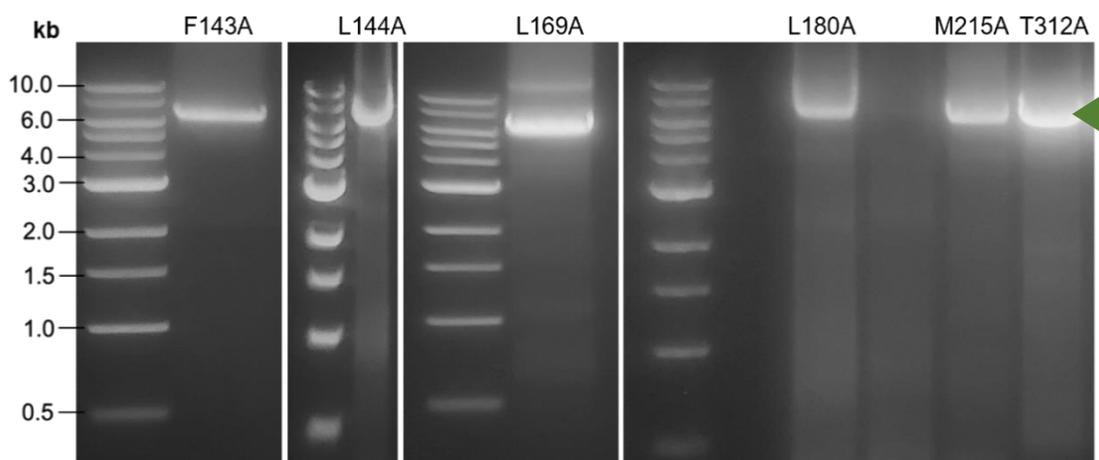


Figure 80: PCR product bands from MATOUAmDH2 alanine screen site directed mutagenesis, analysed via gel electrophoresis. The left-hand lanes of each cropped gel represent a 1 kb QuickLoad NEB marker and associated kb lengths. All lanes are labelled with specific mutant PCR products. This figure represents a collage of cropped images of bands that were excised from individual agarose gels. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangle represents the bands of interest.

As all the molecular weights of the resulting products were very similar, only differing in one codon substitution, for InFusion® cloning 50 ng of DNA was used in each case. The volume required for the InFusion® reaction (as per *Methods 3.1.3*) was calculated using the concentrations determined by a NanoPhotometer (Implen). Successful ligation proceeded; colonies were selected, recombinant plasmids isolated, and sequencing results confirmed the presence of codon substitutions at specific mutant sites.

Recombinant mutant plasmids were then transformed into BL21(DE3) cells, as in *Methods 3.1.5*, and expressed and then purified via NiNTA and SEC (*Methods 3.2*). Chromatograms display one peak on both the nickel affinity traces as well as the size exclusion traces for all mutants (Figure 81 and 82); protein did not seem to be co-eluted or co-overexpressed with anything else. SDS-PAGE analysis revealed one over expressed band at the expected molecular weight for each mutant (Figure 83 and 84). Yet, during initial nickel affinity runs both F143A and L144A precipitated during fraction collection. Aggregates were pelleted using centrifugation (~13,000 rpm for 5 minutes) and then the supernatant was collected and filtered through diameter-based syringe filters for loading onto Superdex® columns. Consequently, after these size exclusion runs precipitate again formed for both these mutants, and therefore F143A and L144A were not taken further in subsequent biochemical analysis.

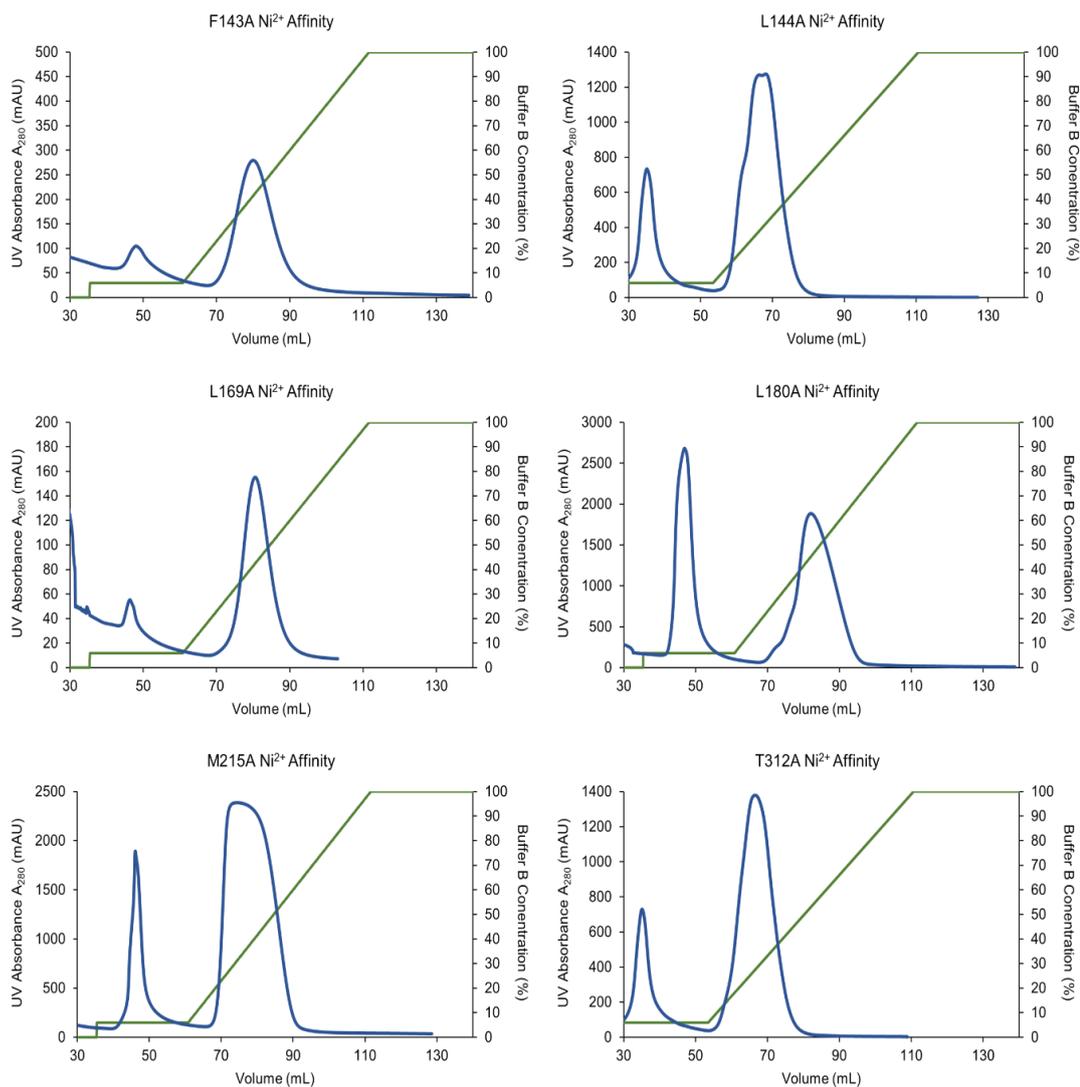


Figure 81: Nickel Affinity Chromatogram for MATOUAmDH2 alanine mutants expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line. All chromatograms are labelled with corresponding mutant names.

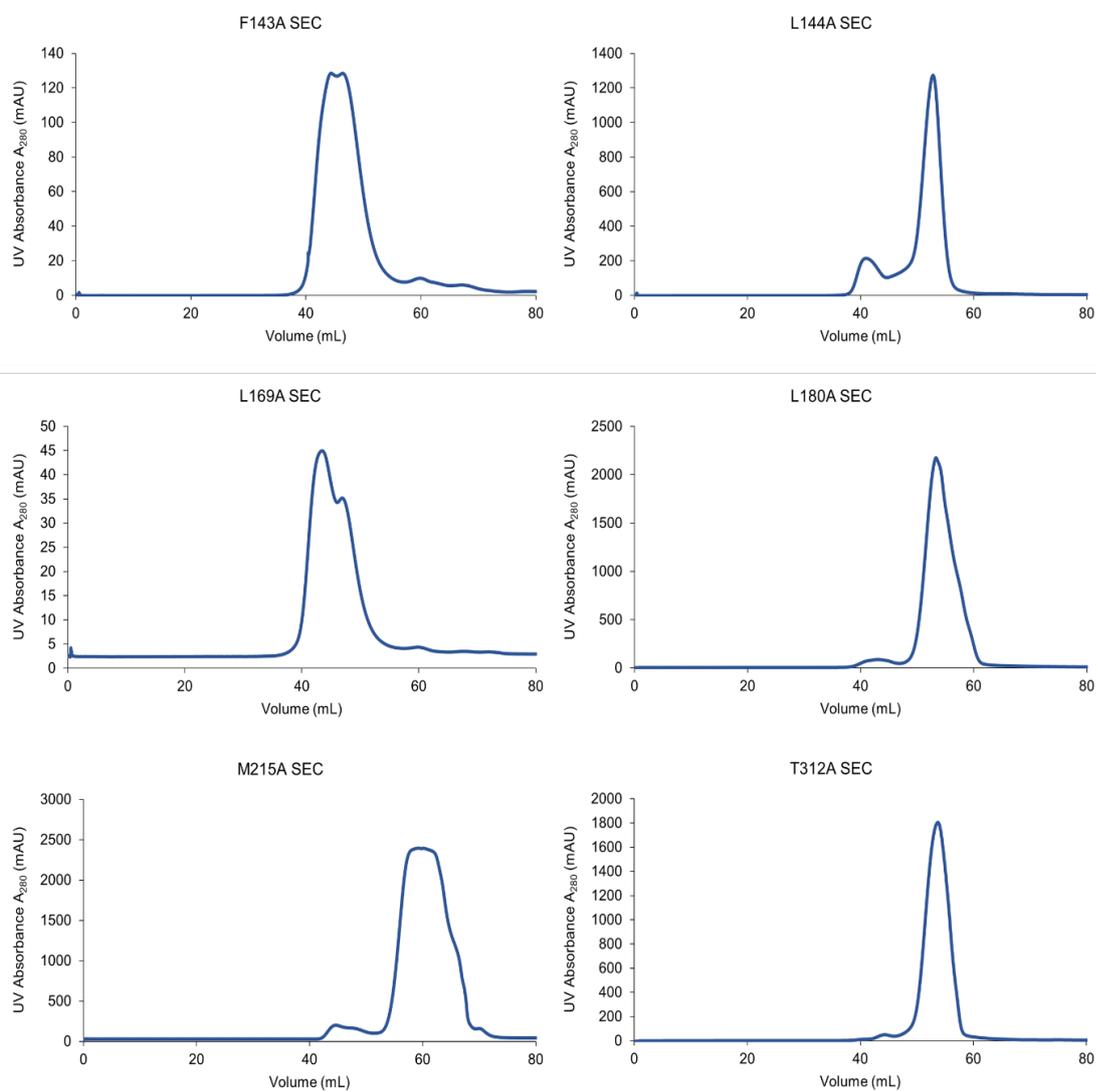


Figure 82: Size Exclusion Chromatogram for MATOUAmDH2 alanine mutants expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). All chromatograms are labelled with corresponding mutant names.

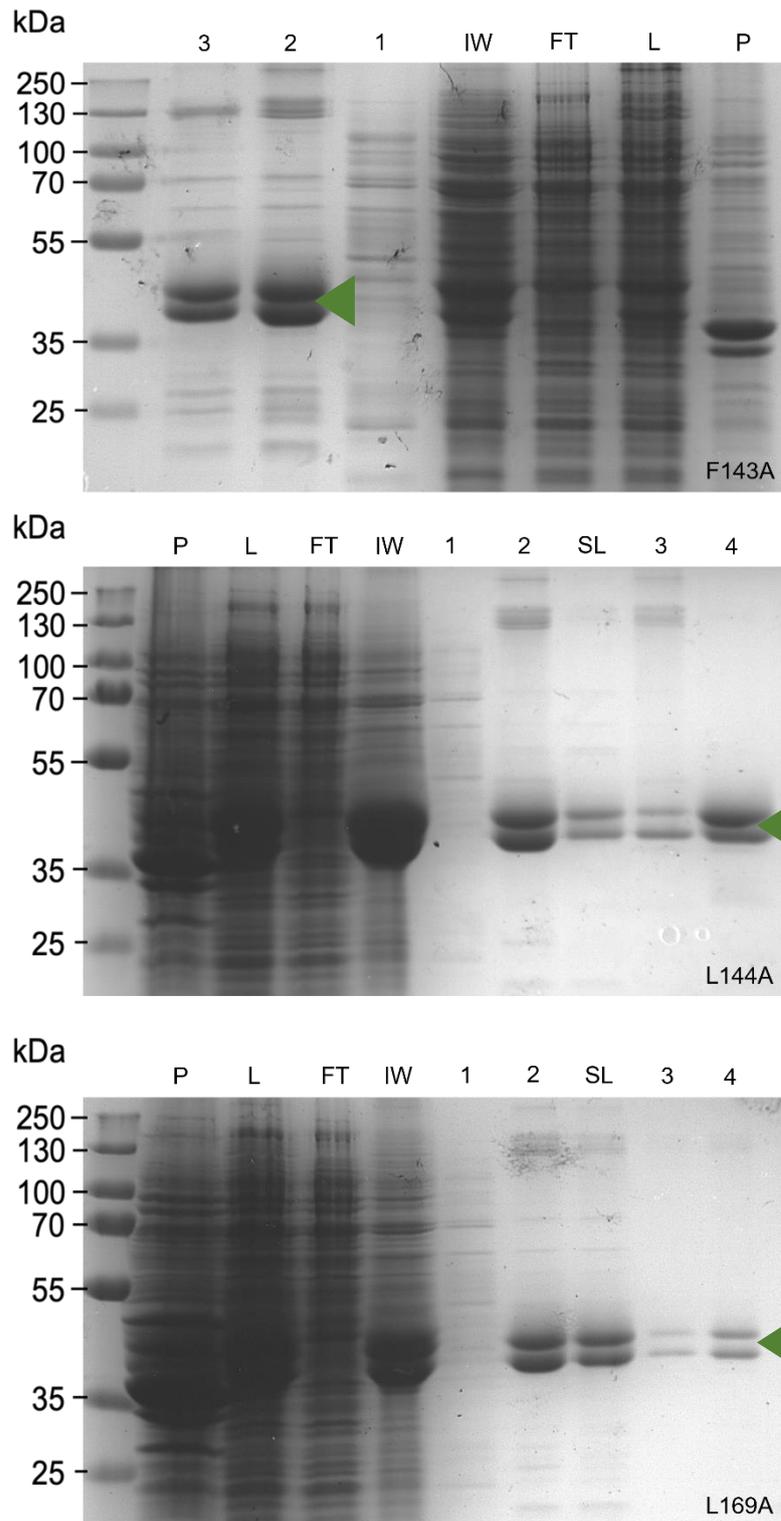


Figure 83: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C MATOUAmDH2 F143A, L144A and L169A protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane P*: 2 x diluted sample of cell pellet after lysis, resuspended in 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane L*: 2 x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lane FT*: 2 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane IW*: 2 x diluted sample of the initial wash, before the imidazole gradient, on the ÄKTA using 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lanes 1&2*: 2 x diluted combined, pre-determined, non-specific and specific samples from the ÄKTA fraction collector after Nickel Affinity Chromatography. *Lane SL*: 2 x diluted pooled and concentrated protein loaded onto the Superdex® column. *Lanes 3&4*: 2 x diluted samples from fractions during Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangles represent the bands of interest.

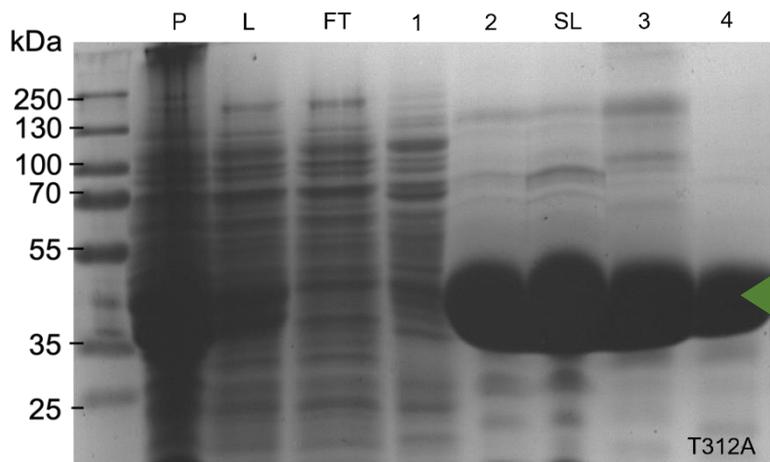
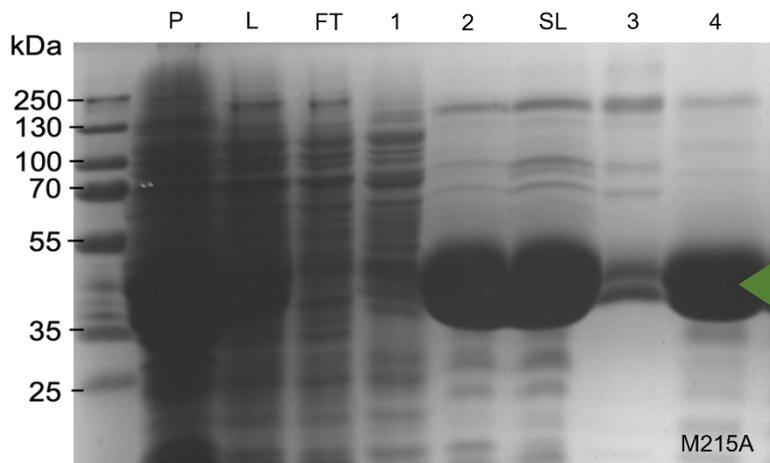
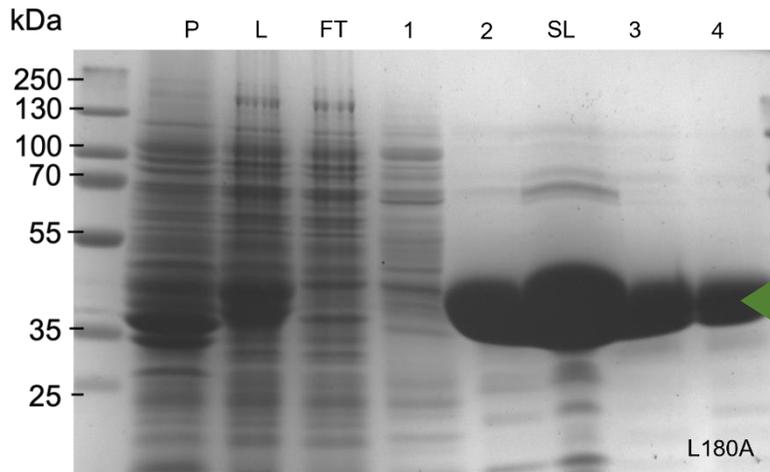


Figure 84: SDS-PAGE analysis of protein fractions from pETYSB LIC-3C MATOUAmDH2 L180A, M215A and T312A protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane P*: 2 x diluted sample of cell pellet after lysis, resuspended in 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane L*: 2 x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lane FT*: 2 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lanes 1&2*: 2 x diluted combined, pre-determined, non-specific and specific samples from the ÄKTA fraction collector after Nickel Affinity Chromatography. *Lane SL*: 2 x diluted pooled and concentrated protein loaded onto the Superdex® column. *Lanes 3&4*: 2 x diluted samples from fractions during Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangles represent the bands of interest.

Alanine screen mutants were used for biochemical characterisation to analyse catalytic characteristics. Kinetic parameters were determined for all mutants using UV-vis spectrometry (*Methods 3.4.3*) and compared with the wild-type enzyme also expressed and purified in the pETYSBLIC-3C vector. L169A precipitated on addition to solutes (cofactor and ligand) in 2 M Ammonium Formate buffer pH 8.0, therefore no accurate change in absorbance for this mutant was able to be recorded. L169A appeared to not be stable at room temperature and so was not taken further for kinetic parameter analysis or use in subsequent biotransformations. As mentioned in *Methods 3.4.3*, the amount of protein used was increased 10-fold, to 0.25 mg.mL⁻¹, in comparison to the pET22b(+) expressed wild-type enzyme as well as a reduction in cuvette block temperature from 50°C to 25°C. ΔA_{340} was converted to velocity which was plotted against the cyclohexanone concentrations used and Michaelis Menten curves were fitted (Figure 85); kinetic parameters were then calculated from this ideal fit and are reported in Table 12.

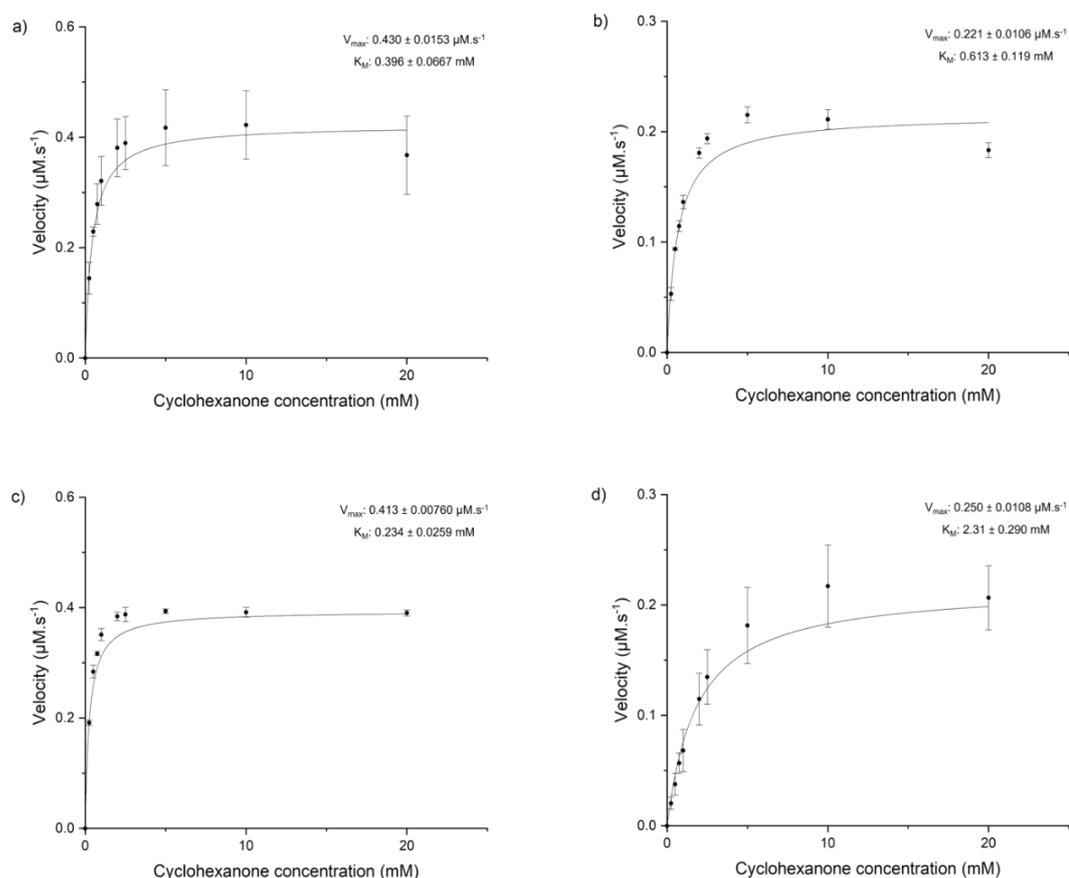


Figure 85: Kinetic plot for pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone. X-axis represents cyclohexanone concentrations in mM. Y-axis represents the velocity in µM.s⁻¹. Michaelis Menten curve fit is represented in black and Y-error bars and caps are displayed at each concentration/velocity point. a) MATOUAmDH2 WT:NADPH:cyclohexanone. b) L180A:NADPH:cyclohexanone. c) M215A:NADPH:cyclohexanone. d) T312A:NADPH:cyclohexanone.

Table 12 – Kinetic parameters for pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone

Variant	V_{\max} ($\mu\text{M}\cdot\text{s}^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{s}^{-1}\cdot\text{mM}^{-1}$)
pETYSBLIC-3C MATOUAmDH2	0.430 ± 0.0153	0.396 ± 0.0667	0.0820	0.210
L180A	0.221 ± 0.0106	0.613 ± 0.119	0.0330	0.0500
M215A	0.413 ± 0.00760	0.234 ± 0.0259	0.0630	0.270
T312A	0.250 ± 0.0108	2.31 ± 0.290	0.0380	0.0200

All mutant and wild-type enzymes were subjected to biotransformations using cyclohexanone **21** as a model substrate and using either ammonia or methylamine as amine donors during the reactions. Reactions were run at 25°C and time points were worked up every hour for 24 h. Time points were analysed via GC-FID using standards as references for peak analysis. Ratio percentages were calculated and plotted as conversions over time (Figure 86). For reactions concerning the formation of cyclohexylamine **21b** (ammonia-based reactions) the following conversions were observed after 24 h: WT: 99.5%, L180A: 80.7%, M215A: 99.9%, T312A: 93.9%. For reactions where methylamine was the amine donor to produce *N*-methyl cyclohexylamine **21c** the following conversions were seen after 24 h: WT: 8.33%, L180A: 41.0%, M215A: 6.83%, T312A: 14.3%. Peak standards and retention times are reported in *Appendices*: Figure 124 and 125.

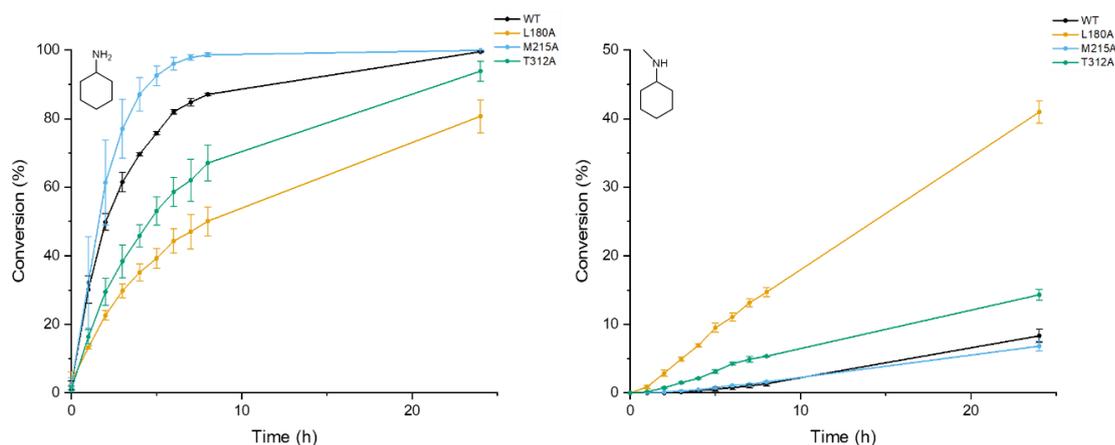


Figure 86: Biotransformations using pETYSBLIC-3C MATOUAmDH2, L180A, M215A and T312A with cyclohexanone. The x-axis represents time (in h) where each time point of the reaction was collected. The y-axis represents the percentage conversion to product (cyclohexylamine or *N*-methyl cyclohexylamine). The left-hand panel represents the reactions where the amine donor is ammonia. The right-hand panel represents the reactions where the amine donor is methylamine. WT points are plotted in black, L180A points in orange, M215A points in blue and T312A points in green.

As a result, a reduced library of site saturation mutagenesis was carried out at both the M215 and L180 positions to optimise and study the effect of these positions on the reactions driven by either ammonia or methylamine respectively. Five point mutations were designed at each site (plus L180A, M215A and WT) for reasons discussed in *Results and Discussion 4.3.1*, for the production of primary and secondary amines. Primers were all manually designed (*Methods 3.1.1.1*) for all codon substitutions in the pETYSBLIC-3C MATOUAmDH2 construct and are detailed in *Appendices Table 17*. PCR was carried out similarly to the alanine screen mutants, without the touch downs (*Methods 3.1.2.1*), which yielded good amplification, showing bands at approximately the expected size for all mutants (Figure 87).

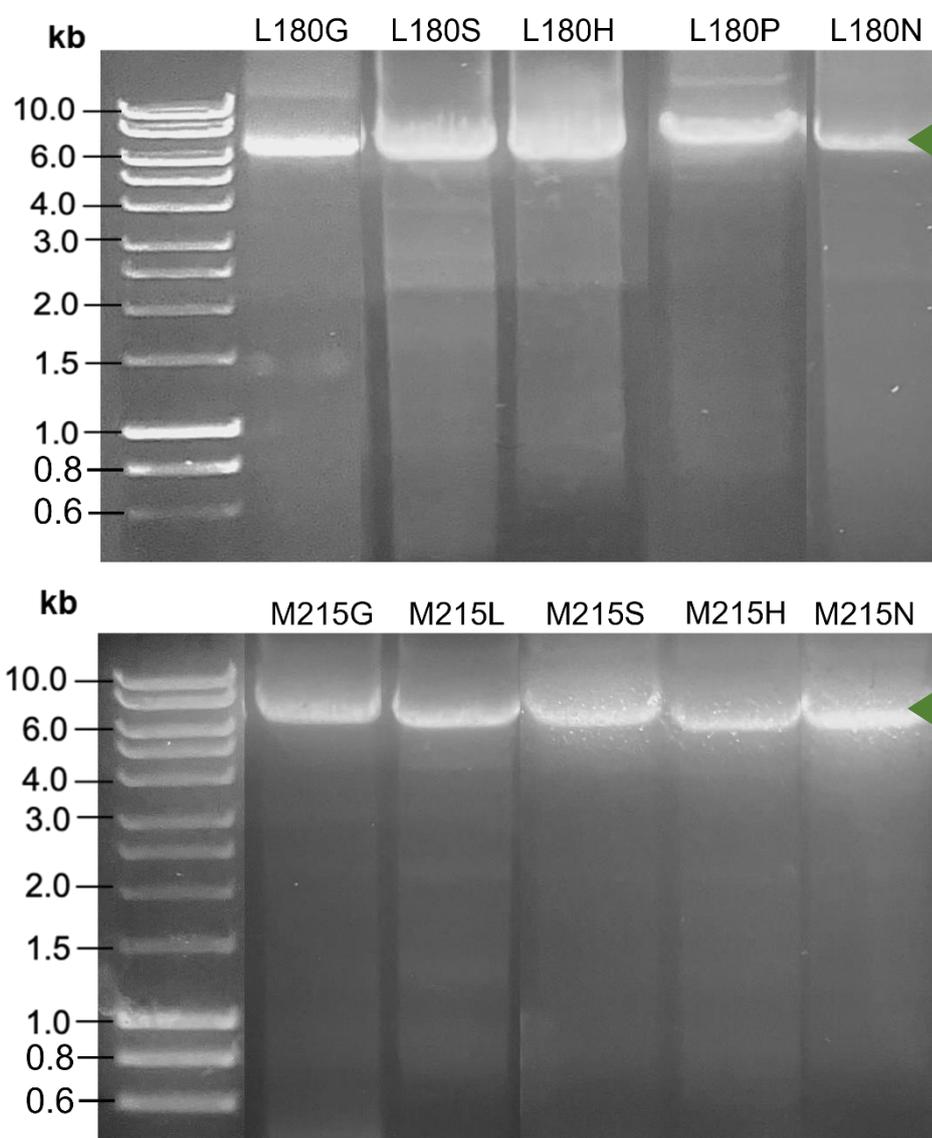


Figure 87: PCR product bands from MATOUAmDH2 site directed mutagenesis at L180 and M215 positions, analysed via gel electrophoresis. The left-hand lanes of each cropped gel represent a 1 kb HyperLadder™ marker and associated kb lengths. All lanes are labelled with specific mutant PCR products. This figure represents a collage of cropped images of bands that were excised from individual agarose gels. All samples were loaded onto a 1% agarose gel with SYBR® Safe stain and run for ~ 1 h at 110 V. The green triangles represent the bands of interest.

InFusion® cloning was carried out similarly to the protocol used for the alanine screen mutants; just one linearised vector with one codon mutation in the middle using the InFusion® premix to ligate and reform the recombinant plasmid. In the case of the mutants produced at both the L180 and M215 positions 100 ng of gel extracted, and purified DNA was used with appropriate amounts of InFusion® premix (*Methods 3.1.3*). Successful ligation proceeded; colonies were selected, recombinant plasmids isolated, and sequencing results confirmed the presence of codon substitutions at specific mutant sites.

Recombinant plasmids were then transformed into BL21(DE3) cells, colonies selected, and overnight cultures transferred into large scale media. These large-scale cultures were grown at 37°C for approximately 2-4 h until the OD₆₀₀ was roughly around 0.7 ± 0.05 for each mutant, and then all induced with 1 mM IPTG. Cell pellets were collected and harvested after expression (*Methods 3.1.6 and 3.2.1*). Cell pellets were weighed and appropriate amounts of 50 mM Tris HCl pH 7.1, 300 mM NaCl buffer was added prior and after lysis (*Methods 3.2.1*), so that the lysates contained approximately the same relative concentration of recombinant mutant enzyme across the board. Relative recombinant protein concentrations and presence was analysed via SDS-PAGE, which showed relatively similar amounts of each mutant protein present in the lysates collected (Figure 88 and 89).

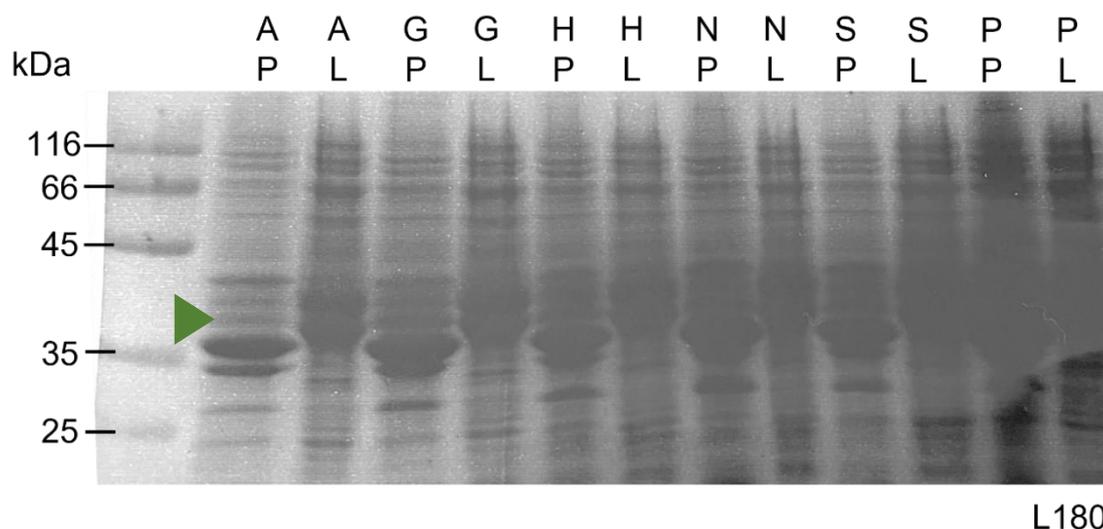


Figure 88: SDS-PAGE analysis of L180 mutant lysates. The left-hand lane represents a low range ladder (ThermoFisher) and associated molecular mass references in kDa. All samples were diluted 2.5 x in 50 mM Tris HCl pH 7.1, 300 mM NaCl buffer. Top letters refer to the amino acid substitution. P refers to a pellet sample post lysis (resuspended in 50 mM Tris HCl pH 7.1, 300 mM NaCl), L refers to a lysate sample i.e. A P stands for the pellet sample of L180A, and A L stands for the lysate sample of L180A. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

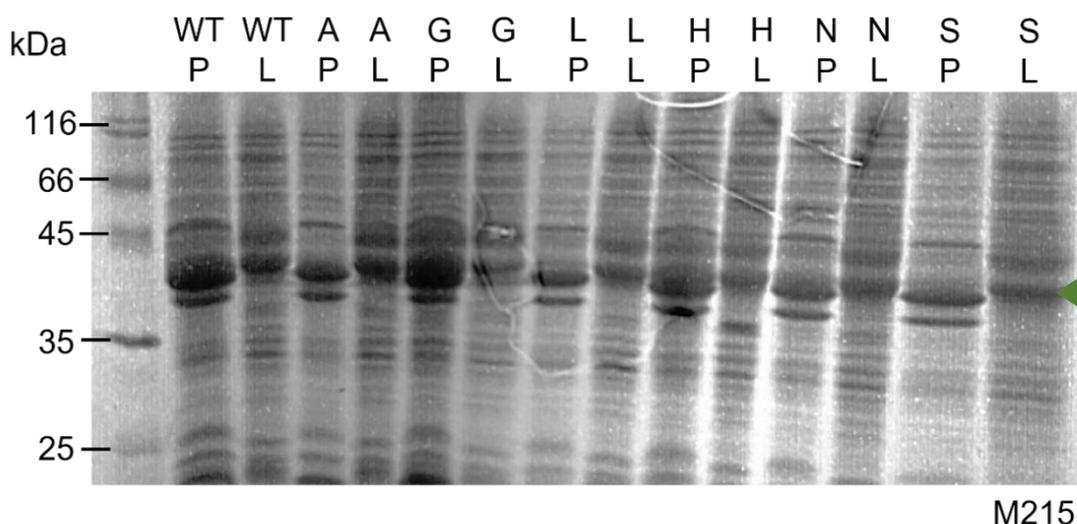


Figure 89: SDS-PAGE analysis of M215 mutant lysates. The left-hand lane represents a low range ladder (ThermoFisher) and associated molecular mass references in kDa. All samples were diluted 2.5 x in 50 mM Tris HCl pH 7.1, 300 mM NaCl buffer. Top letters refer to the amino acid substitution. P refers to a pellet sample post lysis (resuspended in 50 mM Tris HCl pH 7.1, 300 mM NaCl), L refers to a lysate sample i.e. A P stands for the pellet sample of M215A, and A L stands for the lysate sample of M215A. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

Lysates were then subjected to biotransformations using a range of substrates to assess the biocatalytic potential of each mutant to produce either primary or secondary amines for a larger substrate scope. Reactions were set up as described in *Methods* 3.5.2, using 20 mg.mL⁻¹ of lysate, 10 mM of various aliphatic and aromatic ketones with in either 2 M Ammonium Formate buffer or 25 equivalents of methylamine for the production of primary and secondary amines respectively. Reactions were run at 25°C for 24 h where time points were taken every 2 h between 0-8 h and then a 24 h time point collected. Similar results as to previous biotransformations using WT, L180A, M215A and T312A to produce either cyclohexylamine **21b** or *N*-methyl cyclohexylamine **21c** were obtained for this library of mutants (Figure 90). In the case of cyclohexanone ammonia-based reactions M215S performed the best with yielding a conversion of 91.5% after 24 h. In the case of cyclohexanone methylamine-based reactions L180A and L180S performed the best yielding 29.6% and 27.9% conversion respectively, after 24 h. Reactions were carried out using larger aromatic compounds such as 1,2-cyclohexanedione **31** and norcamphor **29a** which have previously been shown to be compatible substrates for MATOUAmDH2.¹⁰³ In the case of norcamphor **29a** ammonia-based reactions M215A and M215S performed the best out of all enzymes producing a 46.2% and 45.9% conversion respectively after 24 h (Figure 90). Unfortunately, no reaction was seen with norcamphor **29a**, for any of the mutants or wild-type enzyme, when using methylamine as the amine donor for this reaction. For 1,2-cyclohexanedione **31** the wild-type enzyme performed the best giving a 92.0% conversion after 24 h with

M215A, M215L and M215S also giving excellent conversions of 89.8%, 89.4% and 87.9% respectively (Figure 90). Larger aromatic compounds such as 2-tetralone **15** were also used in preliminary ammonia-based reactions but unfortunately no reaction was observed. Smaller aliphatic compounds such as 2-pentanone **20**, which has been shown to be catalysed by MATOUAmDH2¹⁰³, proved challenging to analyse via GC-FID as for reasons discussed in *Results and Discussion 4.3.1*. Additionally larger aliphatic compounds such as 2-heptanone **32** were also tried with the wild-type and mutant enzymes but no measurable reaction was observed with ammonia. Figure 90 outlines the findings from all the small-scale preliminary, and full reactions, tried with all mutants for both ammonia and methylamine-based reactions with a range of substrates. Interestingly for all reactions histidine substitutions yielded either no conversion for ammonia-based reactions at position M215 (M215H) whereas L180H also produced little to no conversion in methylamine-based reactions (Figure 90 and 91). Peak standards and retention times are reported in *Appendices Figures 124-130*.

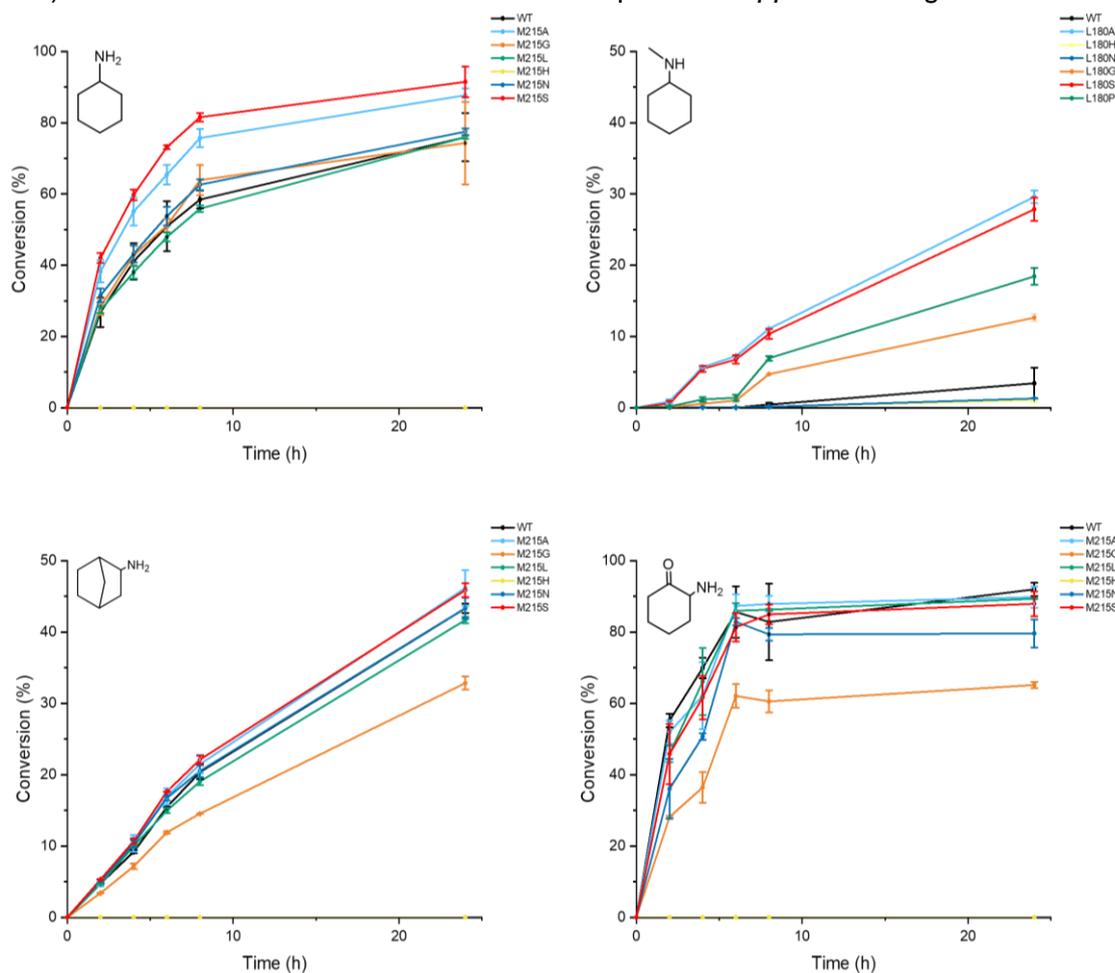


Figure 90: Biotransformations using pETYSBLIC-3C MATOUAmDH2 L180 and M215 mutants. The x-axis represents time (in h) where each time point of the reaction was collected. The y-axis represents the percentage conversion to product. The top left-hand panel represents the reactions with cyclohexanone and ammonia. The top right-hand panel represents the reactions with cyclohexanone and methylamine. The bottom left-hand panel represents the reactions with norcamphor and ammonia. The bottom right-hand panel represents the reactions with 1,2-cyclohexanedione and ammonia. Mutants are labelled by colour in individual legends.

		WT	M215A	M215G	M215L	M215H	M215N	M215S	WT	L180A	L180G	L180H	L180N	L180S	L180P
Conversion (%) t = 24 h		75.9 ± 6.77	87.3 ± 1.95	74.3 ± 11.6	75.9 ± 0.434	0	77.5 ± 0.912	91.5 ± 4.29	3.42 ± 2.21	29.6 ± 0.872	12.7 ± 0.517	1.00 ± 0.160	1.29 ± 0.0362	27.9 ± 1.63	18.4 ± 1.20
		92.0 ± 1.88	89.8 ± 3.00	65.2 ± 0.890	89.4 ± 0.0551	0	79.6 ± 3.94	87.9 ± 3.50	-	-	-	-	-	-	-
		43.3 ± 0.641	46.2 ± 2.50	32.9 ± 0.930	41.7 ± 0.454	0	43.4 ± 1.48	45.9 ± 0.968	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-	-	-
		0	0	0	0	0	0	0	-	-	-	-	-	-	-

Figure 91: Conversions of a range of ketones using pETYSBLIC-3C MATOUAmDH2 L180 and M215 mutants. All reactions were run with NADPH as cofactor and 20 mg.mL⁻¹ cell lysate. Reactions were analysed via GC-FID. Conversions and standard deviation errors from an average of two repeats are displayed to three significant figures in percentages (%). Different ketone substrates are displayed in the left-hand column and mutants for either ammonia-based or methylamine-based reactions are highlighted in a header row. 0% conversions represent reactions where there was no observable product peak formed. – denotes reactions that either were not carried out or were not able to be analysed via

Across the board M215S showed promising biocatalytic abilities towards not only cyclohexanone **21** but more sterically hindered, larger, aromatic compounds such as norcamphor **29a**. Therefore, attempts were made to crystallise this mutant to better understand the structural properties that allude to this improved activity. This mutant was transformed and expressed in BI21(DE3) (*Methods 3.1.5 and 3.1.6*) as standard using the recombinant plasmid generated in the pETYSBLIC-3C vector. M215S was then purified via NiNTA in tandem with SEC to yield 65 mg per 2 L of LB culture. Both chromatograms yielded one nice A280 peak indicating relatively pure protein (Figure 92 and 93) in which overexpression was confirmed by SDS-PAGE analysis (Figure 94). The band runs slightly lower than the expected weight due to shorter electrophoresis run time; the ladder has not separated as well as expected (Figure 94).

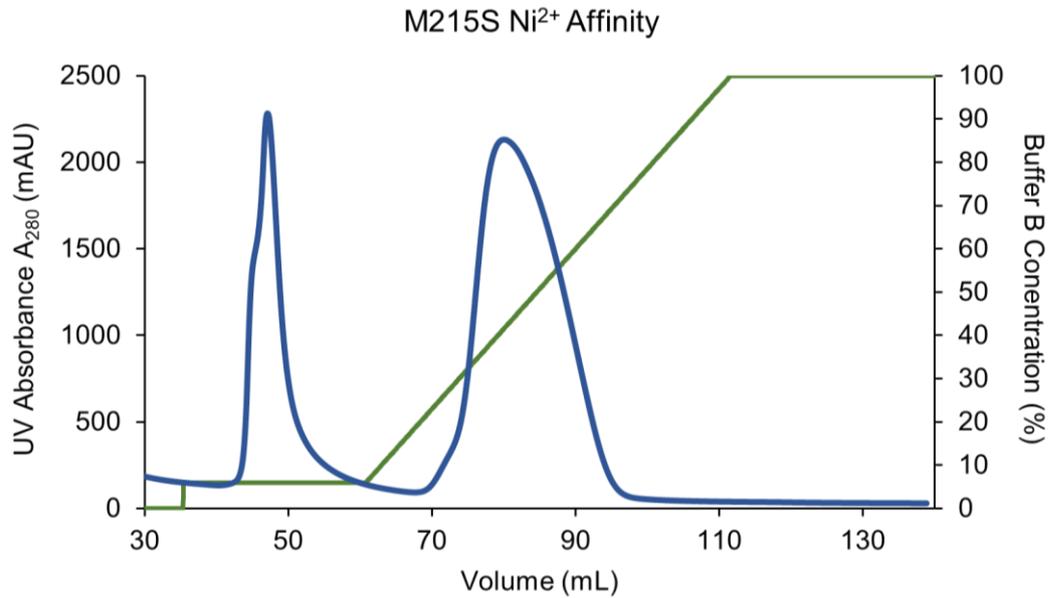


Figure 92: Nickel Affinity Chromatogram for M215S MATOUAmDH2 expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line.

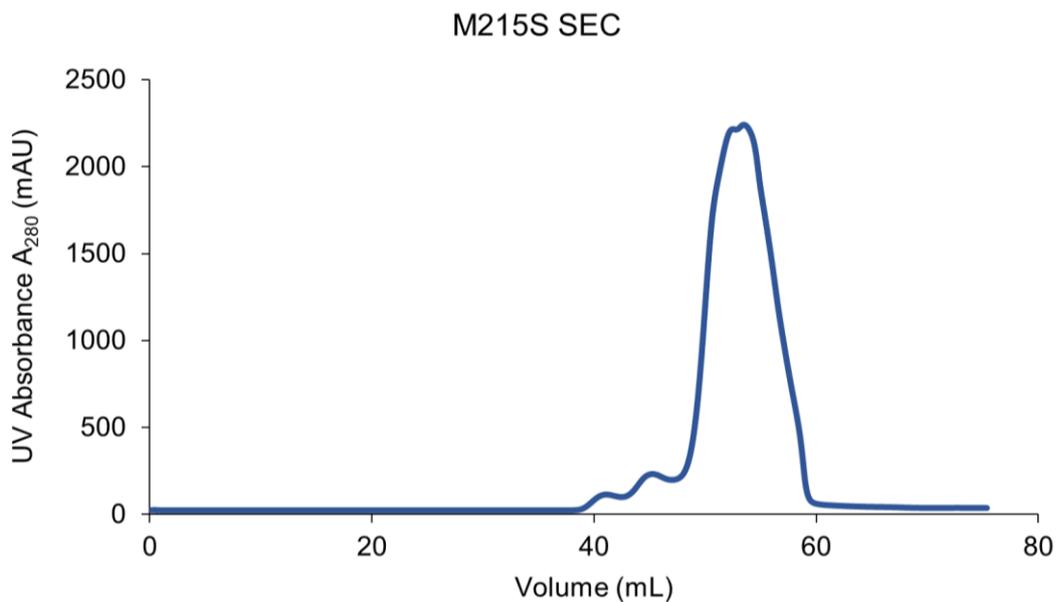


Figure 93: Size Exclusion Chromatogram for M215S MATOUAmDH2 expressed in the pETYSBLIC-3C vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue).

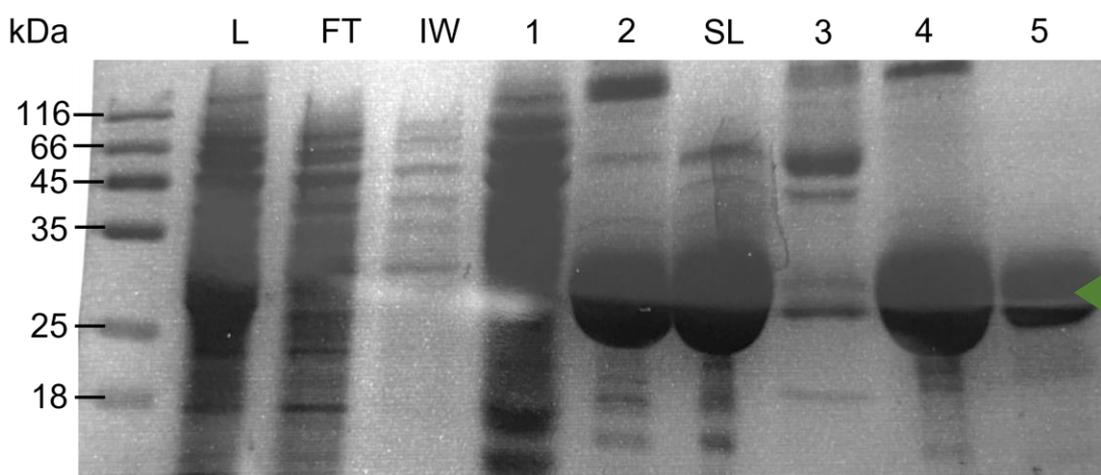


Figure 94: SDS-PAGE analysis of protein fractions from pETYSBLIC-3C M215S MATOUAmDH2 protein purification. The left-hand lane represents a low range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane L*: 5 x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lane FT*: 5 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane IW*: 5 x diluted sample of the initial wash, before the imidazole gradient, on the ÄKTA using 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lanes 1&2*: 2 x diluted combined, pre-determined, non-specific and specific samples from the ÄKTA fraction collector after Nickel Affinity Chromatography. *Lane SL*: 5 x diluted pooled and concentrated protein loaded onto the Superdex® column. *Lanes 3&4*: 2 x diluted samples from non-specific and specific fractions during Size Exclusion Chromatography. *Lane 5*: 10 x diluted pooled and concentrated pure M215S. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

M215S was subjected to a round of initial pre-dispensed crystallisation screens, in a sitting drop manor using a 1:1 ratio (*Methods 3.3.1.1*). Initially M215S was concentrated to 20 mg.mL⁻¹ and incubated with 4 mM NADP⁺ (2 mM final in drop) and 15 mM 2-aminonorbornane hydrochloride **29** (in H₂O) or quenched and extracted 2-aminonorbornane **29** in ethyl acetate. Ligand dissolved in H₂O formed much nicer cuboid like crystals, than that dissolved in ethyl acetate. A dataset was resolved for this structure but unfortunately no NADP⁺ was bound in the active site. Therefore the NADP⁺ concentration was increased up to 10 mM (5 mM final in crystal drops) and incubated with 20 mg.mL⁻¹ M215S and 15 mM 2-aminonorbornane **29** in H₂O. Initial Index screens, 48-well and 96-well optimisations all produced fishable crystals, resembling cuboid like morphologies (Figure 95). Crystals were fished and sent off to the Diamond synchrotron and a data set was resolved to 1.61 Å. These crystals were fished from a drop containing 0.3 M MgCl₂, 0.1 M Tris-HCl pH 8.5 and 25% PEG 3350 with no additional cryoprotectant added. This dataset occupied a *P12₁1* space group with 4 molecules in the asymmetric unit. The final solution was resolved to 1.61 Å and refined to R values of 20.4% and 23.0% for *R*_{crystal} and *R*_{free} respectively. The data and refinement statistics for this dataset are reported in Table 13.

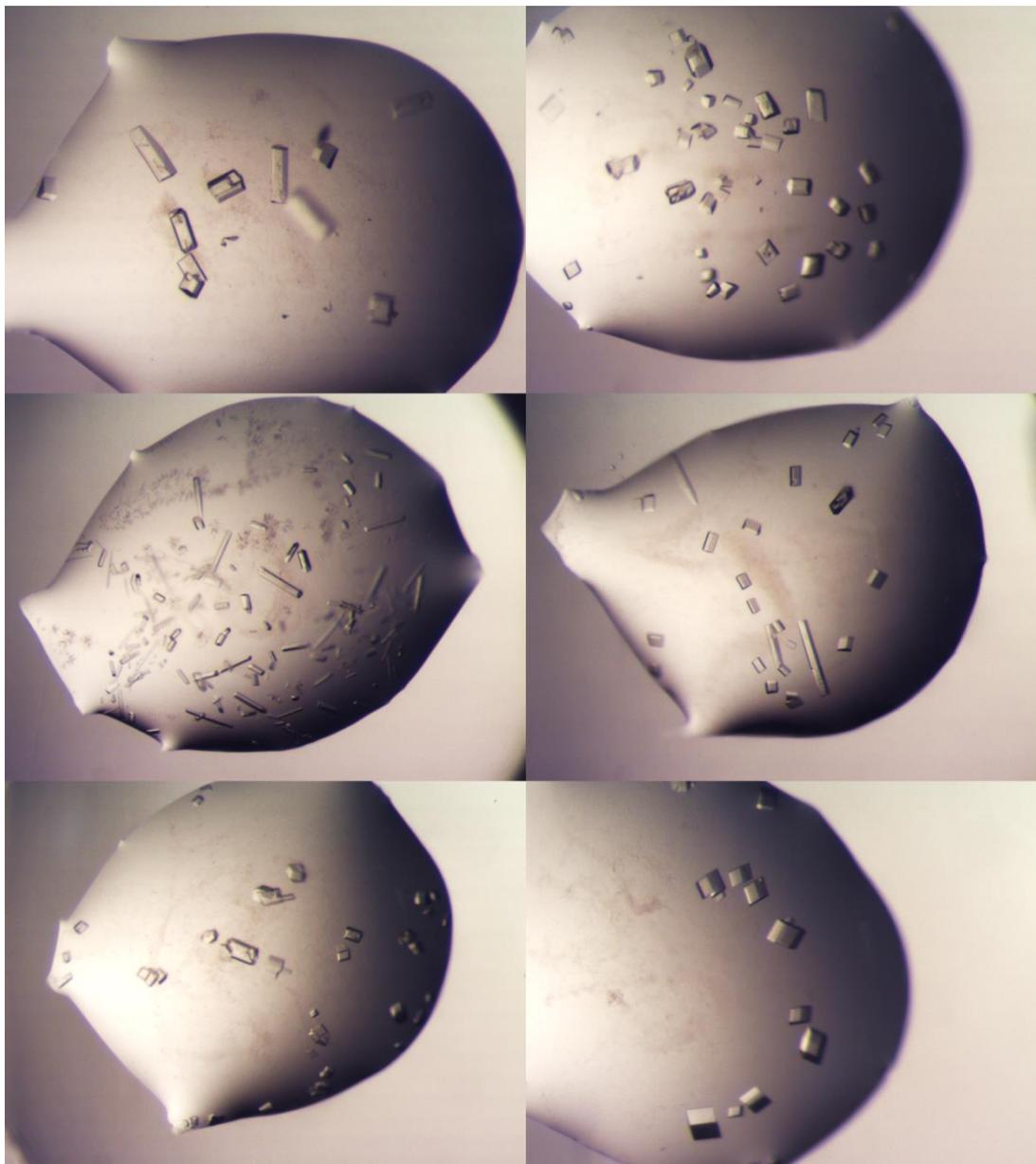


Figure 95: Crystal forms from M215S MATOUAmDH2:pETYSBLIC-3C in complex with NADP⁺ and 2-aminonornornane. Microscopy images of cuboid plates and rods from a pre-dispensed Index (MRC-96 well) screen and 96-well optimisations.

Table 13 – Data collection and refinement statistics for the 1.61 Å MATOUAmDH2 M215S in complex with NADP⁺ and 2-aminonorbornane.

1.61 Å M215S MATOUAmDH2:NADP⁺:2-aminonorbornane

Beamline	Diamond I03
Wavelength (Å)	0.8235
Resolution (Å)	46.06-1.61 (1.64-1.61)
Space Group	<i>P</i> 12 ₁ 1
Unit cell (Å)	a = 75.45; b = 107.91; c = 89.71 α = 90° β = 99.66° γ = 90°
No. of molecules in the asymmetric unit	4
Unique reflections	182654 (1136)
Completeness (%)	100.0 (100.0)
R _{merge} (%)	0.025 (0.57)
R _{p.i.m.}	0.025 (0.57)
Multiplicity	2.0 (1.8)
<I/σ(I)>	11.1 (1.4)
Overall B from Wilson plot (Å ²)	48
CC _{1/2}	1.61 (0.61)
R _{cryst} / R _{free} (%)	20.4/23.0
r.m.s.d 1-2 bonds (Å)	0.00088
r.m.s.d 1-3 angles (°)	1.474
Avgc main chain B (Å ²)	20
Avgc side chain B (Å ²)	23
Avgc water B (Å ²)	22
Avgc NADP ⁺ B (Å ²)	35
Avgc 2-AMN B (Å ²)	42

Brackets refer to data in the highest resolution shell.

As expected, the overall structure of this mutant is very similar to the wild-type enzyme, displaying a large N-terminal Rossman domain and a C-terminal beta pleated sheet comprised of 8 beta strands. The biological assembly consists of two dimers chains A and B and chains C and D where each monomer associates at a series of beta pleated sheets (Figure 96). All chains are in the 'closed' conformation like the cyclohexylamine bound structure of the wild-type enzyme (PDB: 7R09) (*Results and Discussion 4.2*); the average SSM RMSD value for each chain to the one chain in the ligand bound wild-type structure varied from 0.75-0.84 Å for approximately 341 residues. The diameter of the binding pocket is similar to the wild-type cyclohexylamine bound structure, giving measurements of approximately 6.5 Å at the narrowest point and 12.0 Å at the widest. In each binding pocket there was clear $F_o - F_c$ electron density which could be modelled as 2-aminonorbornane, specifically the (*S*)-*endo* epimer, in all chains. Surprisingly even when increasing the NADP⁺ concentration during co-crystallisation only this one dataset showed some cofactor density. Albeit this cofactor electron density was incomplete; only appearing in chains B and D and the nicotinamide nucleotide portion did not show up at all in the map. As much of the cofactor as possible was built into chains B and D, Figure 97 highlights the omit maps for both 2-aminonorbornane **29** and NADP⁺. Tris molecules were also built into the structure as well as some chloride ions in the solvent density.



Figure 96: Dimer of M215S bound to 2-aminonorbornane and NADP⁺. All images have been visualised using CCP4mg (v.2.10.11). Biological assembly of MATOUAmDH2 M215S mutant, chains C and D associated at a series of beta pleated sheets (blue). Part of the NADP⁺ cofactor is in grey (chain D) and 2 molecules of 2-aminonorbornane (yellow), one bound to each monomer.

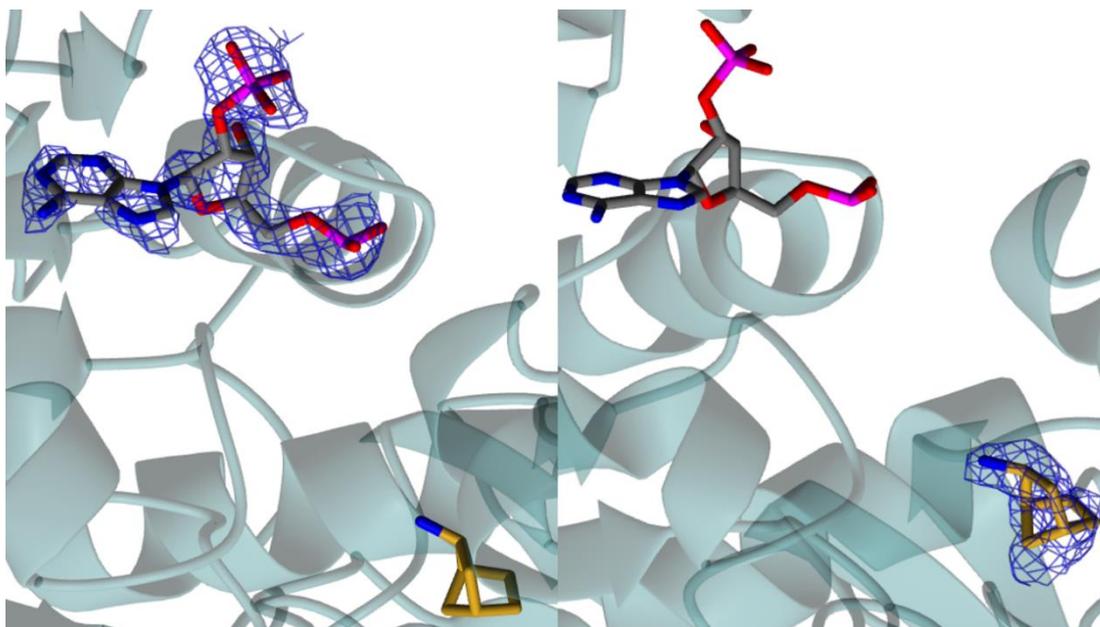


Figure 97: Omit maps for both NADP⁺ and 2-aminonorbornane bound to M215S. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel displays the omit map for NADP⁺ (grey) displayed in cylindrical wire form in blue. The right-hand panel displays the omit map for 2-aminonorbornane (yellow) where the F_0-F_c electron density is displayed in cylindrical blue wire. The map contour level is set to 3 σ .

4.3.1 Discussion

Initial alanine mutants were designed to test the functionality of residues that lined both the floor and roof of the active sites. The residues that were targeted were either bulkier residues conferring some possible steric role or ‘un-typical’ hydrophobic residues which added additional hydrophobicity to the wild-type active site. Most residues that were targeted for mutation were found in the floor or ceiling of the active site and so probably play a role in how dynamic and caged the active site is on ligand binding. Previously, it was shown that in MATOUAmDH2 A147, which replaces typically bulkier residues from other nat-AmDHs such as F144 from *Cfus*AmDH, lead to an increased activity towards larger substrates.¹⁰³ Any changes in hydrophobicity or size of the available space in the binding pocket may have a large effect on substrate scope and specificity of MATOUAmDH2.

Sub-cloning was not as straightforward as for previous experiments (*Results and Discussion 4.2*), various optimisations were carried out but ultimately removing the touch down steps yielded single pure bands at the correct molecular weight for all six mutants. Bands were excised and purified, and the correct alanine substitution codons were confirmed for all six mutants using Sanger sequencing. Purification for L180A, M215A and T312A all gave high yields and purities with one prominent single

peak and band on the chromatograms and SDS-PAGE gels respectively. However, for F143A and L144A aggregation and protein precipitation was first observed when fractions were collected during the NiNTA chromatography run. As a result, the aggregation was spun down and the resulting supernatant protein was concentrated through a MWCO filter and loaded onto a Superdex® column. However, protein aggregation for both F143A and L144A was observed again in the fraction collector showing that the aggregation was indeed from these recombinant mutant proteins and not from any *E. coli* expressed background, co-eluted protein. This is also reflected in the overall loss of yield between the NiNTA purification and SEC purification chromatograms. SDS-PAGE analysis of these mutants also revealed some proteolysis present. Interestingly this is also similar for L169A however no clear aggregation was seen during purification. However, when L169A was added to the cuvette during kinetic runs aggregation was immediately spotted. Evidently this may be due to the relatively high concentration in the cuvette and relatively low concentration of protein during purification (as evident from the low measurable mAU) which is why aggregation for L169A initially went unnoticed. Overall, these three mutants were not able to be used for further characterisation but reflect how these mutants may have a larger role in the stability of the protein. This is not unexpected as highlighted in *Results and Discussion 4.2.1*, these residues either confer some added steric or hydrophobic effect that is a more novel and specific property of MATOUAmDH2 when compared to other nat-AmDHs.

The resulting three mutants L180A, M215A and T312A were all used for further characterisation. Kinetic parameters determined from scanning UV-vis spectrometry revealed that M215A showed an improved k_{cat}/K_m value when compared to the wild-type enzyme. This is largely due to its much-improved K_m value (relative to WT enzyme) showing that M215A may perform faster at the very start of the reaction with stronger protein-substrate interactions at the lower range of concentrations (<1mM). However, these mutants may not have such a tight binding affinity to cyclohexanone when compared to the wild-type enzyme as the protein concentration needed to be increased 10 x fold to see any observable activity with any of the mutants. It should also be noted that the wild-type enzyme k_{cat}/K_m value for this pETYSBLIC-3C expressed protein is also very similar to the pET22b(+) expressed protein which is to be expected but confirms no change in activity on sub-cloning. No parameter determination could be carried out using methylamine as the amine donor as the reaction was deemed to slow to be observed using UV-vis spectrometry.

L180A, M215A and T312A were all subsequently used in biotransformation reactions for ammonia-based and methylamine-based reactions using cyclohexanone as a model substrate. Reactions were monitored over a 24 h time period at 25°C and analysed using GC-FID using standards as retention time references. All standard peaks are reported in *Appendices: Figure 122-130 and Table 19*. M215A showed slightly improved activity towards cyclohexanone when using ammonia as the amine donor. L180A showed a near 5-fold increase in activity towards cyclohexanone when methylamine was the amine donor. M215A most likely results in a mutant in which more space is afforded in the floor of the active site and so the substrate cage is more tightly stabilising the substrate/product. L180A shows a drastic increase in activity, the likely cause being the extra space afforded in the ceiling of the active site allowing extension in the *N*-methyl group of *N*-methyl cyclohexylamine product/intermediate to be stabilised in the active site. For this reason, ten other mutants were designed at positions M215 and L180 to probe the enzyme for improved biocatalysis with a range of substrates using either ammonia or methylamine respectively.

NDT degeneracy was used to bias codons with equal probabilities of insertions. Acid based residues were avoided as there was potential to interfere with the amine-based binding. Amino acid substitutions were largely designed to confer hydrophobicity, reduction in size and rigidity. All proteins were subcloned as before and lysates were produced as per *Methods 3.2.1/Results and Discussion 4.3*. All cell pellets were resuspended in and lysed in the appropriate amount of buffer and lysates were confirmed to have approximately similar amounts of protein present. Lysates were used for biotransformations using a range of substrates. Across all reactions approximately a 10% reduction in overall activity was seen when comparing to purified protein; which is expected due to the presence of other *E. coli* background protein i.e. a reduction in specific activity. Reactions performed well with the model substrate cyclohexanone **21** for both ammonia-based and methylamine-based reactions. M215A, M215N and M215S outperformed the wild-type enzyme when producing cyclohexylamine; most likely due to the decrease in side chain size conferring more tight binding, with M215S performing the best out of all mutants. M215G in all reactions seemed to confer little functionality at this position and does not improve the overall reaction conversions regardless of reduced relative size. For reactions where methylamine was the amine donor L180A performed the best with most mutants outperforming the wild-type enzyme except for L180N and L180H. L180N reflects on the importance of hydrophobicity at this position for high specific activities and turnovers. Interestingly in all reactions involving mutations to histidine either resulted

in very little activity or complete inactivation of the enzyme following aggregation, showing that charge stabilisation may play a bigger role than anticipated during the reaction mechanism at these positions.

With MATOUAmDH2's promising active site size; affording more space and being more 'open' than previously characterised nat-AmDHs (*Results and Discussion 4.1 & 4.2*) a range of larger substrates were also used to test for activity. Norcamphor **29a**, the analogue of the natural product camphor **29b**, is routinely used as a building block in medicinal and material chemistry. Due to norcamphor's chemical significance it was chosen to be an interesting target to produce the resulting amine products with both potential biological and chemical significance. Additionally, norcamphor has been shown to be active with the wild-type MATOUAmDH2¹⁰³; and so, probing and improving this reaction using mutagenesis proved to be attractive. As was seen in the cyclohexanone-based reactions, mutants with smaller residues at position 215 proved to be the best enzymes with the highest conversions after 24 h. M215A and M215S performing the best relative to the wild-type activity to produce 2-aminonorbornane **29**. Unfortunately, no reaction was observed for any of the mutants or wild-type enzyme when using methylamine as the amine donor.

Additional larger substrates were also screened including 1,2-cyclohexanedione **31**, which has been previously shown to have a high specific activity for the wild-type enzyme.¹⁰³ Overall good conversions were monitored using GC-FID with the wild-type enzyme having the best conversion for this substrate. M215S and M215A still performed relatively well with this substrate (and M215L). This may be since 1,2-cyclohexanedione may be more closely related to the enzyme's natural product than either cyclohexanone **21** or norcamphor **29a** and so the current active site is highly evolved towards this di substituted carbonyl group. The natural product is unknown with the enzyme having been mined from metagenomic DNA. There is no conclusive reason as to why the wild-type, with the additional functionality and size that the methionine confers, performs the best. Moreover, this reaction proved slightly trickier to analyse using GC-FID than both norcamphor and cyclohexanone-based reactions. 1,2-cyclohexanedione appeared to degrade during the reaction time course or at least concentration dependence decreased without any accumulation of product. In the GC-FID traces it is evident that as time goes on the peak shifts further to the left and a generation of a 'sub-peak' appears with no generation or indication of any different product peaks such as alcohol formation. This is indicative of a degradation of 1,2-cyclohexanedione to a compound with a similar GC profile (Figure 98). This was also observed in the M215H mutant where no product was produced (<1% conversion

after 24 hours or undetectable) in the reaction at all, but the substrate peak was shifted again, and a second sub peak was observed (Figure 98). The substrate peak when processing was taken as the second right half, where the initial peak started, and this was used to calculate relative conversions. Taking the whole peak into consideration felt like an underestimation of true product conversion when some 'loss of conversion' i.e., 'total substrate' peak would be attributed to the degradation product. Figure 99 shows the plot for this reaction if the 'whole' substrate peak was used in calculating relative conversions. Furthermore, this reaction would need to be optimised, to prevent degradation of the substrate, to see any accurate and meaningful results if it was to be attempted using methylamine as the amine donor.

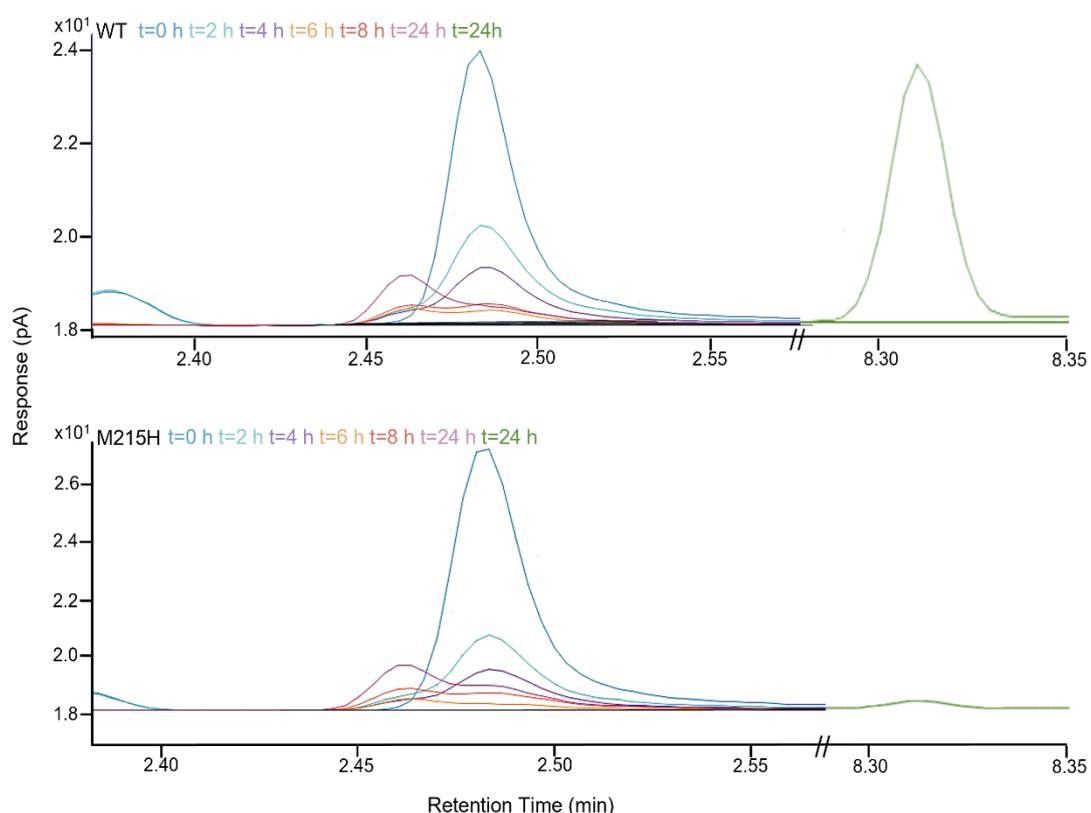


Figure 98: GC-FID traces of substrate and product peaks for the 1,2-cyclohexanedione:ammonia based reactions monitored over 24 h. The x-axis represents retention times in min an x-axis split is denoted by //. The y-axis represents the response current in pA. The top panel displays the reaction carried out with 1,2-cyclohexanedione, ammonia and wild-type lysate. The bottom panel represents the reaction carried out with 1,2-cyclohexanedione, ammonia and M215H lysate. Time points are labelled accordingly based off a colour scheme with the time point in green at 24 h representing the product peak.

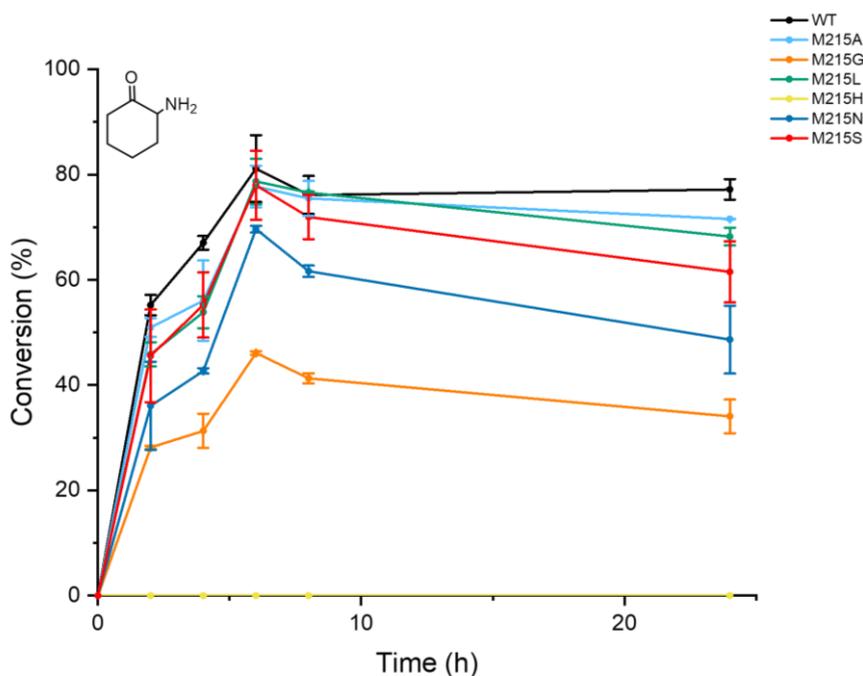


Figure 99: Underestimated biotransformations calculated from the combined peak of 1,2-cyclohexanedione in reactions using ammonia and M215 mutants. The x-axis represents time (in h) where each time point of the reaction was collected. The y-axis represents the percentage conversion to product. The conversion over-time scheme was calculated using the total peak, including the degradation product of 1,2-cyclohexanedione as a reference during processing. All mutants are labelled by colour.

Even larger, untested aliphatic and aromatic substrates were screened including 2-tetralone **15** and 2-heptanone **32**, each of which are intermediates routinely used in organic synthesis and often can be derived from naturally occurring precursors such as naphthalene with 2-heptanone occurring as a natural product. Unfortunately, both compounds did not produce any observable reaction with ammonia and M215 mutants, possibly due to being too large to occupy the binding pocket even with the afforded extra space by mutants particularly M215A and M215S. Methylamine based reactions were not attempted for this reason also. Reactions were scaled back to smaller aliphatic compounds to assess biocatalytic activity with both sets of mutants. 2-pentanone **20** was used in ammonia-based reactions but unfortunately no observable product peak was seen due to the possible volatile nature; the peak appeared masked by the solvent peak. If this reaction was to be scaled up with ammonia and methylamine the GC-FID ramp temperatures would need to be optimised. Current optimisations of these oven temperatures were not successful in producing readable and reliable peaks for either product or substrate.

M215S was an interesting mutant resulting from this screening showing high activity for the model substrate cyclohexanone and larger substrates such as 1,2-cyclohexanedione **31** and norcamphor **29a**. Hence, M215S was subjected to crystallisation trials to better understand the impact of this substitution on the dynamic space and nature of the active site. Obtaining crystals for this mutant was relatively straightforward where better crystal morphologies were observed when the ligand, 2-aminonorbornane **29** (HCl), was dissolved in H₂O as opposed to ethyl acetate (quenched and extracted). Initially co-crystallisation was set up using NADP⁺ concentrations of 4 mM (2 mM drop concentration) which has been successful for resolving other NADP⁺ bound crystals (*Results and Discussion 4.1, 4.2 and 4.4*). Data obtained from these initial trials were of good quality and resolution with a resolvable and refined solution; however, no cofactor was bound in any observed chains, but clear ligand density was seen in all cases. The NADP⁺ concentration was increased to 10 mM during incubation and two datasets were obtained showing promising diffraction data with resolvable solutions. Again, however no cofactor electron density was observed in all chains albeit one dataset where chains B and D showed sparse cofactor density. This density was largely limited to the adenine-based nucleotide and the nicotinamide nucleotide density was not present. This is the extension of the cofactor that reaches down into the binding pocket and stacks usually opposite with the ligand to close the cage. It is likely that the additional steric effect of 2-aminonorbornane pushes the cofactor into a more flexible orientation or it becomes more dynamic to allow for a larger substrate to bind rendering it harder to crystallise in a fixed orientation. Another accumulative effect could be due to extra saturation of water in the active site from the ligand; this is more noticeable in chains A and C where no cofactor density whatsoever was observed. Additionally in chains A and C with no NADP⁺ density K43 seems to orientate into the space that would be otherwise be occupied by NADP⁺ whereas in chains B and D the lysine side chain orientates away from the site as in the wild-type ligand bound structures (with NADP⁺) (Figure 100). The flexibility of this K43 side chain might be a reason for the lack of consistent NADP⁺ binding between chains. However, this may just be the result of no cofactor being bound in these chains allowing for additional flexibility around this region. It is also hard to conclusively say as the electron density in this area is also relatively sparse.

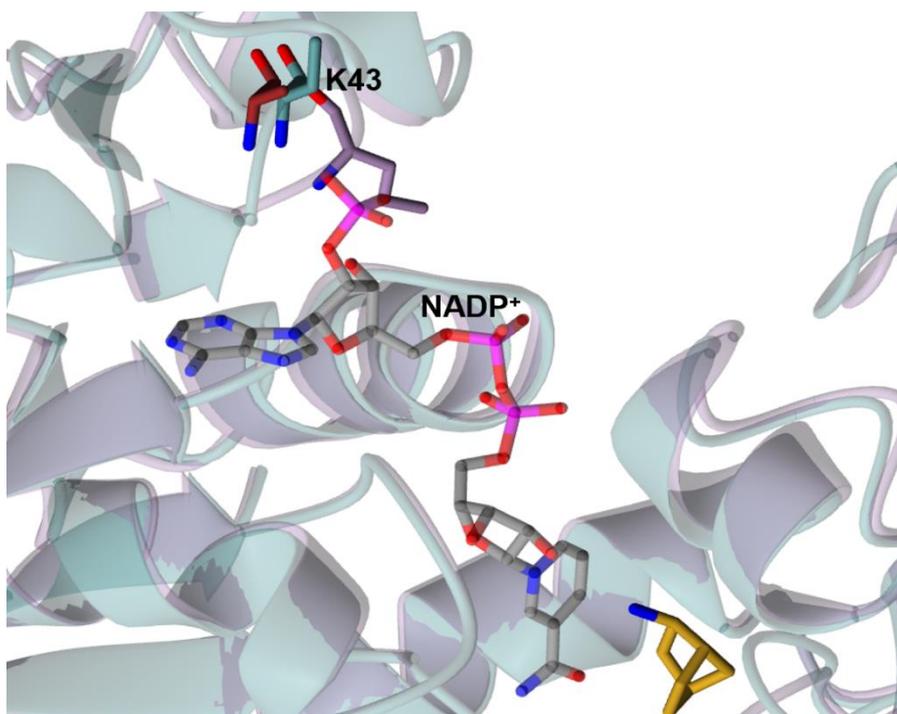


Figure 100: K43 orientations between chains of M215S and WT MATOUAmDH2. All images are visualised in CCP4mg (v.2.10.11). K43 residues from chain A (lilac) SSM superimposed with chain B (blue) and wild-type MATOUAmDH2 (PDB:7R09) (red). NADP+ (grey) has been modelled from the wild-type MATOUAmDH2 structure. 2-aminonorbornane is represented in yellow from chain B.

As with the cyclohexylamine **21b** bound wild-type structure (*Results and Discussion* 4.2) (PDB: 7R09), M215S displays all the typical binding site residues in correct orientations to enclose 2-aminonorbornane **29** in the binding pocket (Figure 101). In each chain 2-aminonorbornane was found in a single (*S*)-enantiomer configuration of the *endo* epimer, highlighting the potential selectivity during the reaction. This would suggest that during the reductive amination the hydride from NADPH may attack at the *exo* face of the resulting iminium ion and the hydride would insert at the *exo* face to give the *endo* amination product (Figure 102). The *endo* product of 2-aminonorbornane is also an ideal distance from E111 at 2.46 Å; ideal proximity for the stabilised ammonia to attack at the carbonyl centre of norcamphor **29a** (Figure 101). Unless there was a rotation of the ligand within the active site it is likely that the *exo* product would increase the distance for E111. It is noticeable when looking at the binding pocket of M215S that the substitution to serine allows for more space in the floor of the active site. Interestingly however, you can see that the side chain of L144 slightly shifts in and occupies some of this free space (Figure 101). This may explain why M215S works extremely well for smaller substrates like cyclohexylamine but only just outperforms the wild-type enzyme with larger substrates. The small amount of

extra space is enough to stabilise smaller substrates but the shift of L144 may reduce the potential maximised amount of space for larger substrates like norcamphor.

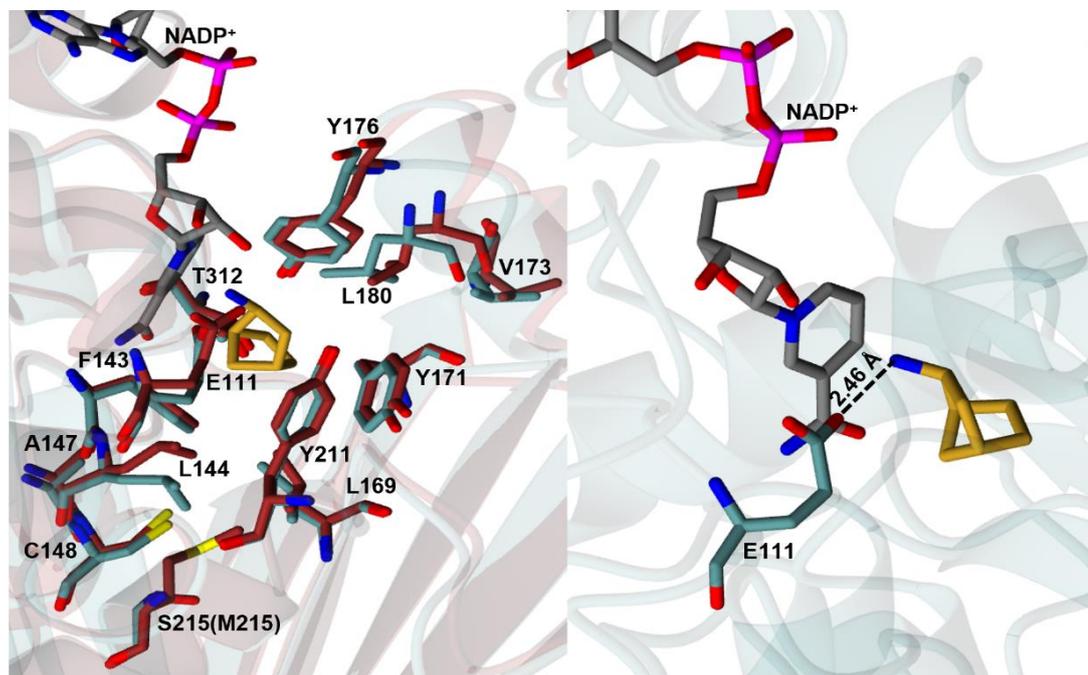


Figure 101: Binding pocket of M215S MATOUAmDH2 in complex with NADP⁺ and 2-aminonorbornane. All images are visualised in CCP4mg (v.2.10.11). The left-hand panel represents M215S chain D (blue) SSM superimposed with the wild-type MATOUAmDH2 enzyme (red) (PDB:7R09) and 2-aminonorbornane (yellow). Of particular importance is the shift of L144 into the space afforded by the M215S mutation. Amino acid residues are labelled accordingly. The right-hand panel displays E111 (blue) from chain D of M215S and the annotated distance to 2-aminonorbornane (yellow). The measured distance is displayed in Ångstroms and highlighted by a dashed black line. NADP⁺ in all cases has been modelled from the wild-type MATOUAmDH2 and is displayed in grey.

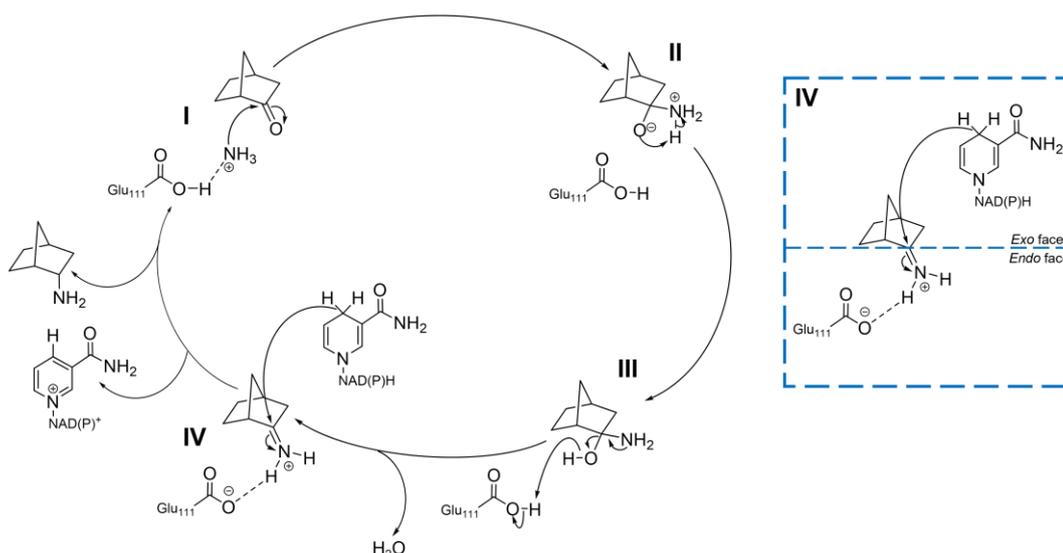


Figure 102: Proposed reductive amination mechanism of norcamphor in the active site of MATOUAmDH2. Stepwise theoretical mechanism of the reductive amination of norcamphor to 2-aminonorbornane. I) Attack of an amine donor at the carbonyl centre, in this case ammonia stabilised by E111. II and III) Formation of the iminium ion through loss of water. IV) Attack of the imine or iminium ion by NAD(P)H with hydride transfer occurring at the *exo* face of the intermediate. This results in the formation of the *endo* amine product, generation of NAD(P)⁺ and regeneration of protonated E111.

4.4 Other nat-AmDHs of interest

4.4.1 *CfusAmDH* W145A

Recently published work found that the W145A mutant of *CfusAmDH* (*Introduction 1.3.1*) worked on a range of substrates including longer chain aliphatic aldehydes such as heptanal **34** and octanal **35**.⁹² Not only did this mutant increase its substrate scope to longer chains but it also showed some interesting cofactor dependent specificities. For example, for most aldehydes the preference, like the wild-type enzyme, was for NADPH over NADH. However, in the case of heptanal this was reversed; with the W145A mutant displaying a much higher specific activity with NADH over NADPH.⁹² This was interesting as it highlights a factor to consider when engineering these nat-AmDHs for a range of substrates. To better understand why the preferences for substrates seemed to be dependent on cofactors bound, W145A was subjected to structural characterisation using X-ray crystallography.

A sample of purified *CfusAmDH* W145A was provided to us by Genoscope (*L. Ducrot*) in the pET22b(+) modified vector. A separate sample was also expressed from the pET22b(+) modified plasmid and purified in house via NiNTA and SEC as routinely described (*Methods 3.1.6 and 3.2.1*). However, this sample was subsequently not needed for further crystallisation trials. W145A was subjected to initial crystallisation trials using pre-dispensed 96-well screens (*Methods 3.3.1.1*). Crystallisations were typically set up by incubating 20-30 mg.mL⁻¹ of W145A with 4 mM NADP⁺ or NAD⁺ (pre-drop concentrations). The best crystals were subjected to optimisation screens set up in 48-well format. A range of singular crystal morphologies were observed in plate and cuboid form (Figure 103). The best, largest, crystals observed were crystallised with NADP⁺, these crystals were soaked with 10 mM *n*-pentylamine **30** for approximately 30 min prior to fishing.

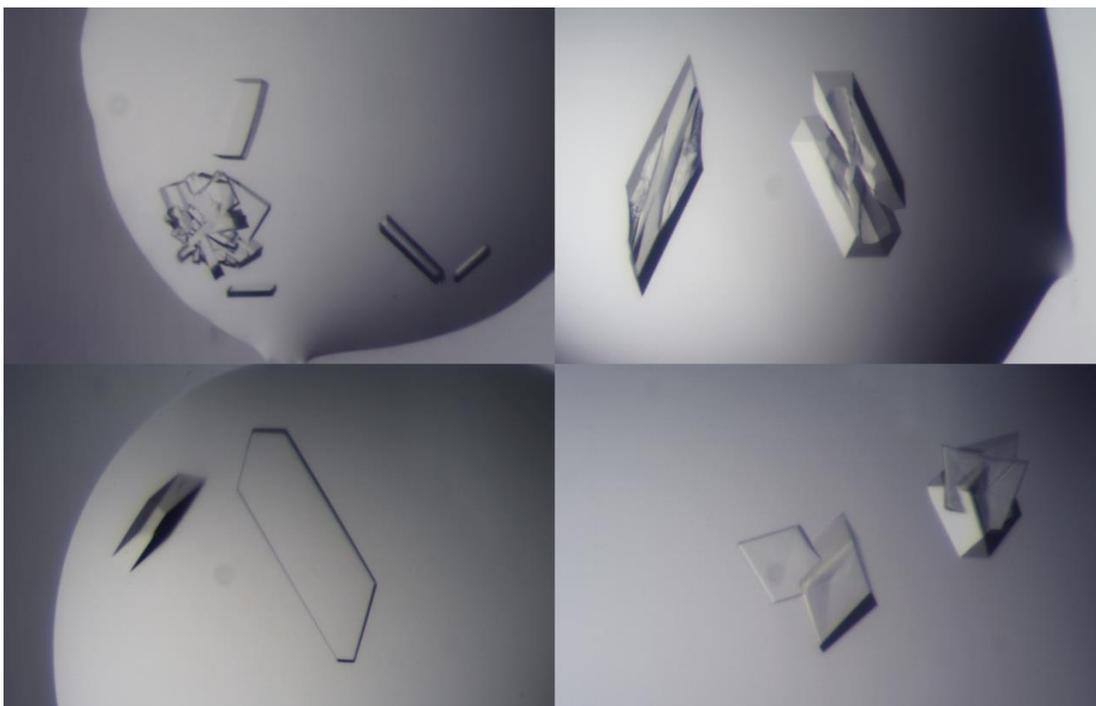


Figure 103: Crystal forms from *CfusAmDH W145A* in complex with NADP⁺ or NAD⁺. Microscopy images of plates and cuboids from a pre-dispensed Index and CSS (MRC-96 well) screens and 48-well optimisations.

Crystals were fished, flashed cooled and sent off to the Diamond synchrotron. The best NAD⁺ crystals were obtained using 0.1 M Tris-HCl buffer pH 8.5 with 0.2 M MgCl₂, 8 % (w/v) PEG 20,000 and 8 % (v/v) PEG 500 MME. The best NADP⁺ crystals were obtained using 0.1 M Bis-Tris buffer pH 6.5 with 25% (w/v) PEG 3350. Two data sets were resolved to 1.64 Å and 1.50 Å for complexes with NAD⁺ and NADP⁺ and *n*-pentylamine respectively. The NAD⁺ complex occupied a $P2_12_12_1$ space group with 2 molecules in the asymmetric unit. The final solution was refined to R values of 17.8% and 21.4% for R_{crystal} and R_{free} respectively. Concerning the NADP⁺ and *n*-pentylamine bound complex it occupied a $C222_1$ space group with two molecules in the asymmetric unit. The final solution for this complex was refined to R values of 16.4% and 18.8% for R_{crystal} and R_{free} respectively. Coordinates for the *CfusAmDH W145A* structures in complex with NAD⁺ or NADP⁺ and *n*-pentylamine were deposited in the PDB under accession codes 7QZN and 7QZL respectively, held for publication. The data and refinement statistics for both datasets are reported in Table 14 and 15.

Table 14 – Data collection and refinement statistics for the 1.64 Å *CfusAmDH* W145A structure in complex with NAD⁺.

1.64 Å *CfusAmDH* W145A:NAD⁺

Beamline	Diamond I03
Wavelength (Å)	0.976261
Resolution (Å)	48.05-1.64 (1.67-1.64)
Space Group	<i>P2₁2₁2₁</i>
Unit cell (Å)	a = 51.02; b = 86.23; c = 142.94 $\alpha = \beta = \gamma = 90^\circ$
No. of molecules in the asymmetric unit	2
Unique reflections	78152 (3786)
Completeness (%)	100.0 (100.0)
R _{merge} (%)	0.07 (1.30)
R _{p.i.m.}	0.03 (0.54)
Multiplicity	11.8 (12.8)
<I/σ(I)>	19.1 (2.0)
Overall B from Wilson plot (Å ²)	22
CC _{1/2}	1.00 (0.73)
R _{cryst} / R _{free} (%)	17.8/21.4
r.m.s.d 1-2 bonds (Å)	0.011
r.m.s.d 1-3 angles (°)	1.71
Avg main chain B (Å ²)	27
Avg side chain B (Å ²)	30
Avg water B (Å ²)	38
Avg NAD ⁺ B (Å ²)	42

Brackets refer to data in the highest resolution shell.

Table 15 – Data collection and refinement statistics for the 1.50 Å *Cfus*AmDH W145A structure in complex with NADP⁺ and *n*-pentylamine.

1.50 Å <i>Cfus</i>AmDH W145A:NADP⁺:<i>n</i>-pentylamine	
Beamline	Diamond I03
Wavelength (Å)	0.976284
Resolution (Å)	47.36-1.50 (1.53-1.50)
Space Group	<i>C</i> 222 ₁
Unit cell (Å)	a = 56.69; b = 84.31; c = 284.13 α = β = γ = 90°
No. of molecules in the asymmetric unit	2
Unique reflections	109306 (5350)
Completeness (%)	100.0 (100.0)
R _{merge} (%)	0.08 (0.96)
R _{p.i.m.}	0.04 (0.40)
Multiplicity	12.2 (13.0)
<I/σ(I)>	17.1 (3.1)
Overall B from Wilson plot (Å ²)	17
CC _{1/2}	1.00 (0.90)
R _{cryst} /R _{free} (%)	16.4/18.8
r.m.s.d 1-2 bonds (Å)	0.012
r.m.s.d 1-3 angles (°)	1.74
Avgc main chain B (Å ²)	20
Avgc side chain B (Å ²)	23
Avgc water B (Å ²)	31
Avgc NAD ⁺ B (Å ²)	15
Avgc Ligand B (Å ²)	30

Brackets refer to data in the highest resolution shell.

As expected, overall, the NAD⁺ and NADP⁺:*n*-pentylamine mutant structures are very similar to the wild-type enzyme (PDB: 6IAU) where residues 3-157 and 297-342 make up the large N-terminal Rossman domain. Residues 158-296 from the C-terminal domain, largely associate to form a beta pleated sheet comprised of 8 strands (Figure 104). The flexible loop hinge region (within the C-terminal domain) which is largely responsible for the hinging during the 'opening and closing' of the substrate binding pockets (Figure 104), as seen with all other nat-AmDHs described previously (*Results and Discussion 4.1, 4.2 & 4.3*). RMSD values between chains A of both W145A structures and wild type enzyme (PDB: 6IAU) were as follows: for the NAD⁺ structure and 6IAU 0.53 Å across 340 Cα atoms, for the NADP⁺:*n*-pentylamine structure and 6IAU 0.35 Å across 340 Cα atoms. RMSD values for the SSM superimposition of chains A between the NAD⁺ and NADP⁺:*n*-pentylamine structures was 0.56 Å for 340 Cα atoms. Again, the biological assemblies are dimers in which both monomers associate at a series of beta pleated sheets and each monomer is responsible for binding one molecule of NAD⁺ or NADP⁺ (and one molecule of ligand where applicable) (Figure 105 and 106). Both structures, similarly to the wild-type enzyme, appear in 'closed' conformations. This is likely since they both have something bound in their active site, in addition to the electron density seen for either NAD⁺ or NADP⁺ (Figure 106), there was additional electron density observed in the binding pocket. In the case of the NADP⁺ bound structure this was fitted as *n*-pentylamine **30** which was soaked into the crystal prior to fishing (Figure 106). In the case of the NAD⁺ structure the additional electron density could be fitted as imidazole which was present in purification buffers (Figure 106).

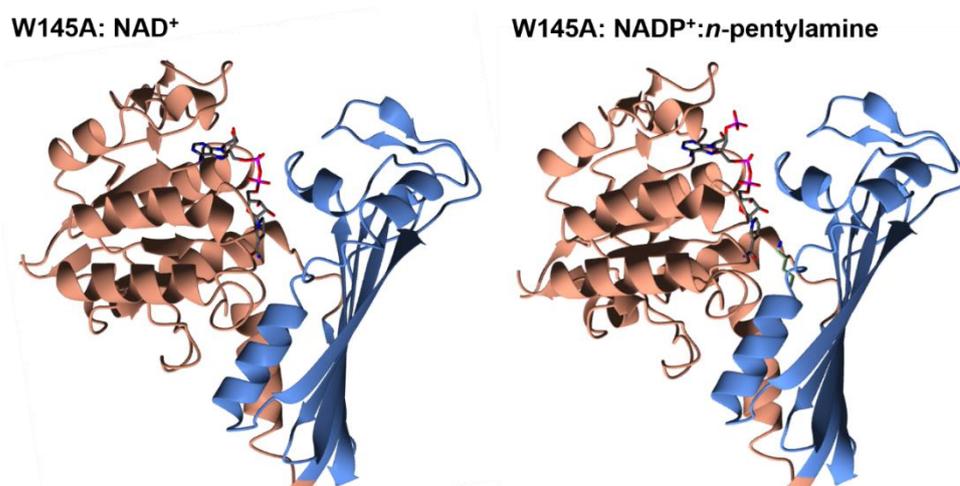


Figure 104: Monomeric forms of *CfusAmDH* W145A bound to either NAD⁺ or NADP⁺ and *n*-pentylamine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the NAD⁺ bound structure of W145A with NAD⁺ visualised in grey and the N-terminal Rossman domain (orange) and C-terminal beta sheet domain (blue) highlighted. The right-hand panel represents the NADP⁺ and *n*-pentylamine bound structure of W145A with NAD⁺ visualised in grey and *n*-pentylamine visualised in green. The N-terminal Rossman domain (orange) and C-terminal beta sheet domain (blue) are highlighted.

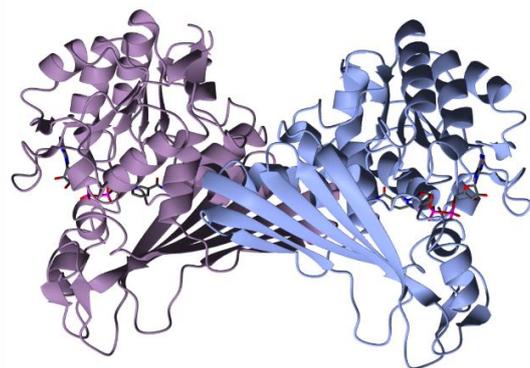
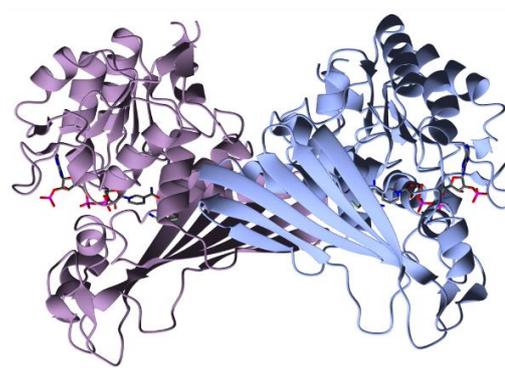
W145A: NAD⁺W145A: NADP⁺:*n*-pentylamine

Figure 105: Biological assembly forms of *Cfus*AmDH W145A bound to either NAD⁺ or NADP⁺ and *n*-pentylamine. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the dimeric NAD⁺ bound structure of W145A with NAD⁺ visualised in grey. Chains A and B are represented in lilac and blue respectively. The right-hand panel represents the dimeric NADP⁺ and *n*-pentylamine bound structure of W145A with NAD⁺ visualised in grey and *n*-pentylamine visualised in green. Chains A and B are represented in lilac and blue respectively.

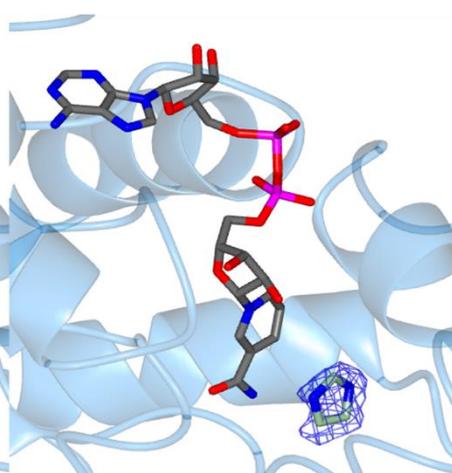
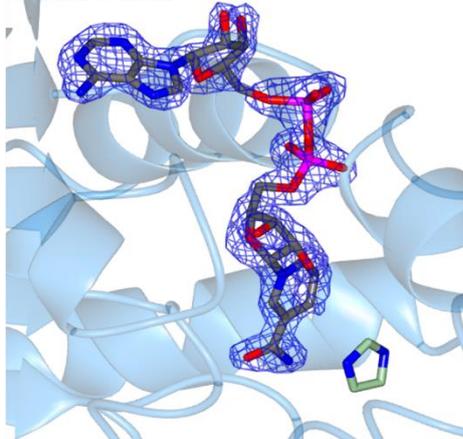
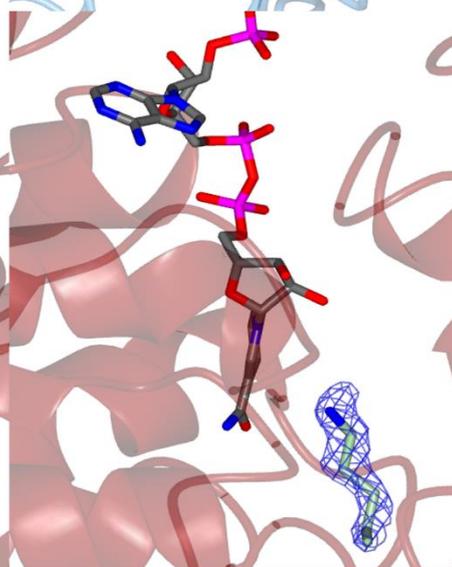
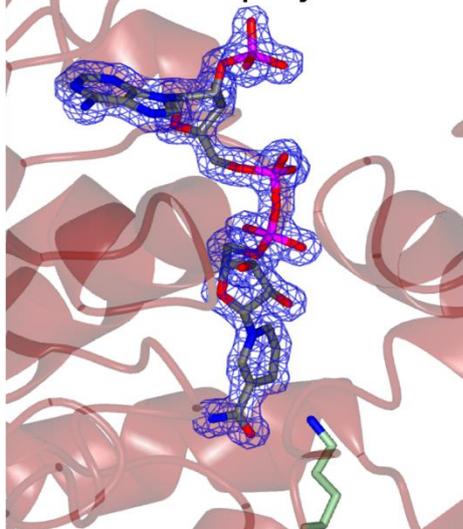
W145A: NAD⁺W145A: NADP⁺:*n*-pentylamine

Figure 106: Omit maps for the cofactors and ligands of both *Cfus*AmDH W145A structures. All images have been visualised using CCP4mg (v.2.10.11). All ligands are visualised in green, with NAD(P)⁺ in grey and omit maps in blue cylindrical form. The top left-hand panel represents the NAD⁺ omit map. The top right-hand panel represents the imidazole omit map. The bottom left-hand panel represents the NADP⁺ omit map with the bottom right-hand side representing the *n*-pentylamine omit map. All $F_o - F_c$ maps are displayed in blue cylindrical form at a map level of 3 σ .

4.4.2 *Tther*AmDH

*Tther*AmDH was discovered by the genomic mining of environmental samples which were validated against reference nat-AmDHs such as *Micro*AmDH, *Msm*eAmDH and *Cfus*AmDH as described previously in *Introduction 1.3.1*. The biological role of *Tther*AmDH is thought to be associated with ornithine degradation where it acts as a diaminopentanoate dehydrogenase, similarly to 2,4-DAPDH (*Introduction 1.3*).^{89,90} Given the sequence similarity of *Tther*AmDH to the first extensively characterised nat-AmDH, AmDH4 (*Introduction 1.3*)^{89,91}, and its potential thermophilic properties (*Results and Discussion 4.4.4*)¹⁰³, it was thought that *Tther*AmDH could be exploited and engineered for substrate promiscuity and activity towards a range of unfunctionalised carbonyl centres. Therefore, understanding the nature of this enzyme, structurally, was of importance.

*Tther*AmDH was expressed from a pET22b(+) plasmid, optimised for E.coli expression, provided by Genoscope. *Tther*AmDH was transformed into BL21(DE3) cells and expressed at 37°C as routinely described (*Methods 3.1.5 and 3.1.6*). Subsequently *Tther*AmDH was purified via NiNTA and SEC (*Methods 3.2*) to yield approximately 24 mg per 2 L of LB culture. From the nickel affinity chromatogram there appears to be one broad peak which is probably the accumulation of poor resolution from multiple peaks (Figure 107). This was confirmed using SDS-PAGE analysis in which the latter half of the peak contained pure protein (Figure 109). In the size exclusion chromatogram, there were two distinct peaks (Figure 108), both peaks contained *Tther*AmDH but the second peak contained less co-eluent i.e. purer *Tther*AmDH, as confirmed by SDS-PAGE analysis (Figure 109). ze

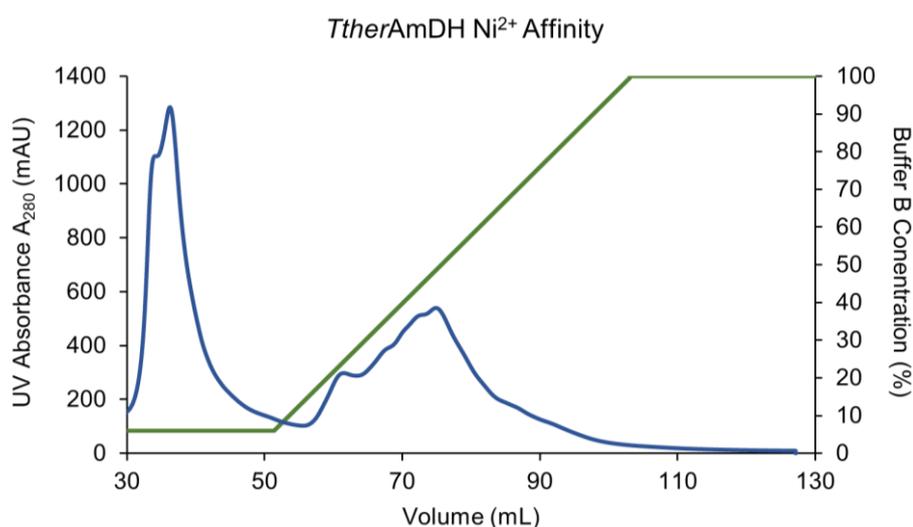


Figure 107: Nickel Affinity Chromatogram for *Tther*AmDH expressed in the pET22b(+) vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line.

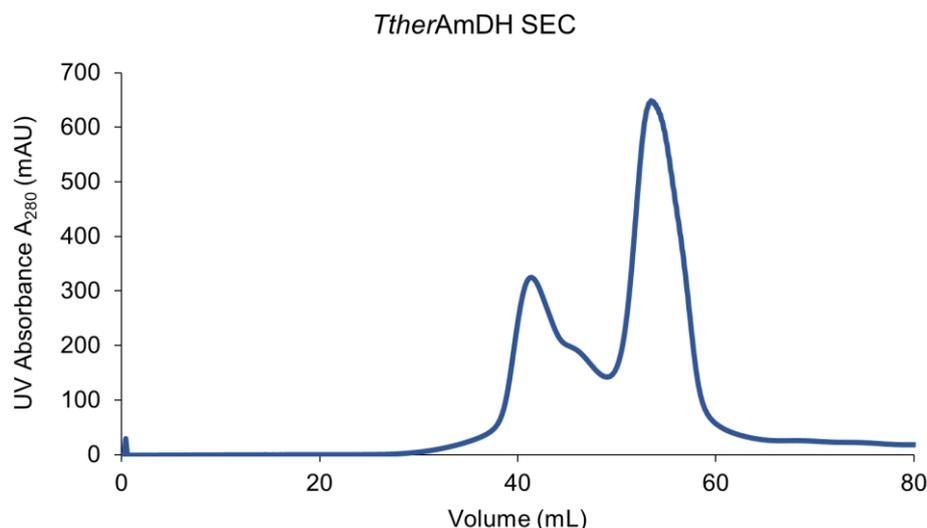


Figure 108: Size Exclusion Chromatogram for *TtherAmDH* expressed in the pET22b(+) vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue).

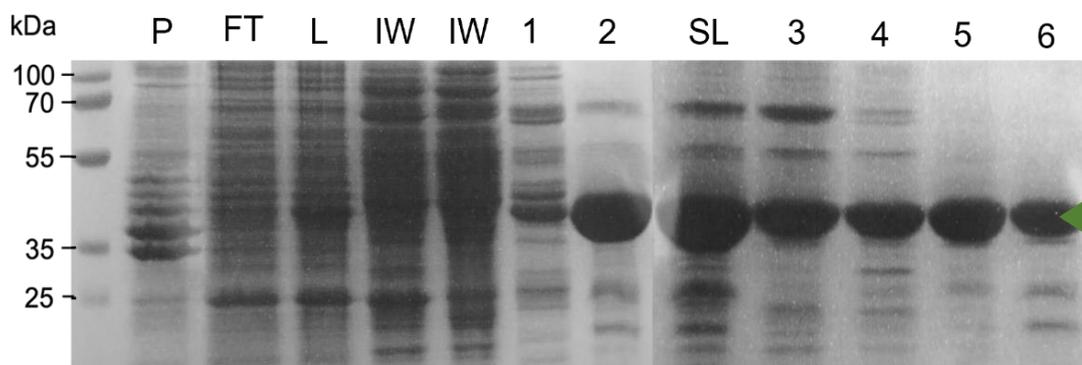


Figure 109: SDS-PAGE analysis of protein fractions from pET22b(+) *TtherAmDH* protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane P*: 5 x diluted sample of cell pellet after lysis, resuspended in 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane FT*: 5 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane L*: 5 x diluted sample of cell lysate that was loaded onto the 5 mL FF HisTrap™ column. *Lane IW*: 2 x diluted samples of the initial wash, before the imidazole gradient, on the ÄKTA using 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane 1*: Sample from an ÄKTA fraction relating to the first half of the broad peak as determined from the Nickel Affinity chromatogram. *Lane 2*: 2x diluted sample from an ÄKTA fraction relating to the second half of the broad peak as determined from the Nickel Affinity chromatogram. *Lane SL*: 5 x diluted pooled and concentrated protein loaded onto the Superdex® column. *Lanes 3&4*: Samples from ÄKTA fractions of the first peak during Size Exclusion Chromatography. *Lanes 5&6*: 2x diluted samples from ÄKTA fractions of the second peak during Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

*Tther*AmDH underwent initial crystallisation trials using a range of commercially available screens provided by Hampton Research, namely the Index, CSS and PACT screens. CSS screens were set up using two buffers: Bis-Tris pH 6.5 and HEPES pH 7.5 with final concentrations of 0.1 M. After initial trials, conditions were optimised for 48-well MRC plates using drops containing 0-2 M NaCl or MgCl₂, 0.1 M HEPES or Bis-Tris pH 6.5-9.0 and PEG 3350 20-25% (w/v). The best collected crystals were from drops containing 0.2 M NaCl, 0.1 M HEPES pH 8.0, PEG 3350 20% (w/v) with final concentrations of 10 mg.ml⁻¹ *Tther*AmDH:2 mM NADP⁺ (Figure 110). Crystals were frozen in liquid nitrogen without any additional cryoprotectant added. A data set was resolved to 2.71 Å and occupied a *I*121 space group with 2 molecules in the asymmetric unit. The final solution was resolved to R factors values of 23.1% and 27.9% for R_{crystal} and R_{free} respectively. Data and refinement statistics are reported in Table 16. Unfortunately, no NADP⁺ electron density was resolved, and the cofactor therefore was not fitted into the final structural solution.

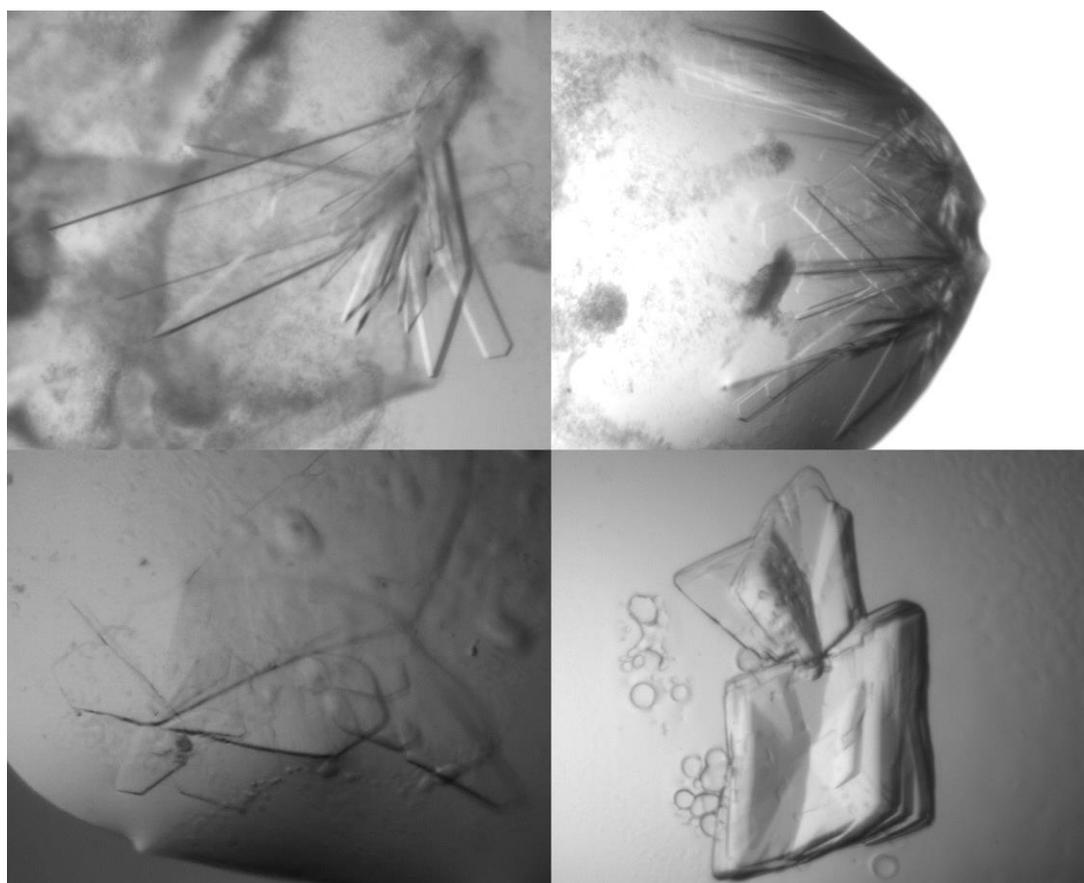


Figure 110: Crystal forms from *Tther*AmDH in complex with NADP⁺. Microscopy images of plates and cuboids from a pre-dispensed Index, PACT and CSS (MRC-96 well) screens and 48-well optimisations.

Table 16 – Data collection and refinement statistics for the 2.71 Å apo-*TtherAmDH* structure.

2.71 Å <i>TtherAmDH</i>:NADP⁺	
Beamline	Diamond I03
Wavelength (Å)	0.976261
Resolution (Å)	50.40-2.71 (2.84-2.71)
Space Group	I121
Unit cell (Å)	a = 114.477; b = 53.896; c = 182.171 $\alpha = \gamma = 90^\circ, \beta = 97.894^\circ$
No. of molecules in the asymmetric unit	2
Unique reflections	30369 (1524)
Completeness (%)	100.0 (99.6)
R _{merge} (%)	0.067 (1.173)
R _{p.i.m.}	0.067 (1.173)
Multiplicity	1.9 (1.9)
<I/σ(I)>	20.4 (5.0)
Overall B from Wilson plot (Å ²)	86.55
CC _{1/2}	0.998 (0.700)
R _{cryst} / R _{free} (%)	23.1/27.9
r.m.s.d 1-2 bonds (Å)	0.006
r.m.s.d 1-3 angles (°)	1.668
Avg main chain B (Å ²)	86
Avg side chain B (Å ²)	88
Avg water B (Å ²)	59

Brackets refer to data in the highest resolution shell.

As expected, due to the high sequence similarity to AmDH4, *Tther*AmDH displays the typical 'AmDH fold'. The biological assembly consists of a dimer (Figure 111) in which each monomer is thought to be catalytic and would typically bind one molecule of NAD(H)/NADP(H) and appropriate ligand(s). Residues 1-153 and 295-344 make up the large N-terminal Rossman domain with residues 154-294 making up the C-terminal beta pleated sheet domain comprised of 8 beta strands as illustrated in Figure 109. As mentioned previously no NADP⁺ was able to be fitted into the binding pocket despite co-crystallisation attempts. In the latter discussion of *Tther*AmDH (*Results and Discussion 4.4.4*), the structural properties will be explored in relation to other nat-AmDHs.

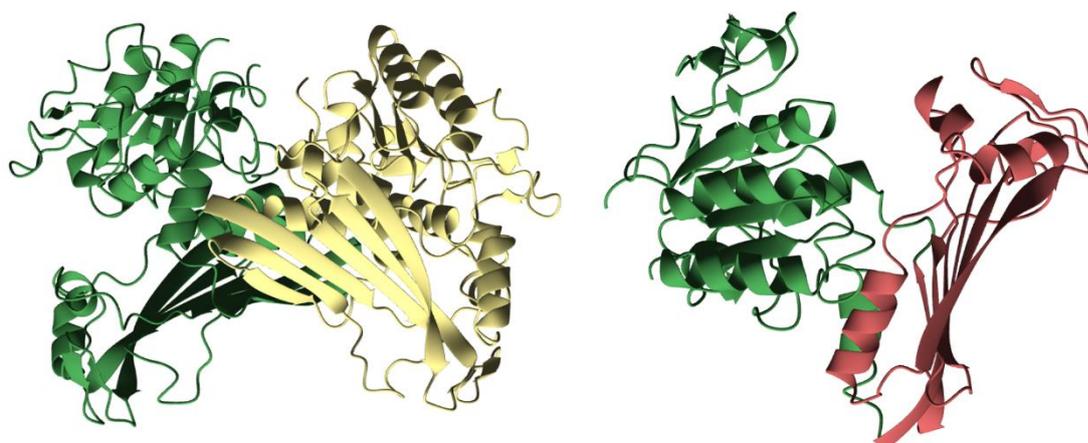


Figure 111: Biological assembly and domain assignment of apo-*Tther*AmDH. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the dimeric form of *Tther*AmDH with chain A represented in green and chain B in yellow. The right-hand panel represents the monomeric form of *Tther*AmDH with the N-terminal Rossman domain highlighted in green and the C-terminal beta sheet morphology highlighted in red.

4.4.3 *Micro*AmDH

*Micro*AmDH, as discussed in *Introduction 1.3.1*, which offers the same substrate capacity as *Cfus*AmDH and *Msme*AmDH, contrastingly, has not been structurally studied via X-ray crystallography. Due to *Micro*AmDH's wide substrate scope and specificity it was an interesting target to study potentially allowing for the structurally informed engineering of this enzyme.⁹¹

*Micro*AmDH was expressed in the pET22b(+) vector which was provided and optimised for *E.coli* expression by Genoscope. *Micro*AmDH was transformed into BL21(DE3) cells and expressed at 37°C as previously detailed (*Methods 3.1.5 and 3.1.6*). *Micro*AmDH was then subjected to purification using NiNTA, and then SEC, to give a yield of 5-20 mg for 2 L of LB culture. From the NiNTA chromatogram there is one *specific* protein peak (Figure 112), albeit at a low concentration, which was

confirmed to be the correct molecular weight by SDS-PAGE analysis (Figure 114). From the Size Exclusion chromatogram there were two distinct peaks with both containing *MicroAmDH* (Figure 113); the second peak contained more *MicroAmDH* with less additional co-eluted proteins, as confirmed by SDS-PAGE analysis (Figure 114). The pure and 'polluted' fractions were pooled separately and concentrated to appropriate concentrations for crystallisation.

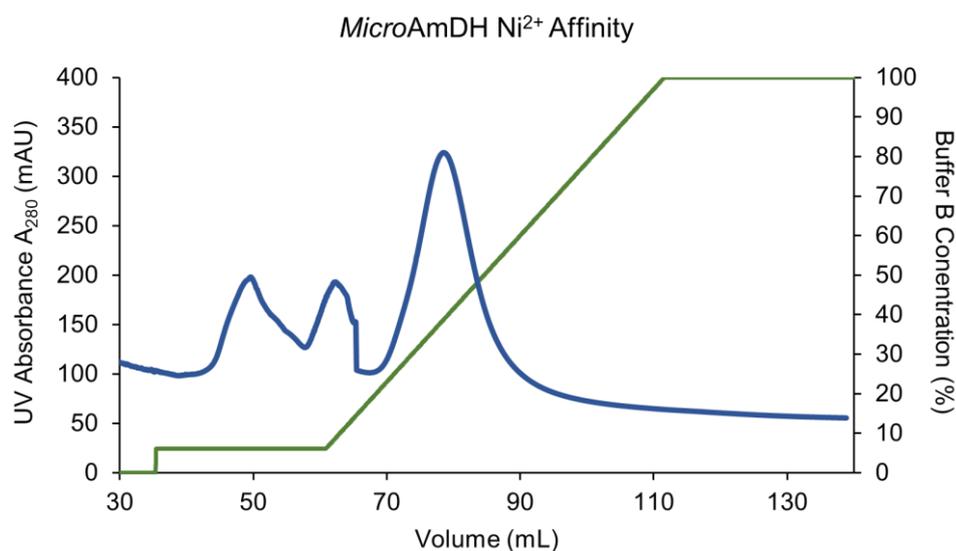


Figure 112: Nickel Affinity Chromatogram for *MicroAmDH* expressed in the pET22b(+) vector. The chromatogram from a round of Nickel Affinity chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue). The right-hand y-axis represents Buffer B concentration (%) which is represented by the green gradient line.

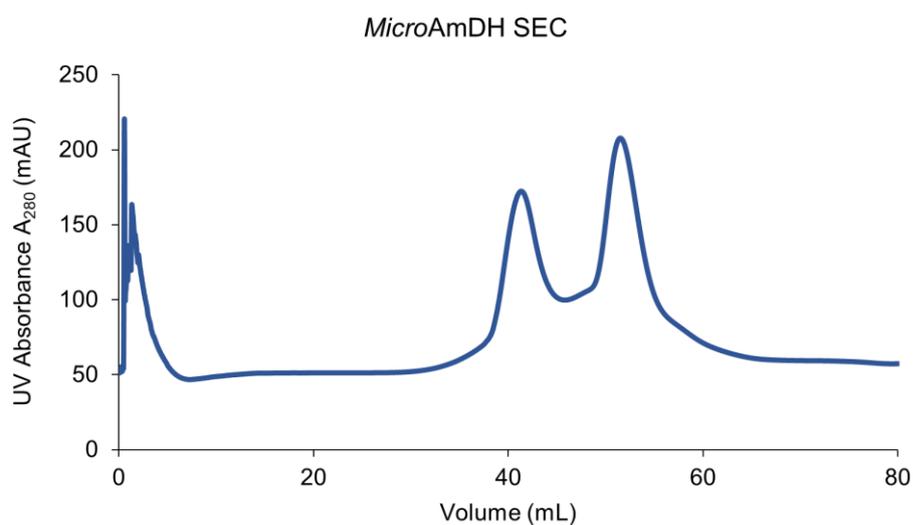


Figure 113: Size Exclusion Chromatogram for *MicroAmDH* expressed in the pET22b(+) vector. The chromatogram from a round of Size Exclusion chromatography. The x-axis represents the volume, in mL, as the column time progresses. Left-hand y-axis represents the UV absorbance at 280 nm in mAU, points are plotted as a smooth line (blue).

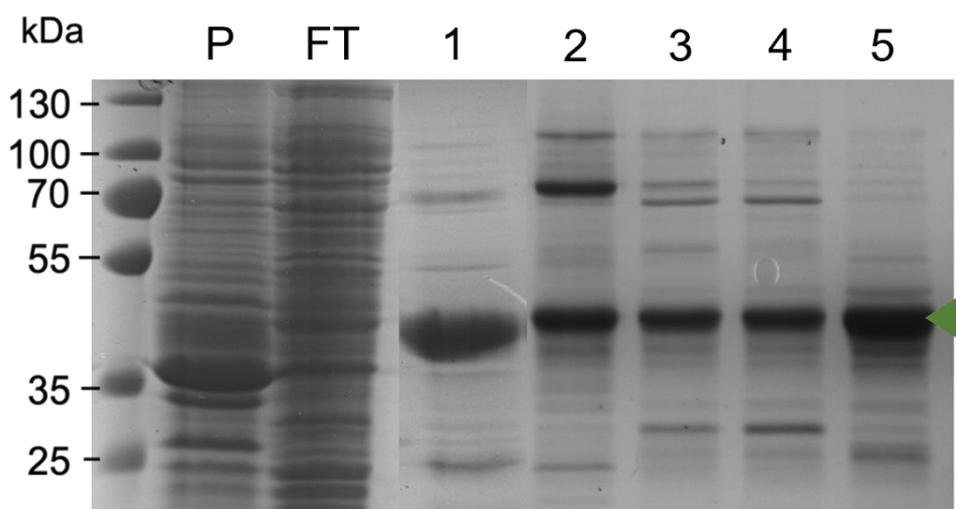


Figure 114: SDS-PAGE analysis of protein fractions from pET22b(+) *MicroAmdH* protein purification. The left-hand lane represents a broad range ladder (ThermoFisher) and associated molecular mass references in kDa. *Lane P*: 2 x diluted sample of cell pellet after lysis, resuspended in 50 mM Tris-HCl pH 7.1, 300 mM NaCl. *Lane FT*: 2 x diluted flow through from the 5 mL FF HisTrap™ column when loading cell lysate. *Lane 1*: 2x diluted sample from an ÄKTA fraction relating to the specific protein peak as determined from the from the Nickel Affinity chromatogram. *Lanes 2&3*: 1.5 x diluted samples from ÄKTA fractions relating to the first peak as determined from the from the Size Exclusion chromatogram. *Lanes 4&5*: 1.5 x diluted samples from ÄKTA fractions of the second peak during Size Exclusion Chromatography. Samples were run on a 12% gel for approximately 1 hour at 200 V. The green triangle represents the bands of interest.

Both pure and polluted fractions were subjected to all pre-dispensed commercially available screens (MRC 96 well) that were present in house, including PACT, CSS, Index, PEG Ion, JCSG and Morpheus screens. Additionally, a range of 96-well and 48-well optimisation screens were manually designed and dispensed for a sitting drop format. Moreover, a range of additives such as ethylene glycol, MPD and glycerol were screened. Furthermore, a range of ligands such as hexylamine **33**, pentylamine **30** and cyclohexylamine **21b** were used for co-crystallisation with either NADP(H) or NAD(H). Unfortunately, screening yielded no substantial crystals that could be fished or even further optimised or seeded. Only some small possible spherulite crystals were observed occasionally (Figure 115).

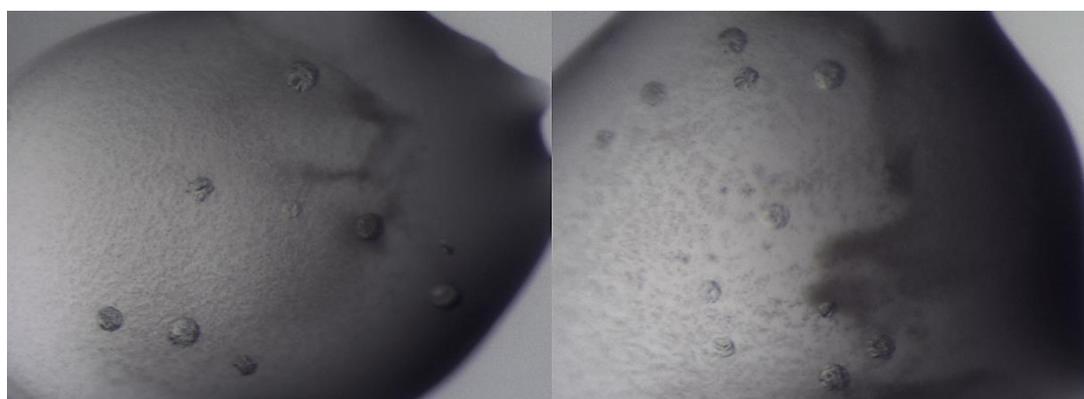


Figure 115: Crystal forms from *MicroAmdH* in complex with NADP⁺ and pentylamine/hexylamine. Microscopy images of potential spherulite crystals from a pre-dispensed Index (MRC-96 well) screen.

4.4.4 Discussion

*Micro*AmDH unfortunately, was not able to be studied structurally as no X-ray crystal structure was obtainable. This could possibly be a solubility or stability issue because of poor optimisation of expression and purification. This is also noticeable by the presence of multiple peaks i.e., co-eluent during the protein purifications (Figure 112 - 114) as well as the noticeably low yield obtained during each round of expression and purification. However, *Micro*AmDH was subsequently sub-cloned into the pETYSBLIC-3C by an MChem student (*J. Jones – work not presented*), which significantly improved yield and purity, post purification, with the hopes of improving crystal morphology quality. However, once again these crystallisation trials and optimisations only yielded either spherulite or salt crystals. Consequently, *Micro*AmDH was not taken any further regarding the evolving, engineering, or structural investigations.

Given *Tther*AmDH's sequence similarity (53.9% Protein:Protein BLAST – *Appendices: Figure 133*) to AmDH4 (from *Petrotoga mobilis*), and its thermophilic nature, it is not surprising that this was an ideal candidate in screening for the reductive amination of a range unfunctionalised carbonyl centres. When analysed by differential scanning fluorimetry, *Tther*AmDH had the highest reported T_m of all analysed AmDHs at a value of $\sim 70^\circ\text{C}$.¹⁰³ Therefore, structurally characterising *Tther*AmDH using X-ray crystallography was an attractive goal. The apo structure of *Tther*AmDH was resolved with both chains appearing in the 'open' form. This is not surprising given the lack of cofactor (and ligand) that was able to be trapped in the active site via co-crystallisation. The open form of *Tther*AmDH appears to be very similar to the open form of *P. mobilis* AmDH4 (PDB: 6G1M) with RMSD values of 0.87 Å for 328 Ca atoms (Figure 114). Looking at the active site reveals that *Tther*AmDH contains almost identical residues to AmDH4 that make up the substrate cage (Figure 116). Crucially E104 which is positioned in a similar orientation as E102 from AmDH4 (Figure 116); which is in the ideal position and distance from modelled natural substrates⁹¹, facilitating the activation of ammonia (or an amine donor) for attack at the carbonyl centre. Additionally, R163, N165, H199 and H266, and equivalents in AmDH4, which are thought to be responsible for maintaining electrostatic interactions, are found within the binding pocket (Figure 116).⁹¹ Moreover F170 (F168 in AmDH4) is responsible for closing the substrate cage with the nicotinamide ring of NAD(P)⁺/NAD(P)H (Figure 116). It is hard to determine the true open structure of *Tther*AmDH without at least a cofactor bound structure. Also, it is difficult to make any structural deductions about how the 'closed' form may look on ligand binding without

this structure available to us. It would be interesting to see how dynamic the open structure vs the closed structure is of *Tther*AmDH, potentially owing to some of its thermophilic properties. Preliminary screening data on *Tther*AmDH (*L. Ducrot et al – Unpublished*) indicated that *Tther*AmDH was in fact active against functionalised substrates such as 2,4-diaminopentanoate (2,4-DAP) **25a** (*Introduction 1.3*) but not unfunctionalised carbonyls of interest such as cyclohexanone **21** or isobutyraldehyde **28**. Therefore, due to its narrow and specific substrate scope and lack of structural novelty it was decided not to characterise this enzyme any further in terms of structurally informed engineering.

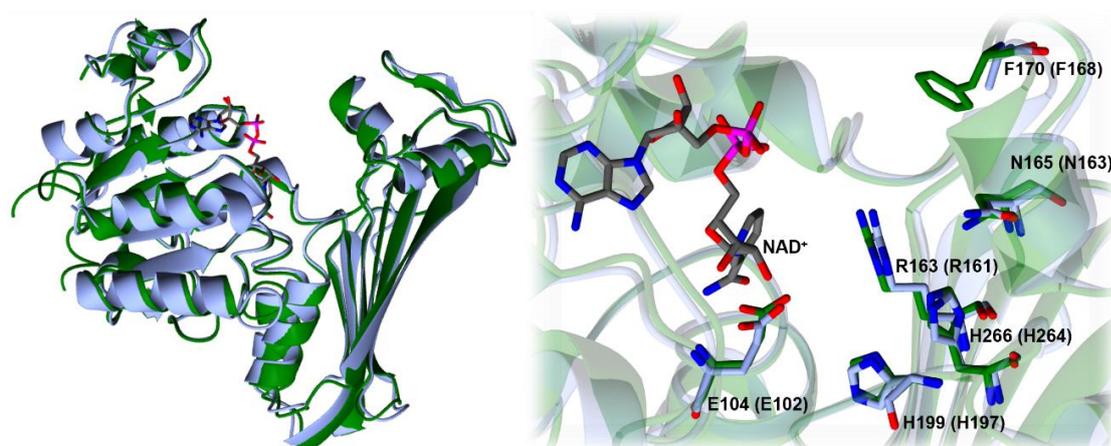


Figure 116: A structural comparison of *Tther*AmDH and AmDH4. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents an SSM superimposition of the monomeric forms of *Tther*AmDH (green) and AmDH4 (PDB: 6G1M) (blue) in ‘open’ conformations from chains A. The right-hand panel represents the binding site residues from chain A of *Tther*AmDH (green) and chain A of AmDH4 (PDB: 6G1M) (blue) SSM superimposed. Residues are labelled accordingly with AmDH4 residues bracketed. *N.B.* NAD⁺ is modelled from AmDH4 (PDB: 6G1M) and highlighted in grey throughout.

*Cfus*AmDH W145A crystals formed large plates or cuboids with the best singular crystals being formed via the co-crystallisation with NADP⁺. Subsequently these crystals were selected for soaking with appropriate ligands such as *n*-pentylamine **30**. Both the NAD⁺ and NADP⁺:*n*-pentylamine structures were resolved with all chains in the ‘closed’ conformation. Overall, both these structures were very similar to the closed conformations of the cyclohexylamine **21b** bound wild-type enzyme (PDB: 6IAU) (Figure 117). RMSD values for all closed chains in the wild-type and mutant structures are reported above (*Results and Discussion 4.4.1*). Both of the *Cfus*AmDH W145A structures contain E108 in the optimal position for activating the amine donor as does the wild-type enzyme (Figure 117). As expected, the substrate cage is closed in the same way by hydrophobic and aromatic residues F140, T166, Y168, Y173 and Y202 (Figure 117). What is clear from these crystal structures is the extra afforded space at the bottom of the active site that is present in the mutant structures (Figure

117). This would allow for straight chain, aliphatic substrates larger than *n*-pentylamine, to bind within the active site that is either not permitted in the wild-type structure, or requires an active rotation of W145 away from the centre (Figure 117). This aligns with recent findings depicting that *Cfus*AmDH W145A has a much higher specific activity than that of the wild-type enzyme towards larger substrates such as heptanal and octanal.⁹² Interestingly in the NAD⁺ structure Y202 could be fitted in two conformations sharing 50% occupancy each. Either Y202 rotates into the space previously occupied by the indole ring of W145 or away from this mutation held in the typical conformation displayed by the wild-type enzyme (Figure 117). In the NAD⁺:*n*-pentylamine structure Y202 is held in place away from the floor of the active site in the orientation that is also adopted in the wild-type enzyme. The extra flexibility of this tyrosine side chain could explain a specific preference towards substrates of different chain length. Albeit this finding could be an artifact since the NAD⁺ structure is not bound to a ligand which could potentially provide the extra rigidity to the Y202 side chain. However, this not without merit; the extra flexibility in this region is an interesting finding which could be used, along with cofactor specificity, to fine tune these enzymes for either specificity or promiscuity towards different substrates which is cofactor dependent.

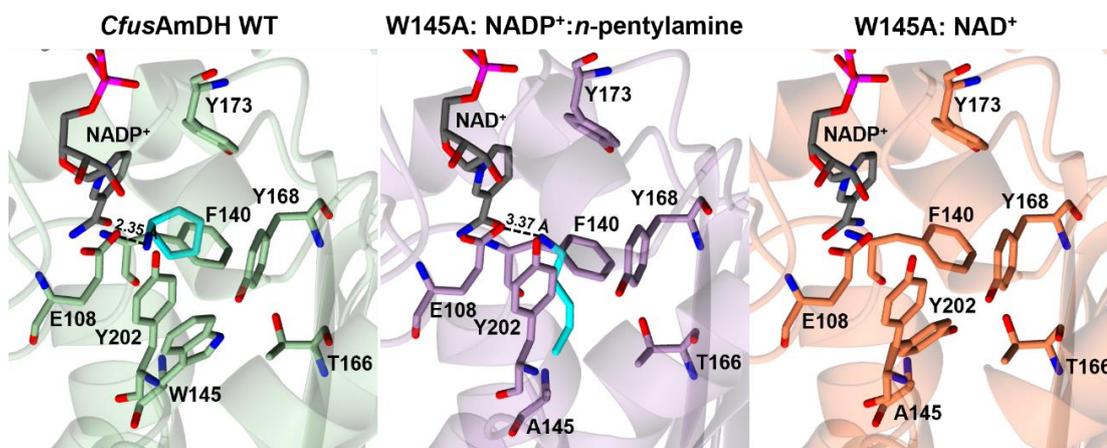


Figure 117: A structural comparison of *Cfus*AmDH wild-type and W145A structures. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents the active site residues of the wild-type enzyme in green (PDB: 6IAU) in complex with cyclohexylamine (blue) and NADP⁺ (grey). The middle panel represents the active site residues of the W145A structure (purple) (PDB: 7QZL) in complex with NADP⁺ (grey) and *n*-pentylamine (blue). The right-hand panel represents the active site residues of the W145A structure (orange) (PDB: 7QZN) in complex with NAD⁺ (grey). Interactions are highlighted with a dashed line and all distances are represented in Ångstroms. Residues are labelled accordingly.

Additionally, there are notable differences between both W145A structures that could offer explanations for cofactor specificity. In both the NAD⁺ and NADP⁺:*n*-pentylamine structure the 3'OH positions are secured by interactions from the side chains of D36 and R41 (Figure 118). In the NADP⁺ bound structure, the ribose 2' phosphate group is stabilised by interactions from the side chains of H37 and N38 (Figure 118). However, in the case of the NAD⁺ structure when substituted for the 2' hydroxyl group the only interaction comes from the side chain of H37 (Figure 118). The overall reduction in interactions in the NAD⁺ bound structure, compared to the NADP⁺ bound structure, could explain the general observable preference of NADPH over NADH for the reductive amination of a range of ketones and aldehydes.⁹² In general, *CfusAmDH* and associated mutants, display a 4-fold higher specific activity towards aliphatic carbonyls when using NADPH as the cofactor in comparison to when utilising NADH.⁹² Interestingly in the NADP⁺:*n*-pentylamine structure R41 is rotated downwards whereas in the wild-type and NAD⁺ structures it is fixed towards the cofactor (Figure 118). This fixed position facilitates the ionic interaction between R41 and D36. This interaction could be competing with the interactions towards stabilising the cofactor which could be a further explanation as to why *CfusAmDH* W145A:NADP⁺ seems to outperform the wild-type enzyme, in terms of reductive amination capabilities.⁹²

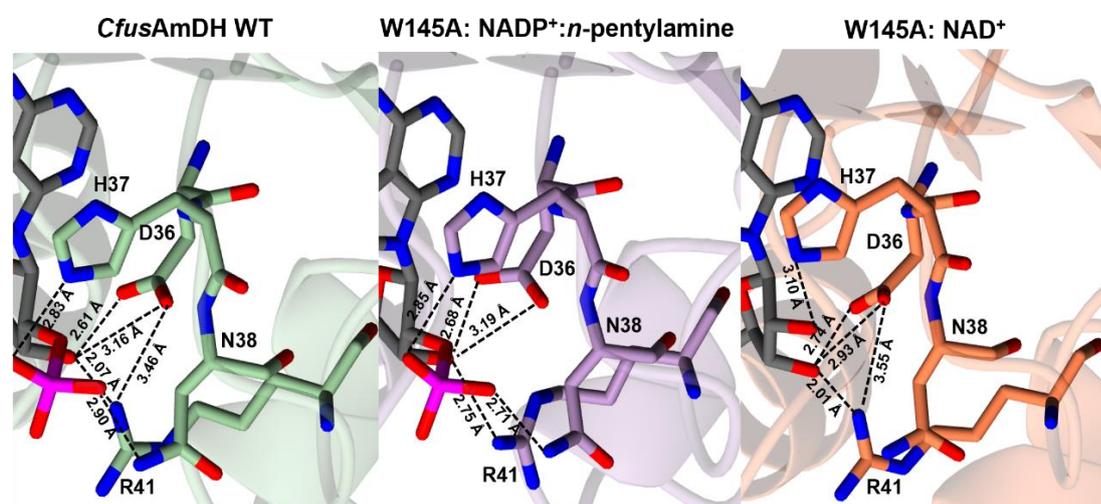


Figure 118: Cofactor stabilisation and interactions of *CfusAmDH* wild-type and W145A structures. All images have been visualised using CCP4mg (v.2.10.11). The left-hand panel represents residues of the wild-type enzyme in green (PDB: 6IAU) in complex with NADP⁺ (grey). The middle panel represents the active site residues of the W145A structure (purple) (PDB: 7QZL) in complex with NADP⁺ (grey). The right-hand panel represents the active site residues of the W145A structure (orange) (PDB: 7QZN) in complex with NAD⁺ (grey). Interactions are highlighted with a dashed line and all distances are represented in Ångstroms. Residues are labelled accordingly.

5. Summary and Conclusions

Amine dehydrogenases (AmDHs) are a class of NAD(P)H dependent enzymes that can carry out asymmetric reductive aminations to produce chiral amines. The advantages of using a biocatalytic route for the production of chiral amines is clear; enzymes offer improved enantioselectivity with the promise of greener and more sustainable syntheses. In previous years AmDHs, and subsequent engineering, has improved the access to a wider range of significant chiral amine functionalities. This is often owed to their promiscuity for carbonyl substrates, and in more recent years, amine donor spectra, and even the acceptance of hydroxyl functionalities. Recently, attention has turned to finding enzymes present in nature that are able to carry out these reductive aminations. Screening of metagenomic samples and databases has subsequently led to the discovery of the native amine dehydrogenase family (nat-AmDHs). These nat-AmDHs offer advantages over previously characterised enzymes involved in the production of amines, such as transaminases and amino acid dehydrogenases, due not only to nat-AmDHs improved substrate promiscuity, but also the ability to accept amine donors not limited to just ammonia. Additionally, nat-AmDHs offer advantages over imine reductases (IREDs) or even reductive aminases (RedAms) because they also work very well with ammonia which IREDs and RedAms typically do not. Therefore nat-AmDHs offer an exceptional scaffold for the rational engineering or directed evolution to create a remarkable biocatalyst for the formation of many chemically and biologically significant primary, secondary and tertiary chiral amines.

MATOUAmDH2, a nat-AmDH, was initially screened and displayed noteworthy substrate and donor scopes with excellent activities. MATOUAmDH2 has successfully been studied structurally using X-ray crystallography to obtain both cofactor-bound and co-crystallised cyclohexylamine **21b** bound structures. The results of these structural investigations led to the discovery of an active site which displayed typical canonical residues such as E111, which along with its counterpart in other nat-AmDHs, is thought to be involved in the activation of the amine donor to allow for nucleophilic attack at the carbonyl centre of the substrate. More extraordinarily the active site of MATOUAmDH2 was found to be much more 'open' than that of previously characterised nat-AmDHs such as *Cfus*AmDH and *Msme*AmDH. The extra space afforded in the MATOUAmDH2 binding pocket was largely due to shifts at the ceiling and smaller substitutions in the floor of the active site (C148 and A147).

With the promising activities displayed by MATOUAmDH2 towards a range of substrates, and its slight structural novelty, it was targeted for rational engineering. Initially, a range of alanine point mutations concerning residues lining the binding pocket, were designed and created to elucidate specific non-catalytic roles of these positions, dictated by their chemistry and steric effects. Both M215A and L180A mutants showed biocatalytic improvements for the production of primary and secondary amines respectively. The extra space afforded at the floor (M215A) or ceiling (L180A) of the active site, that lead to the improvements in activities, was further probed by creating a reduced library of mutations at these positions. Ultimately, mutations concerning the reduction in size i.e., alanine or serine-based mutants, performed the best and were capable of improved binding for larger, more sterically hindered, substrates such as norcamphor **29a** and 1,2-cyclohexandione **31**. Good conversions were observed concerning the production of a range of cyclic primary amines and some cyclic secondary amines.

A selection of other nat-AmDHs of interest, mainly selected from their promising substrate promiscuities or thermophilic properties, were to be characterised structurally. *Micro*AmDH was not pursued after initial crystallisation trials due to experimental complications. Likewise, *Tther*AmDH was not pursued any further in co-crystallisation studies due to the evidence of limited substrate scope. A mutant of *Cfus*AmDH, W145A, showed improved activity, compared to the wild-type enzyme, towards a range of larger aliphatic substrates such as heptanal **34** and octanal **35**. In addition, this activity also seemed to be dictated by cofactor preferences showing switching activities with either NAD(H) or NADP(H) depending on chain length of the substrate. *Cfus*W145A was characterised structurally to uncover these novel features. It was clear from the structures obtained that the W145A mutant afforded more space in the floor of the active site, allowing for larger substrates of chain lengths >5 to bind, something which the wild-type binding pocket would not permit. The flexibility seen from Y202, and cofactor-residue mediated interactions could be used as an explanation as to why this mode of binding is dictated by differences in cofactors.

The structural and biochemical studies of these nat-AmDHs described in this thesis has furthered our understanding as to the modes of binding displayed in these enzymes as well as a deeper understanding of their displayed substrate selectivity and promiscuity. The rational engineering of MATOUAmDH2 has been crucial in providing new scaffolds for the biocatalysis of a range of pharmaceutically reactions that can be expanded in the future using random mutagenesis techniques.

Appendices

Table 17 – List of Primers for pETYSBLIC-3C subcloning and mutagenesis

Primer	Sequence (5'-3')
MATOUAmDH2 Forward	TTCCAGGGACCAGCATCGCCTCTGCGTGTTGCG
MATOUAmDH2 Reverse	GGAGAAGGCGCCTTTAAAGACGGCCATGAACGTACGC
pETYSBLIC-3C Forward	AGGCGCGCCTTCTCCTCAC
pETYSBLIC-3C Reverse	TGCTGGTCCCTGGAACAGAACTT
F143A Forward	TCTGATCGGCGGTGGT GCT CTGGACGGTG
F143A Reverse	TCACAGCAAGCACCGTCCAG AGC ACCACCG
L144A Forward	CGGCGGTGGTTT GCT GACGGTGCTTGCTGT
L144A Reverse	CAGCAAGCACCGT AGC GAAACCACCGCCGATC
L169A Forward	AAGTTAGATGGGGT GCT CAGTACAACGTTG
L169A Reverse	TCAACGTTGACTG AGC ACCCCCATCT
L180A Forward	CGGTCAGGT AGCT GCTATCGCCCAC
L180A Reverse	TGGGCGATAGC AGCT TACCTGACCGT
L180G Forward	CGGTCAGGT GGT GCTATCGCCCAC
L180G Reverse	TGGGCGATAGC ACCT TACCTGACCGT
L180H Forward	CGGTCAGGT CAC GCTATCGCCCAC
L180H Reverse	TGGGCGATAGC GTG TACCTGACCGT
L180N Forward	CGGTCAGGT AAT GCTATCGCCCAC
L180N Reverse	TGGGCGATAGC ATT TACCTGACCGT
L180P Forward	CGGTCAGGT CCT GCTATCGCCCAC
L180P Reverse	TGGGCGATAGC AGGT TACCTGACCGT
L180S Forward	CGGTCAGGT TCT GCTATCGCCCAC
L180S Reverse	TGGGCGATAGC AGAT TACCTGACCGT
M215A Forward	TACGTATAACAAC GCT AACGACTGGTTCGC
M215A Reverse	GCGAACCAGTCGTT AGC GTTGTATACGTAT
M215G Forward	TACGTATAACAAC GGT AACGACTGGTTCGC
M215G Reverse	GCGAACCAGTCGTT ACC GTTGTATACGTAT
M215H Forward	TACGTATAACAAC CAC AACGACTGGTTCGC
M215H Reverse	GCGAACCAGTCGTT GTG GTTGTATACGTAT
M215L Forward	TACGTATAACAAC CTT AACGACTGGTTCGC
M215L Reverse	GCGAACCAGTCGTT AAG GTTGTATACGTAT
M215N Forward	TACGTATAACAAC AAT AACGACTGGTTCGC
M215N Reverse	GCGAACCAGTCGTT ATT GTTGTATACGTAT
M215S Forward	TACGTATAACAAC TCT AACGACTGGTTCGC
M215S Reverse	GCGAACCAGTCGTT AGA GTTGTATACGTAT
T312A Forward	GAAAAACCCGCCG GCT CCGGCGATGA
T312A Reverse	GGTCATCGCCGG AGC CGGCGGGTTTT

Red indicates the codon substitution site

Table 18 – Protein sequences of nat-AmDHs

Protein	Sequence
<i>Cfus</i> AmDH W145A*	VPRGSSKRPIRIIQWGCGLMGQTLIRTLREKGAELVGAI DH NAARRDRDAGEVAGLGQSLGVRIHPPDQADAVFREARAD VCILCTRSIMSELAGALRVAARHGVNAITIGEEAFYPWTT SQ ALTEELDQLARANDCTLTGSGA QDVFWGNLITVLAGATHRI DRIVGLTQYNADDYGSALAQKHGVLDPETFAARIGASNS PSYVWNSNEWLCAQLGWRVDIRQQLLPTTHTGTLRSAS LGREVPAGHATGMKAVVVTETHEGPVIETHCVGKLYAPGE VDLNEWTLRGEPDTTVTIRQPATPALTCATVLNRLPQLLAA PPGFVTTDRFTPATYVSRLETEA
MATOUAmDH2:pET22b(+)	MHHHHHHSPLRVALYGFNGNKNEMAKMLVERKDVEIVAVI SNKSNVKGDFGEVIGLAPQGILVTAGLDAAETLRTTNPQIA MLSTLSTVGDIESQLRACAENKVNVTIAEELTFSWTSAP E KTKEMDEL FKEHNVSLIGGGFLDGACCDMARTMAAMMHKI DKLDGGLQYNVDHYGQVLAIAHGVGLSEEEFYAENGP GW TSPTSYPKSYVYNMNDWFASAFGLTVIKTEEVKTP TKAPIE LYSEAIGRAIPVGQCTGMIVTATTTTEEGVI VEKQVVKCYE DGDEDMVFMNLEGNPTGGVGFTMKNP PTPAMTNTIAISR MFQTV DAPAGYITTDKLP TMEAYVHGRL
MATOUAmDH2:pETYSBLIC-3C	HHHHHHSGLVLFQGPASPLRVALYGFNGNKNEMAKML VERKDVEIVAVISNKSNGKDFGEVIGLAPQGILV TAGLDAA ETLRTTNPQIAMLSTLSTVGDIESQLRACA ENKVNVTIAEE LTFWTSAP EKTKEDEL FKEHNV SLIGGGFLDGACCDMA RTMAAMMHKIDKLDGGLQ YNVDHYGQVLAIAHGVGLSEEE FYAENGP GWTSPTSYPKSYVYNMNDWFASAFGLTVIKTEE VKTPTKAPIELYSEAIGRAIPVGQCTGMIVTATTT TEEGVIIV EKQVVKCYEDGDEDMVFMNLEGNPT GGVGFTMKNPPTP AMTNTIAISRMFQTV DAPAGY ITTDKLP TMEAYVHGRL
<i>Micro</i> AmDH	MHHHHHHTNIRAVVYGVGAMNSVITRYLLDKDVEIV GAISR SPDKVGKDLGEVTGLDRRLGVSISDDPHEV FTRTSPDIAVV AITSYLVDAAEHFRIALSHGVN VITLSEEALYPWNTAPELTA ELDALAKEHGV TITGGGFQDSFVWNAVAQLMGTAHRIDSV TGTSSWNVDEYGPELAELQQVGATIEEFDAW CREAVRPP TFGRIALDALVAGAGLTPKQILTRTE PELAHETLHCAALGID VPPGKCI GFTDIDEIR TEEGPVFVFRMSGRLYGPDDSDVNE WTIHGEPDLVMSNGTPPTMATTCTQLVNRIPD VLDADPGF VTVVDL PRLRYRHGRLHDHLSR WSSDRYIVREEL
<i>Tther</i> AmDH* **	MENIKVVVWGLGAMGSGIAKMILFKKGMEIVGAID TDPNKR GKDLNEILGTNSKPVYITSEPQDIKKS ADIIVITSSYVEK VFPLIKLAVENGINVITAE EMAYPSAQHLELAKEIDRLAREN GVSVLGT GINPGFVLDYLIALTGVCVDVDSIKAARINDLSP F GKAVMEEQGVGLTPEEFEEGVKNGTVAGHIG FPESISMIC DALGWKLSGIEQTREPIVSKTYRE TPYARVEPGYVAGCRQI GYGKVDGEVKIE LEHPQQILPQKEGVETGDYIEIKGTPNIKL SIKPEIPGGLGTIALCVNMIPHVINAEPGLV TMLDLPVPRAIM GDARDMIRRR

*= listed without purification tags

**= from UnitProt Accession code: M8DJE2

Red indicates the site of appropriate mutation(s)

Table 19 – Retention times of standards analysed by GC-FID

Compound	Retention Time (min)
21 pET22b(+):NH ₃	3.13
21b pET22b(+)	2.65
21 pET22b(+):CH ₃ NH ₂	2.40
21c pET22b(+)	2.54
21 pETYSBLIC-3C	1.85 - 1.88
21b pETYSBLIC-3C	1.66
21c pETYSBLIC-3C	2.07
31	2.48
31a	8.32
29	2.26
29a	2.44
15	2.85
15b	3.68
20	1.27
20b	1.50
32	2.85
32a	2.76

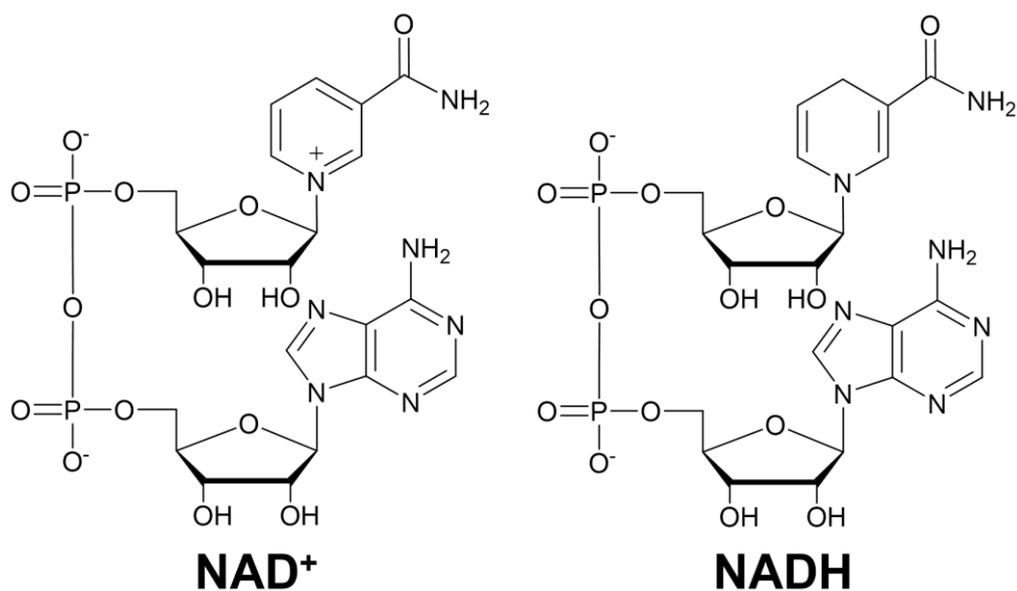


Figure 119a: Chemical structure of NAD⁺ and NADH. Note in the reduced form the reduction of the nicotinamide ring in the structure.

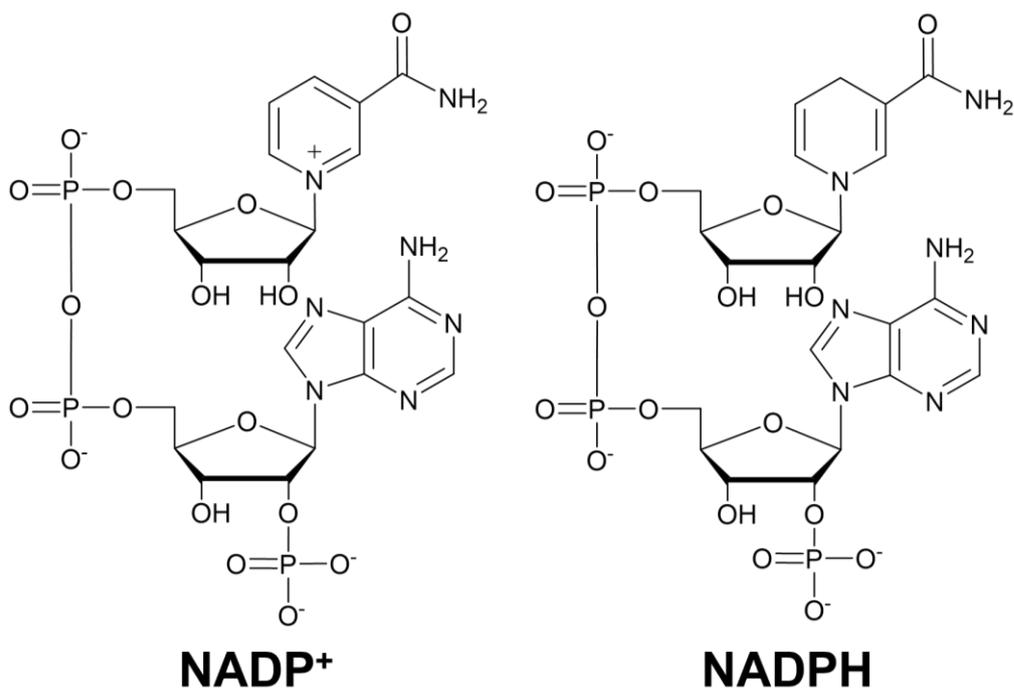


Figure 119b: Chemical structure of NADP⁺ and NADPH. Note in the reduced form the reduction of the nicotinamide ring in the structure.

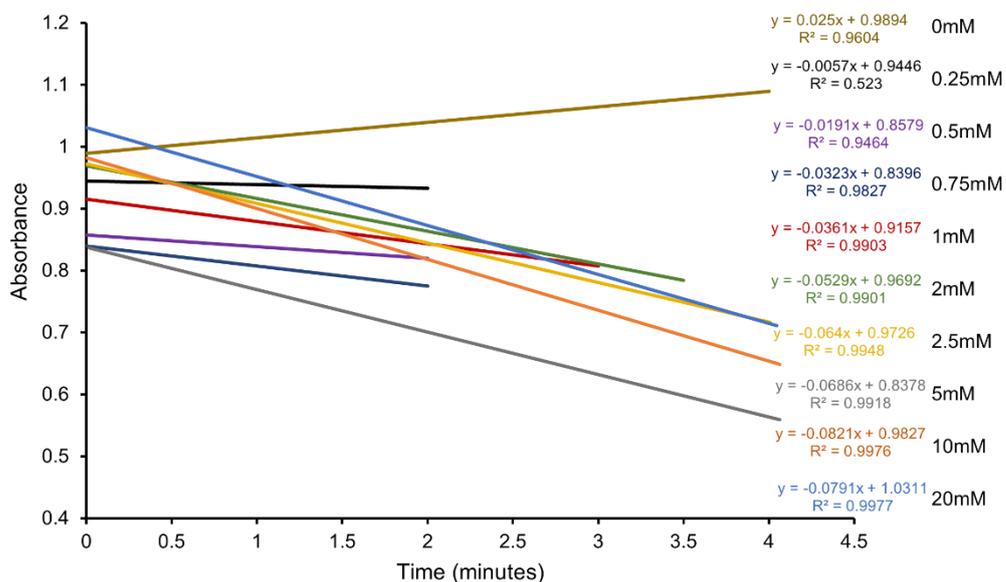


Figure 120: Initial kinetic gradients for MATOUAmDH2 in the pET22b(+) vector. Initial gradients used to calculate velocities for kinetic parameter determination at 50°C. The x-axis represents time in minutes where data was cut off for processing. The y-axis represents the absorbance recordings at 340nm. Trendlines and equations are labelled by colour for respective concentrations of cyclohexanone. Note that only one repeat is displayed for clarity of depiction.

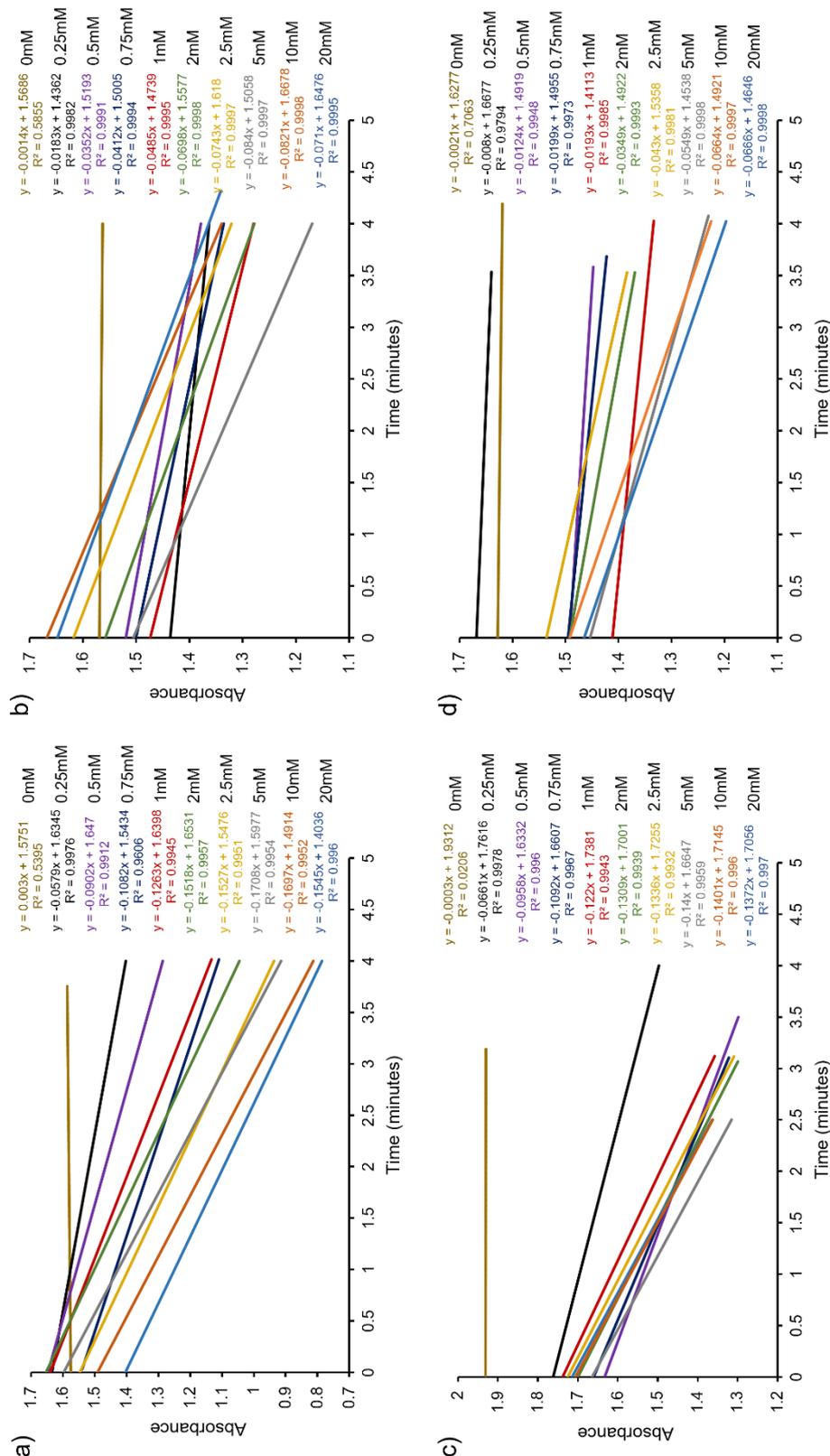


Figure 121: Initial kinetic gradients for MATOUAmDH2 alanine screen mutants. Initial gradients used to calculate velocities for kinetic parameter determination at 25°C. The x-axis represents time in minutes where data was cut off for processing. The y-axis represents the absorbance recordings at 340nm. Trendlines and equations are labelled by colour for respective concentrations of cyclohexanone. Note that only one repeat is displayed for clarity of depiction. a) MATOUAmDH2:YSBLIC-3C WT, b) L180A, c) M215A, d) T312A.

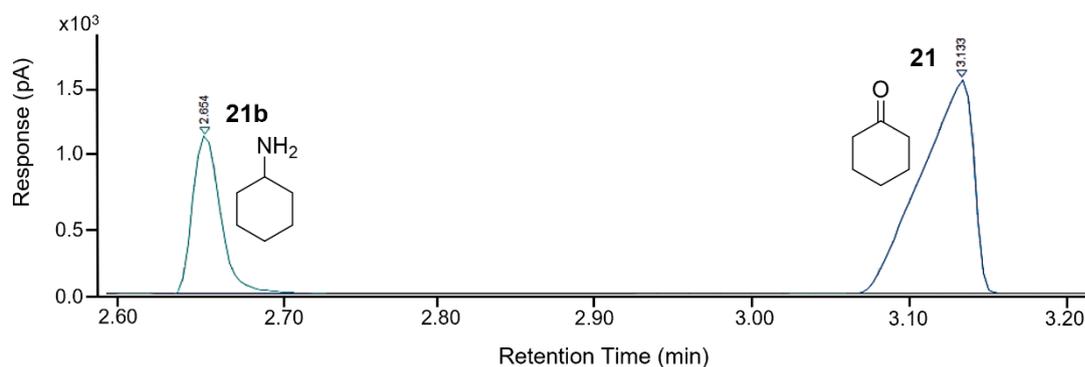


Figure 122: GC-FID traces for pET22b(+) cyclohexanone and cyclohexylamine standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **21** and **21b** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

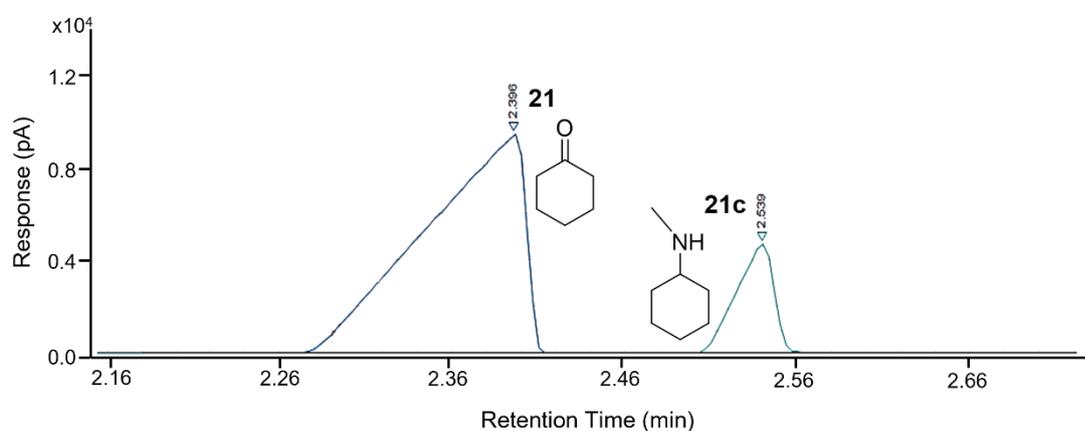


Figure 123: GC-FID traces for pET22b(+) cyclohexanone and N-methylcyclohexylamine standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **21** and **21c** used in methylamine-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

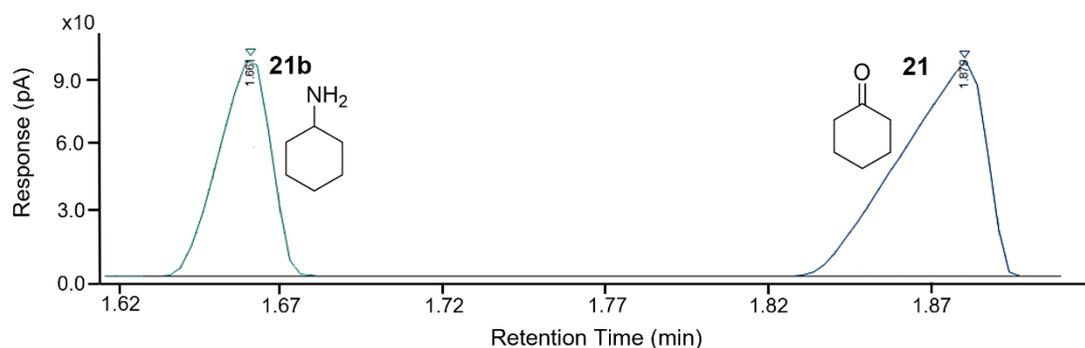


Figure 124: GC-FID traces for pETYSBLIC-3C cyclohexanone and cyclohexylamine standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **21** and **21b** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

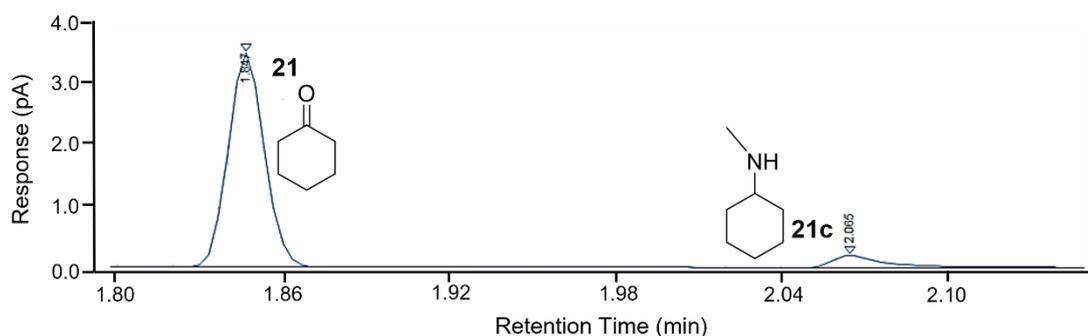


Figure 125: GC-FID traces for pETYSBLIC-3C cyclohexanone and N-methylcyclohexylamine standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **21** and **21c** used in methylamine-based reactions, standards as from an exemplar enzymatic based run and are appropriated from Figure 124 **21** and distance in **21** and **21c** from Figure 123. Retention times linked to chemical standards are labelled and tabulated accordingly.

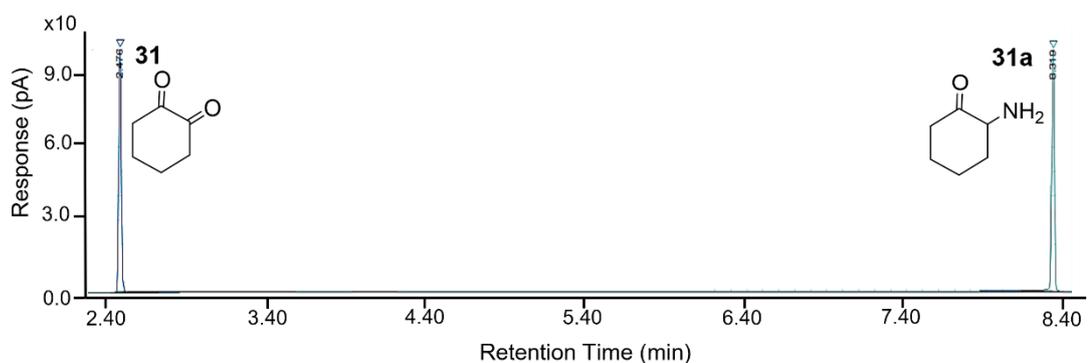


Figure 126: GC-FID traces for 1,2-cyclohexanedione and 2-aminocyclohexanone standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **31** and **31a** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

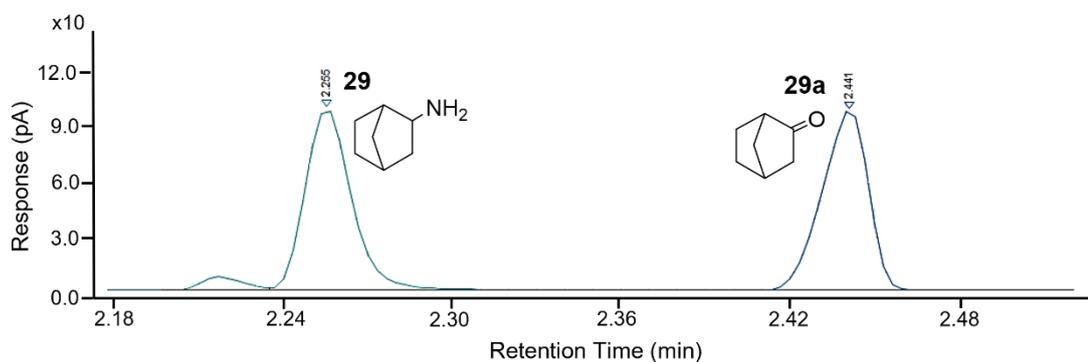


Figure 127: GC-FID traces for norcamphor and 2-aminonorbornane standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **29** and **29a** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

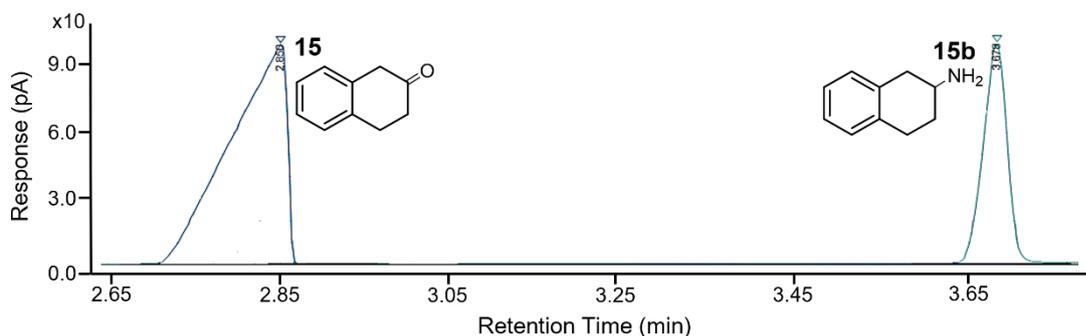


Figure 128: GC-FID traces for 2-teralone and 2-aminotetralin standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **15** and **15b** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

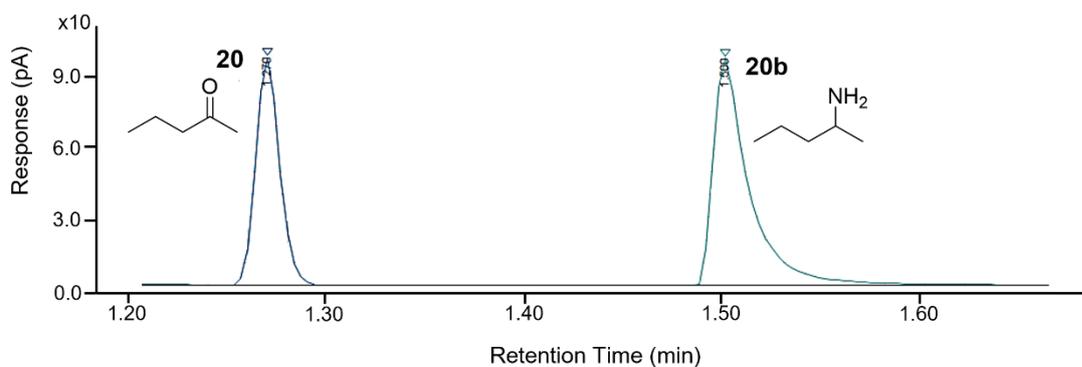


Figure 129: GC-FID traces for 2-pentanone and 2-aminopentane standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **20** and **20b** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

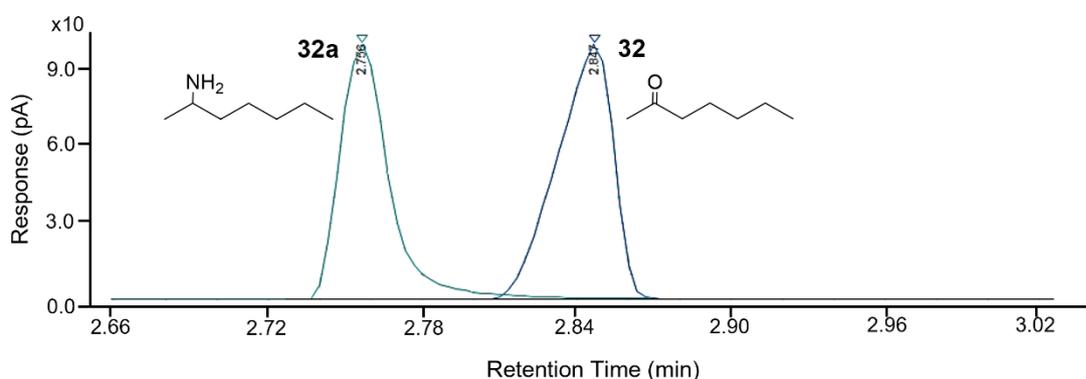


Figure 130: GC-FID traces for 2-heptanone and 2-aminoheptane standards. The x-axis represents retention times in min. The y-axis represents the response current in pA. Standards for **32** and **32a** used in ammonia-based reactions. Retention times linked to chemical standards are labelled and tabulated accordingly.

	Score	Expect	Method	Identities	Positives	Gaps
	172 bits(435)	3e-50	Compositional matrix adjust.	114/334(34%)	167/334(50%)	11/334(3%)
MATOUAmDH2	9		PLRVALYGFQNGQKEMAKMLVERKDVEIVAVIS-NKSNVKGDFGEVIGLAPQ-GILVTAG	66		
<i>CfusAmDH2</i>	9		P+R+ +G G + + + L E K E+V I N + +D GEV GL G+ +			67
MATOUAmDH2	67		LDAAETLRTTNPQIAMLSTLSTVGDIESQLRACAENKVNVTIAEELTFSWTSAPKTK	126		
<i>CfusAmDH2</i>	68		A R + +L T S + ++ LR A + VN TI EE + WT++ T+E			127
MATOUAmDH2	127		MDELFEKHNVSILGGGFLDACCDMARTMAAMMHKIDKLDGGLQYNVDHYGQVLAIAHGV	186		
<i>CfusAmDH2</i>	128		+D+L + ++ +L G GF D ++ +A H+ID++ G QYN D YG LA HG			187
MATOUAmDH2	187		GLSEEFYAENGPWTSPTSYPKSYVYNNMDFASAFGLTVIKTEEVKTPTKAPIELYSE	246		
<i>CfusAmDH2</i>	188		GLDPETFAARIG-----ASNPSYVWNSNEWLCAQLGWRVDIRQQLLPTHTTGLRSA	241		
MATOUAmDH2	247		AIGRAIPVGCtgmivtattttteegviiveKQVGKCYEDGEDMVFMMLEGNPTGGVGF	306		
<i>CfusAmDH2</i>	242		++GR +P G TGM T T EG +I V GK Y G+ D+ L G P V T			299
MATOUAmDH2	307		MKNPPTPAMNTIAISRMFQTVDPAGYITTDKL	340		
<i>CfusAmDH2</i>	300		++ P TPA+T ++R+ Q + AP G++TTD+			333
			IRQPATPALTCATVNLRLPQLLAAPPGFVITDRF			

Figure 131: Sequence alignment of MATOUAmDH2 and *CfusAmDH*. A protein:protein BLAST sequence alignment using the Protein Data Bank proteins (PDB) search specifier.

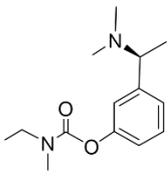
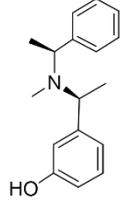
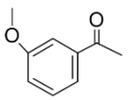
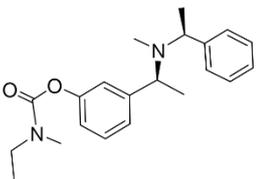
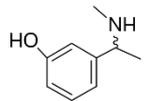
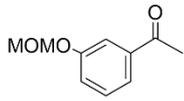
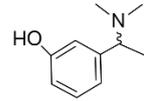
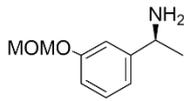
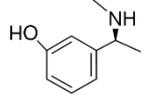
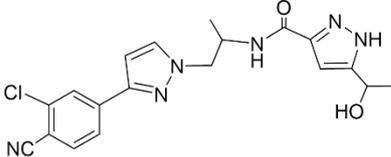
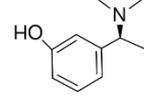
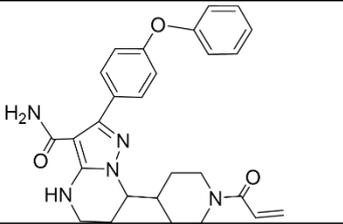
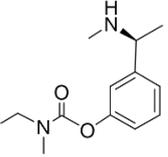
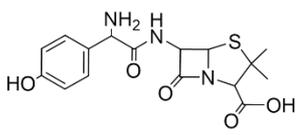
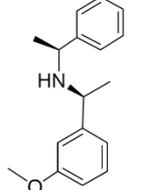
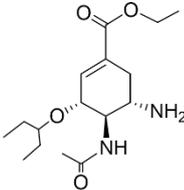
	Score	Expect	Method	Identities	Positives	Gaps
	166 bits(419)	7e-48	Compositional matrix adjust.	116/344(34%)	174/344(50%)	12/344(3%)
MATOUAmDH2	3		HHHHHSPLRVALYGFQNGQKEMAKMLVERKDVEIVAVISNK-SNVGKDFGEVIGLAPQ-G	60		
<i>MsmAmDH2</i>	1		HHHHHS +R +YG G N +A ML++ K V+IV I+ VG+D G+++GL Q G			59
MATOUAmDH2	61		ILVTAGLDAAEELRRTTNPQIAMLSTLSTVGDIESQLRACAENKVNVTIAEELTFSWTS	120		
<i>MsmAmDH2</i>	60		+ V+ DAAE L T+P IA+++ S + D QLR CAE+ VN T+EE+ + W ++			117
MATOUAmDH2	121		PEKTKEMDELFEKHNVSILGGGFLDACCDMARTMAAMMHKIDKLDGGLQYNVDHYGQV	180		
<i>MsmAmDH2</i>	118		PE + E+D L K +L G G+ D +M + H+ID + G +NVD +G L			177
MATOUAmDH2	181		PEKTKEMDELFEKHNVSILGGGFLDACCDMARTMAAMMHKIDKLDGGLQYNVDHYGQV	240		
<i>MsmAmDH2</i>	178		PE + E+D L K +L G G+ D +M + H+ID + G +NVD +G L			232
MATOUAmDH2	241		AIAHGVGLSEEFYAENGPWTSPTSYPKSYVYNNMDFASAFGLTVIKTEEVKTPTKAP	300		
<i>MsmAmDH2</i>	233		A A VG + EF W P ++ N+ D + GLTV P A			292
MATOUAmDH2	301		ATAQQVGRVVAEFDE----WVRGAQRPPFTFGRNVLDALVADTGLTVKSIITATRPDIAS	344		
<i>MsmAmDH2</i>	293		I ELYSEAIGRAIPVGCtgmivtattttteegviiveKQVGKCYEDGEDMVFMMLEGNPT			334
			+ SEA+G + G G TEEG + + G+ Y G+ D+ +EG P			
			AAMRSEALGIDLAPGDVIGFTDIDRIETEEGPVFEFEMSGRVYGPGEIDINWETIEGEPN			
MATOUAmDH2	301		GGVGFMTKNPPTPAMNTIAISRMFQTVDPAGYITTDKLP	344		
<i>MsmAmDH2</i>	293		++ + PT T T ++R+ + AP G +T D+LP +			334
			--LFLSNGTVPTQTTTCTQMVRIPDVIAAPPGIVTVDRLPRLR			

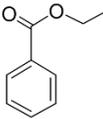
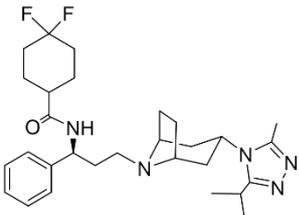
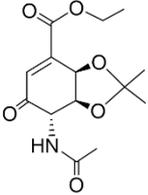
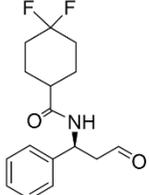
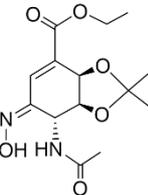
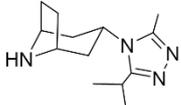
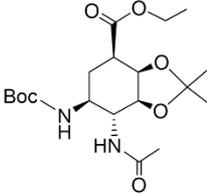
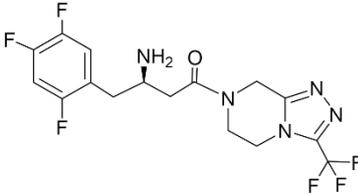
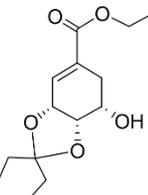
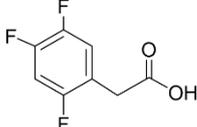
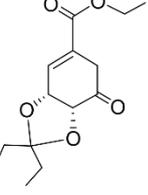
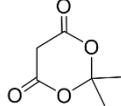
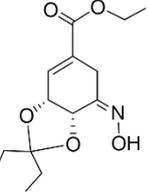
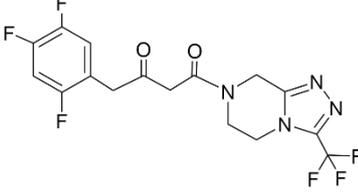
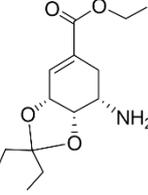
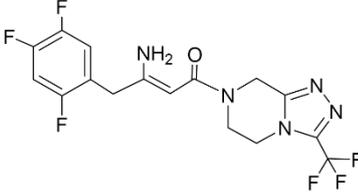
Figure 132: Sequence alignment of MATOUAmDH2 and *MsmAmDH*. A protein:protein BLAST sequence alignment using the Protein Data Bank proteins (PDB) search specifier.

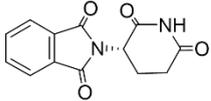
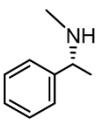
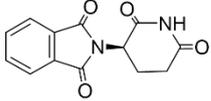
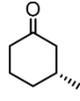
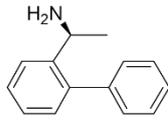
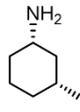
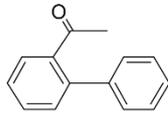
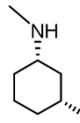
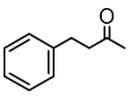
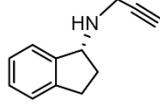
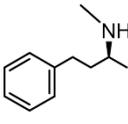
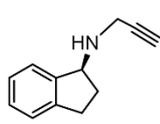
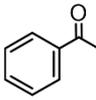
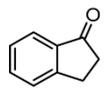
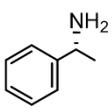
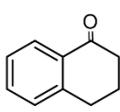
	Score	Expect	Method	Identities	Positives	Gaps
	367 bits(941)	2e-126	Compositional matrix adjust.	181/336(54%)	236/336(70%)	0/336(0%)
<i>TtherAmDH</i>	5		KVVVWNLGAMGSGIAKMILFKKMEIVGAIDTPNKRKDLNEILGTNSKPVYITSEPQD	64		
<i>AmDH4</i>	3		KV +WG GAMGSGIAK IL KK +++VG D KD+ E+LG + + +P			62
<i>TtherAmDH</i>	65		KVGIWGFAMGSGIAKNILSKKNLKLGVVHDFREEYIEKDVGELLGLGKIGIKVYDPDIT	124		
<i>AmDH4</i>	63		I I K K G S A D I A V I T S S Y V E K V F L I K L A V E N G I N V I T T A E E M A Y P S A Q H L E L A K E I D R L A			122
<i>TtherAmDH</i>	125		++K+ D+ VI T+S++ V I ++ NVIT AEEMA+P ++ + A EID +A			184
<i>AmDH4</i>	123		MVKQTPDPLVVIATNSFISVVKDQIISILKENKNVITIAEEMAFPFKDPKAANEIDTVA	182		
<i>TtherAmDH</i>	185		RENGVSVLGTGINPGFVLDYLIILALTGVCVDVDSIKAARINDLSPFGKAVMEEQGVGLTP	244		
<i>AmDH4</i>	183		+++ VSVLGTG+NPGFVLD LII LTG+C++V IKAARINDLSPFG VME QGVG TP			242
<i>TtherAmDH</i>	245		KDHNVSVLGTGVNPGFVLDLTIITLTGICLNVRKAARINDLSPFGPTVMEQGVGTTP	304		
<i>AmDH4</i>	243		EEFEKQIKSGKIVGHIGFEQSIHMIKALGWIEDRIEQKREPIISNMVRETKVYVQVQGM	302		
<i>TtherAmDH</i>	245		EEFEKQIKSGKIVGHIGFEQSIHMIKALGWIEDRIEQKREPIISNMVRETKVYVQVQGM			304
<i>AmDH4</i>	243		VAGCRQIGYKVDGEVKIELEHPQQLPQKEGVETGDYIEIKGTPNIKLSIKPEIPGGLG			302
<i>TtherAmDH</i>	305		VAGC VAGCNHTAKAFYKNELLIELEHPQQLPHELEHPQQLPHELEHPQQLPHELEHPQQLP	340		
<i>AmDH4</i>	303		TIAIATNMIPSVVEARPGLLTMVDPVPRALLAEVH	338		
			TIA+ NMIP V+ A PGL+TM+DLP+PRA++ +			

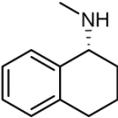
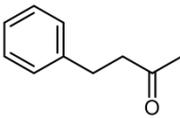
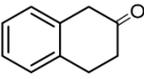
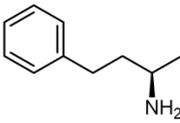
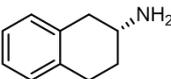
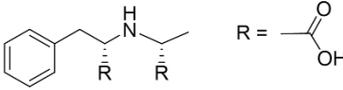
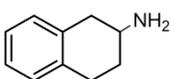
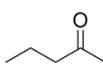
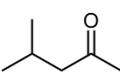
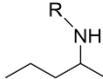
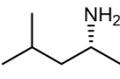
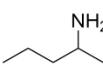
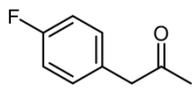
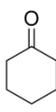
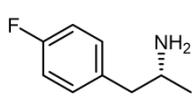
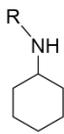
Figure 133: Sequence alignment of *TtherAmDH* and *AmDH4*. A protein:protein BLAST sequence alignment using the Protein Data Bank proteins (PDB) search specifier.

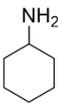
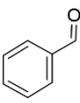
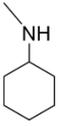
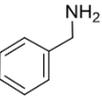
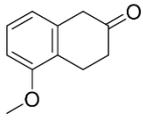
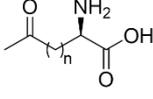
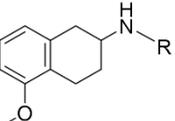
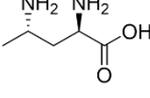
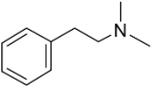
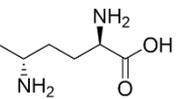
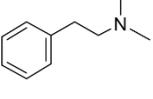
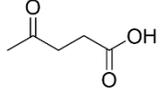
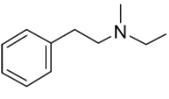
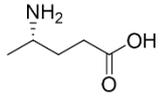
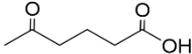
Chemical Glossary

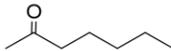
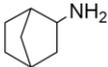
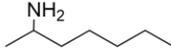
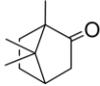
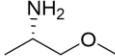
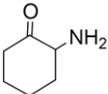
1		1f	
1a		1g	
1.1b		1h	
1.2b		1i	
1.1c		2	
1.2c		3	
1d		4	
1e		5	

5a		6	
5b		6a	
5c		6b	
5d		7	
5e		7a	
5f		7b	
5g		7c	
5h		7d	

8.1		11b	
8.2		12	
9		12a	
9a		12b	
10		13.1	
10a		13.2	
11		13a	
11a		14	

14a		18	
15		18a	
15a		19	
15b		20	
16		20a	
16a		20b	
17		21	
17a		21a	

21b		24a	
21c		24b	
22		25	
22a		25a	
23		25b	
23a		26	
23b		26a	
24		27	

28		32	
29		32a	
29a		33	
29b		34	
30		35	
31		36	
31a			

Abbreviations

1,3-DMBA - 1,3-dimethylbutylamine

2,4-DAP - 2,4-diaminopentanoate

2,4-DAPDH – 2,4-diaminopentanoate dehydrogenase

2,5-DAH - 2,5-diaminohexanoate

Å –Ångstroms

A_{280} – Absorbance at 280 nm

AADH – Amino Acid Dehydrogenase

ADH – Alcohol Dehydrogenase

AIDS – Acquired Immune Deficiency Syndrome

AmDH – Amine Dehydrogenase

bp – Base pair

CSS – Clear Strategy Screen™

d – Days

DMSO - dimethylsulfoxide

DNA – Deoxyribonucleic acid

DPP-4 - Protease Dipeptidyl Peptidase 4

E. coli – *Escherichia coli*

e.e. – Enantiomeric excess

F_o - F_c – Observed structure factors - calculated structure factors

g – Grams

GC-FID – Gas Chromatography equipped with a Flame Ionised Detector

GDH – Glucose Dehydrogenase

G6P-DH – Glucose-6-phosphate Dehydrogenase

GLP-1 - Glucagon-like Peptide 1

h – Hours

HIV – Human Immunodeficiency Virus

IPTG – Isopropyl- β -D-1-thiogalactopyranoside

IRED – Imine Reductase

k_{cat} – Turnover number

kDa – Kilodalton

K_m – Michaelis-Menten constant

LDH – Lactate Dehydrogenase

LIC – Ligation Independent Cloning

M – Molar

MATOU – Marine Atlas of Tara Oceans Unigenes

MIBK - methyl isobutylketone

mg – milligrams

μ L – Microlitres

μ m – Micrometres

μ M – Micromolar

mM – Millimolar

MPD - 2-methyl-2,4-pentanediol

NADH - Nicotinamide adenine dinucleotide

NADPH - Nicotinamide adenine dinucleotide phosphate

Nat-AmDH – Native Amine Dehydrogenase

Ni – Nickel

NiNTA – Nickel Affinity Chromatography

nm – Nanometres

NTA - nitrilotriacetic acid

OD₆₀₀ – Optical Density at 600 nm

OpDH – Opine Dehydrogenase

PCR – Polymerase Chain Reaction

PEG - Polyethylene glycol

PFPA - *p*-fluorophenylacetone

PLP - pyridoxal 5'-phosphate

PMP - pyridoxamine 5'-phosphate

R_{crystal} - Residual factor

RedAm – Reductive Aminase

R_{free} - Free R-Factor

RMSD - Root Mean Square Deviation

rpm – Revolutions per minute

s – Seconds

SDS - Sodium dodecyl sulfate

SDS-PAGE - Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC – Size Exclusion Chromatography

σ – Sigma

(S)-MOIPA - (S)-1-methoxypropan-2-amine

TAE – Tris-acetate-EDTA

t-Bu – Tertiary Butyl

T_m – Melting Temperature

UV-vis - Ultraviolet–visible spectroscopy

V_{max} - maximum velocity

(w/v) – Weight per volume

References

- 1 M. Breuer, K. Ditrich, T. Habicher, B. Hauer, M. Keßeler, R. Stürmer and T. Zelinski, *Angew. Chemie - Int. Ed.*, 2004, **43**, 788–824.
- 2 N. J. Turner, *Nat. Chem. Biol.*, 2009, **5**, 567–573.
- 3 H. E. Schoemaker, D. L. Mink and M. G. WubboLts, *Science*, 2003, **299**, 1694–1697.
- 4 L. Ducrot, M. Bennett, A. A. Caparco, J. A. Champion, A. S. Bommarius, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire, *Front. Catal.*, 2021, **1**, DOI:10.3389/fctls.2021.781284.
- 5 Y. Hsiao, N. R. Rivera, T. Rosner, S. W. Krska, E. Njolito, F. Wang, Y. Sun, J. D. Armstrong, E. J. J. Grabowski, R. D. Tillyer, F. Spindler and C. Malan, *J. Am. Chem. Soc.*, 2004, **126**, 9918–9919.
- 6 P. D. Bailey, J. Clayson and A. N. Boa, *Contemp. Org. Synth.*, 1995, **2**, 173–187.
- 7 N. Uematsu, A. Fujii, S. Hashiguchi, T. Ikariya and R. Noyori, 1996, **7863**, 4916–4917.
- 8 C. G. Espino, K. W. Fiori, M. Kim and J. Du Bois, *J. Am. Chem. Soc.*, 2004, **126**, 15378–15379.
- 9 O. I. Afanasyev, E. Kuchuk, D. L. Usanov and D. Chusov, *Chem. Rev.*, 2019, **119**, 11857–11911.
- 10 S. D. Roughley and A. M. Jordan, *J. Med. Chem.*, 2011, **54**, 3451–3479.
- 11 S. Bhattacharyya, *J. Org. Chem.*, 1995, **88**, 4928–4929.
- 12 S. Stepankova, H.; Hajicek, J.; Simek, *United States Pat. Appl. Publ. WO2004037771A1*, 2006, 1–8.
- 13 K. B. Hansen, Y. Hsiao, F. Xu, N. Rivera, A. Clausen, M. Kubryk, S. Krska, T. Rosner, B. Simmons, J. Balsells, N. Ikemoto, Y. Sun, F. Spindler, C. Malan, E. J. J. Grabowski and J. D. Armstrong, *J. Am. Chem. Soc.*, 2009, **131**, 8798–8804.
- 14 C. Wang and J. Xiao, *Top. Curr. Chem.*, 2014, **343**, 261–282.

- 15 J. J. Murillo, G. J. V. R. S. L.; Armengol, M. M. I. S. A.; Martín and R. S. L., *Eur. Pat. Appl. EP1939172B1*, 2008, 1–24.
- 16 R. . Zhang, F.; Hu, M.; Xie, M.; Lai, A.; Sun, R.; Chen, D.; Bao and H. Bai, *Eur. Pat. Appl. EP2233465A1*, 2010, 1–17.
- 17 M. Hu, F. L. Zhang and M. H. Xie, *Synth. Commun.*, 2009, **39**, 1527–1533.
- 18 L. Werner, A. Machara and T. Hudlicky, *Adv. Synth. Catal.*, 2010, **352**, 195–200.
- 19 N. Chuanopparat, N. Kongkathip and B. Kongkathip, *Tetrahedron*, 2012, **68**, 6803–6809.
- 20 P. Wichienukul, S. Akkarasamiyo, N. Kongkathip and B. Kongkathip, *Tetrahedron Lett.*, 2010, **51**, 3208–3210.
- 21 S. J. Haycock-Lewandowski, A. Wilder and J. Åhman, *Org. Process Res. Dev.*, 2008, **12**, 1094–1103.
- 22 S. Lou, P. N. Moquist and S. E. Schaus, *J. Am. Chem. Soc.*, 2007, **129**, 15398–15404.
- 23 N. Fleury-Brégeot, V. de la Fuente, S. Castellón and C. Claver, *ChemCatChem*, 2010, **2**, 1346–1371.
- 24 A. S. Bommarius, M. Schwarm, K. Stingl, M. Kottenhahn, K. Huthmacher and K. Drauz, *Tetrahedron: Asymmetry*, 1995, **6**, 2851–2888.
- 25 T. C. Nugent, A. K. Ghosh, V. N. Wakchaure and R. R. Mohanty, *Adv. Synth. Catal.*, 2006, **348**, 1289–1299.
- 26 L. Dalko, Peter I., Moisan, *Angew. Chemie - Int. Ed.*, 2001, **12**, 8–27.
- 27 S. Simon, S. Oßwald, J. Roos and H. Gröger, *Zeitschrift fur Naturforsch. - Sect. B J. Chem. Sci.*, 2012, **67**, 1123–1126.
- 28 R. J. Polinsky, *Clin. Ther.*, 1998, **20**, 634–647.
- 29 E. Van Der Ryst, *Front. Immunol.*, 2015, **6**, 1–4.
- 30 T. Biftu, D. Feng, X. Qian, G. B. Liang, G. Kieczkowski, G. Eiermann, H. He, B. Leiting, K. Lyons, A. Petrov, R. Sinha-Roy, B. Zhang, G. Scapin, S. Patel, Y. D. Gao, S. Singh, J. Wu, X. Zhang, N. A. Thornberry and A. E. Weber, *Bioorganic Med. Chem. Lett.*, 2007, **17**, 49–52.

- 31 A. A. Desai, *Angew. Chemie - Int. Ed.*, 2011, **50**, 1974–1976.
- 32 J. Tang, J. Majeti, A. Sudom, Y. Xiong, M. Lu, Q. Liu, J. Higbee, Y. Zhang, Y. Wang, W. Wang, P. Cao, Z. Xia, S. Johnstone, X. Min, X. Yang, H. Shao, T. Yu, N. Sharkov, N. Walker, H. Tu, W. Shen and Z. Wang, *J. Biol. Chem.*, 2013, **288**, 1307–1316.
- 33 D. Buonsenso, D. Serranti and P. Valentini, *Eur. Rev. Med. Pharmacol. Sci.*, 2010, **14**, 845–853.
- 34 S. W. Smith, *Toxicol. Sci.*, 2009, **110**, 4–30.
- 35 T. Eriksson, S. Björkman, B. Roth and P. Höglund, *J. Pharm. Pharmacol.*, 2010, **52**, 807–817.
- 36 T. Eriksson, S. Björkman and P. Höglund, *Eur. J. Clin. Pharmacol.*, 2001, **57**, 365–376.
- 37 W. Heger, H. -J Schmahl, S. Klug, A. Felies, H. Nau, H. -J Merker and D. Neubert, *Teratog. Carcinog. Mutagen.*, 1994, **14**, 115–122.
- 38 A. Schmid, J. S. Doderick, B. Hauer, A. Kiener, M. Wubbolts and B. Witholt, *Nature*, 2001, **409**, 258–268.
- 39 L. Rubio-Pérez, F. Javier Pérez-Flores, P. Sharma, L. Velasco and A. Cabrera, *Org. Lett.*, 2009, **11**, 265–268.
- 40 S. Itsuno, Y. Hashimoto and N. Haraguchi, *J. Polym. Sci. Part A Polym. Chem.*, 2014, **52**, 3037–3044.
- 41 J. R. Marshall, J. Mangas-Sanchez and N. J. Turner, *Tetrahedron*, 2021, **82**, 131926.
- 42 B. A. Sandoval and T. K. Hyster, *Curr. Opin. Chem. Biol.*, 2020, **55**, 45–51.
- 43 B. Testa, *Curr. Opin. Chem. Biol.*, 2009, **13**, 338–344.
- 44 Y. Yang, H. Aloysius, D. Inoyama, Y. Chen and L. Hu, *Acta Pharm. Sin. B*, 2011, **1**, 143–159.
- 45 M. Markovic, S. Ben-Shabat and A. Dahan, *Pharmaceutics*, 2020, **12**, 1–12.
- 46 N. A. Roberts, J. A. Martin, D. Kinchington, A. V. Broadhurst, J. C. Craig, I. B. Duncan, S. A. Galpin, B. K. Handa, J. Kay, A. Kröhn, R. W. Lambert, J. H. Merrett, J. S. Mills, K. E. B. Parkes, S. Redshaw, A. J. Ritchie, D. L. Taylor, G.

- J. Thomas and P. J. Machin, *Science*, 1990, **248**, 358–361.
- 47 J. G. Robertson, *Biochemistry*, 2005, **44**, 8918–8918.
- 48 G. A. Holdgate, T. D. Meek and R. L. Grimley, *Nat. Rev. Drug Discov.*, 2018, **17**, 115–132.
- 49 T. D. H. Bugg, *Introduction to Enzyme and Coenzyme Chemistry: Third Edition*, Wiley, Chichester, 2012.
- 50 S. A. Kelly, S. Pohle, S. Wharry, S. Mix, C. C. R. Allen, T. S. Moody and B. F. Gilmore, *Chem. Rev.*, 2018, **118**, 349–367.
- 51 M. Höhne and U. T. Bornscheuer, *ChemCatChem*, 2009, **1**, 42–51.
- 52 W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz and J. Brands, *Science*, 2010, **329**, 305–310.
- 53 A. Iwasaki, Y. Yamada, N. Kizaki, Y. Ikenaka and J. Hasegawa, *Appl. Microbiol. Biotechnol.*, 2006, **69**, 499–505.
- 54 J. S. Shin and B. G. Kim, *J. Org. Chem.*, 2002, **67**, 2848–2853.
- 55 M. Höhne, S. Kühn, K. Robins and U. T. Bornscheuer, *ChemBioChem*, 2008, **9**, 363–365.
- 56 D. Koszelewski, I. Lavandera, D. Clay, D. Rozzell and W. Kroutil, *Adv. Synth. Catal.*, 2008, **350**, 2761–2766.
- 57 M. Fuchs, D. Koszelewski, K. Tauber, W. Kroutil and K. Faber, *Chem. Commun.*, 2010, **46**, 5500–5502.
- 58 D. F. A. R. Dourado, S. Pohle, A. T. P. Carvalho, D. S. Dheeman, J. M. Caswell, T. Skvortsov, I. Miskelly, R. T. Brown, D. J. Quinn, C. C. R. Allen, L. Kulakov, M. Huang and T. S. Moody, *ACS Catal.*, 2016, **6**, 7749–7759.
- 59 J. Mangas-Sanchez, S. P. France, S. L. Montgomery, G. A. Aleku, H. Man, M. Sharma, J. I. Ramsden, G. Grogan and N. J. Turner, *Curr. Opin. Chem. Biol.*, 2017, **37**, 19–25.
- 60 K. Mitsukura, M. Suzuki, S. Shinoda, T. Kuramoto, T. Yoshida and T. Nagasawa, *Biosci. Biotechnol. Biochem.*, 2011, **75**, 1778–1782.
- 61 H. Man, E. Wells, S. Hussain, F. Leipold, S. Hart, J. P. Turkenburg, N. J. Turner and G. Grogan, *ChemBioChem*, 2015, **16**, 1052–1059.

- 62 D. Wetzl, M. Berrera, N. Sandon, D. Fishlock, M. Ebeling, M. Müller, S. Hanlon, B. Wirz and H. Iding, *ChemBioChem*, 2015, **16**, 1749–1756.
- 63 T. Huber, L. Schneider, A. Präg, S. Gerhardt, O. Einsle and M. Müller, *ChemCatChem*, 2014, **6**, 2248–2252.
- 64 P. N. Scheller, M. Lenz, S. C. Hammer, B. Hauer and B. M. Nestl, *ChemCatChem*, 2015, **7**, 3239–3242.
- 65 D. Wetzl, M. Gand, A. Ross, H. Müller, P. Matzel, S. P. Hanlon, M. Müller, B. Wirz, M. Höhne and H. Iding, *ChemCatChem*, 2016, **8**, 2023–2026.
- 66 P. Matzel, M. Gand and M. Höhne, *Green Chem.*, 2017, **19**, 385–389.
- 67 G. A. Aleku, S. P. France, H. Man, J. Mangas-Sanchez, S. L. Montgomery, M. Sharma, F. Leipold, S. Hussain, G. Grogan and N. J. Turner, *Nat. Chem.*, 2017, **9**, 961–969.
- 68 G. A. Aleku, H. Man, S. P. France, F. Leipold, S. Hussain, L. Toca-Gonzalez, R. Marchington, S. Hart, J. P. Turkenburg, G. Grogan and N. J. Turner, *ACS Catal.*, 2016, **6**, 3880–3889.
- 69 M. Rodríguez-Mata, A. Frank, E. Wells, F. Leipold, N. J. Turner, S. Hart, J. P. Turkenburg and G. Grogan, *ChemBioChem*, 2013, **14**, 1372–1379.
- 70 N. K. Lokanath, N. Ohshima, K. Takio, I. Shiromizu, C. Kuroishi, N. Okazaki, S. Kuramitsu, S. Yokoyama, M. Miyano and N. Kunishima, *J. Mol. Biol.*, 2005, **352**, 905–917.
- 71 L. Ducrot, M. Bennett, G. Grogan and C. Vergne-Vaxelaire, *Adv. Synth. Catal.*, 2020, **363**, 328–351.
- 72 M. Sharma, J. Mangas-Sanchez, N. J. Turner and G. Grogan, *Adv. Synth. Catal.*, 2017, **359**, 2011–2025.
- 73 V. Tseliou, D. Schilder, M. F. Masman, T. Knaus and F. G. Mutti, *Chem. - A Eur. J.*, 2021, **27**, 3315–3325.
- 74 V. Tseliou, T. Knaus, M. F. Masman, M. L. Corrado and F. G. Mutti, *Nat. Commun.*, 2019, DOI:10.1038/s41467-019-11509-x.
- 75 T. Knaus, W. Böhmer and F. G. Mutti, *Green Chem.*, 2017, **19**, 453–463.
- 76 V. Tseliou, M. F. Masman, W. Böhmer, T. Knaus and F. G. Mutti, *ChemBioChem*, 2019, **20**, 800–812.

- 77 Y. Asano, K. Yamaguchi and K. Kondo, *J. Bacteriol.*, 1989, **171**, 4466–4471.
- 78 P. J. Baker, A. P. Turnbull, S. E. Sedelnikova, T. J. Stillman and D. W. Rice, *Structure*, 1995, **3**, 693–705.
- 79 T. Sekimoto, T. Matsuyama, T. Fukui and K. Tanizawa, *J. Biol. Chem.*, 1993, **268**, 27039–27045.
- 80 M. J. Abrahamson, J. W. Wong and A. S. Bommarius, *Adv. Synth. Catal.*, 2013, **355**, 1780–1786.
- 81 M. J. Abrahamson, E. Vázquez-Figueroa, N. B. Woodall, J. C. Moore and A. S. Bommarius, *Angew. Chemie - Int. Ed.*, 2012, **51**, 3969–3972.
- 82 N. M. W. Brunhuber, J. B. Thoden, J. S. Blanchard and J. L. Vanhooke, *Biochemistry*, 2000, **39**, 9174–9187.
- 83 L. J. Ye, H. H. Toh, Y. Yang, J. P. Adams, R. Snajdrova and Z. Li, *ACS Catal.*, 2015, **5**, 1119–1122.
- 84 D. J. Wilkinson, T. Hossain, D. S. Hill, B. E. Phillips, H. Crossland, J. Williams, P. Loughna, T. A. Churchward-Venne, L. Breen, S. M. Phillips, T. Etheridge, J. A. Rathmacher, K. Smith, N. J. Szewczyk and P. J. Atherton, *J. Physiol.*, 2013, **591**, 2911–2923.
- 85 V. R. Silva, F. L. Belozo, T. O. Micheletti, M. Conrado, J. R. Stout, G. D. Pimentel and A. M. Gonzalez, *Nutr. Res.*, 2017, **45**, 1–9.
- 86 S. H. J. Smits, T. Meyer, A. Mueller, N. van Os, M. Stoldt, D. Willbold, L. Schmitt and M. K. Grieshaber, *PLoS One*, 2010, DOI:10.1371/journal.pone.0012312.
- 87 S. N. H. Chen, J. Moore, S. J. Collier, D. Smith, J. Nazor, G. Hughes, J. Janey, G. Huisman, S. Novick, N. Agard, O. Alvizo, G. Cope, W-L. Yeo, J. Sukumaran, *U.S. Pat. Appl. 20130302859*.
- 88 C. V. V. Anne Zaparucha, Véronique de Berardinis, *Modern Biocatalysis: Advances Towards Synthetic Biological Systems, Chapter I: Genome Mining for Enzyme Discovery*, Royal Society of Chemistry, 2018.
- 89 O. Mayol, S. David, E. Darii, A. Debard, A. Mariage, V. Pellouin, J. L. Petit, M. Salanoubat, V. De Berardinis, A. Zaparucha and C. Vergne-Vaxelaire, *Catal. Sci. Technol.*, 2016, **6**, 7421–7428.
- 90 S. Fukuyama, H. Mihara, R. Miyake, M. Ueda, N. Esaki and T. Kurihara, *J.*

- Biosci. Bioeng.*, 2014, **117**, 551–556.
- 91 O. Mayol, K. Bastard, L. Beloti, A. Frese, J. P. Turkenburg, J. L. Petit, A. Mariage, A. Debard, V. Pellouin, A. Perret, V. De Berardinis, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire, *Nat. Catal.*, 2019, **2**, 324–333.
- 92 L. Ducrot, M. Bennett, G. André-Leroux, E. Elisée, S. Marynberg, A. Fossey-Jouenne, A. Zaparucha, G. Grogan and C. Vergne-Vaxelaire, *ChemCatChem*, 2022, DOI:10.1002/cctc.202200880.
- 93 M. Bennett, L. Ducrot, C. Vergne-Vaxelaire and G. Grogan, *ChemBioChem*, 2022, **23**, 6–11.
- 94 G. Winter, J. Beilsten-Edmands, N. Devenish, M. Gerstel, R. J. Gildea, D. McDonagh, E. Pascal, D. G. Waterman, B. H. Williams and G. Evans, *Protein Sci.*, 2022, **31**, 232–250.
- 95 J. P. Glusker, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1993, **49**, 1–1.
- 96 P. Evans, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2006, **62**, 72–82.
- 97 P. R. Evans and G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 1204–1214.
- 98 A. Vagin and A. Teplyakov, *J. Appl. Crystallogr.*, 1997, **30**, 1022–1025.
- 99 A. Vagin and A. Teplyakov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 22–25.
- 100 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1997, **53**, 240–255.
- 101 A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long and G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2004, **60**, 2184–2195.
- 102 P. Emsley and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2004, **60**, 2126–2132.
- 103 A. A. Caparco, E. Pelletier, J. L. Petit, A. Jouenne, B. R. Bommarius, V. de Berardinis, A. Zaparucha, J. A. Champion, A. S. Bommarius and C. Vergne-Vaxelaire, *Adv. Synth. Catal.*, 2020, **362**, 2427–2436.
- 104 L. Holm, *Nucleic Acids Res.*, 2022, **50**, W210–W215.

