

Imine Reductases (IREDs): Structure, Function and Mechanism

Henry Wing-Hong Man

PhD

University of York

Chemistry

October 2016

Abstract

Current abiotic synthetic methods of obtaining chiral amines often involve conditions that are not considered environmentally friendly. One of the biggest concerns for chemists is the diminishing supply of precious metals. Imine reductases (IREDs) and reductive aminases (RedAms) are biocatalysts that have the potential for replacing abiotic synthetic routes for obtaining chiral amines. IREDs are useful biocatalysts for the reduction of preformed imines, whereas, RedAms are capable of catalysing reductive aminations with very broad carbonyl and amine substrate spectrum, utilising a carbonyl-to-amine ratio of 1:1. In this thesis, IREDs and RedAms have been characterised using single crystal X-ray crystallography, substrate profiling with biotransformations and kinetic assays.

IREDs from *Bacillus cereus* BAG3X2 (*BcIRED*) and *Nocardiopsis halophila* (*NhIRED*) were shown to be (*S*)-selective for the substrate 2-methylpyrrolidine. Crystal structures were obtained for a *BcIRED*-NADPH complex and a *NhIRED*-*n*-octyl- β -D-glucoside complex to resolutions of 1.8 Å and 2.4 Å respectively. The models were refined to $R_{factors}/R_{free}$ values of 17.4%/19.7% and 23.0%/26.6% for *BcIRED* and *NhIRED* respectively.

Crystal structures of a series of binary and ternary complexes with the IRED from *Amycolatopsis orientalis* (*AoIRED*) had identified a hydrophobic binding site for the amine (*R*)-methyl-1,2,3,4-tetrahydroisoquinoline. Mutation studies identified Asn171 and Tyr179 to be catalytically relevant. Asn241 was found to play a lesser role in catalysis, but mutations of this residue were used to expand substrate scope.

The crystal structure of the ternary complex of the RedAm from *Aspergillus oryzae* (*AspRedAm*) was determined, permitting the identification of the active site. Mutation studies identified that Asp169 and Tyr177 were key residues required for RedAm activity.

Structural, sequence and phylogenetic tree analysis between IREDs and fungal RedAms have now identified a motif for RedAms, which led to the identification of the bacterial RedAm from *Streptomyces* sp. PRh5 (*StrepRedAm*). *StrepRedAm* was characterised and shown to be capable of catalysing reductive aminations.

Table of Contents

Abstract	3
Table of Contents	5
List of Figures	9
List of Tables	24
Acknowledgements	27
Author's Declaration	29
1. Introduction	33
1.1. Abiotic Chiral Amine Synthesis	33
1.2. Biocatalysts: Enzymes as catalysts	37
1.2.1. Applying Molecular Biology to Biocatalysis	37
1.2.2. Current Biocatalytic Toolbox for Producing Chiral Amines	38
1.3. Imine Reductases (IREDs): Function and Activity	46
1.3.1. Metalloenzymes used for the reduction of imines	46
1.3.2. IREDs with known function	47
1.3.2.1. F420-dependent IREDs	47
1.3.2.2. Ketimine reductase	49
1.3.2.3. Thiazolanyl Imine Reductase	51
1.3.2.4. Dihydrofolate Reductase and Pteridine reductase	53
1.3.2.5. Eukaryotic IRED activity	56
1.3.2.6. Conversion of (<i>S</i>)-reticuline to (<i>R</i>)-reticuline in the opium poppy	57
1.3.2.7. Opine Dehydrogenases	59
1.3.2.8. EC 1.5.1.48 Imine reductases (IREDs): Structure, Function and Activity	60
2. Aim of the project	66

3. Methods	68
3.1. Cloning	68
3.1.1. Primer design	68
3.1.2. Polymerase Chain Reaction (PCR)	68
3.1.3. Touchdown PCR	72
3.1.4. Fragment analysis by agarose gel electrophoresis	73
3.1.5. Ligation independent cloning (LIC) qualified T4 DNA polymerase cloning 74	
3.1.6. In-Fusion® Cloning	77
3.1.7. Site Directed Mutagenesis	79
3.1.8. Bacterial Transformation	81
3.1.9. Double Digest	82
3.2. Gene Expression	83
3.2.1. Protein analysis by sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE) and native-PAGE	83
3.3. Protein Purification	85
3.4. Assays	87
3.4.1. Kinetic Assay by UV-VIS	87
3.4.2. Biotransformation Assay with Substrates of Interest	89
3.5. Protein Crystallography	102
3.5.1. High Throughput Crystal Trial Screening	102
3.5.2. Co-crystallisation	102
3.5.2.1. Incubation with ligand prior to screening	102
3.5.2.2. Ligand Soaking	103
3.5.2.3. Adding the ligand as an additive	104
3.5.3. Crystal Optimisation	105
3.6. Data Collection and Model Building	106

3.6.1.	Radiation damage	106
3.6.2.	Helical data collection	107
3.6.3.	Data processing: Indexing, Integration and Scaling	107
3.6.4.	Structure Solution, Model Building and Refinement	108
4.	Results	110
4.1.	(S)-Selective homologues: <i>Bcl</i>RED and <i>Nh</i>RED	110
4.1.1.	Discussion	127
4.2.	Structure of the IRED from <i>Amycolatopsis orientalis</i>	131
4.2.1.	Discussion	156
4.3.	Structure and mechanism of the reductive aminase (RedAm) from <i>Aspergillus oryzae</i>	161
4.3.1.	Discussion	180
4.4.	The bacterial RedAm from <i>Streptomyces</i> sp. PRh5 (<i>Strep</i>RedAm)	188
4.4.1.	Discussion	199
5.	Conclusion	203
6.	Appendix	205
6.1.	List of Primers	205
	Abbreviations	206
	References	210

List of Figures

- Figure 1 – Generic chiral amine structure with its Cahn-Ingold-Prelog (CIP) priority assigned as $R_1 > N > R_2 > R_3$. Where “>” means greater than.33
- Figure 2 – Examples of chiral amines that are biologically active drugs. Codeine is an over-the-counter drug that is commonly used for pain relief. Sertraline is an antidepressant classed as a selective serotonin reuptake inhibitor. Rasagiline is used in the treatment of Parkinson’s disease and irreversibly inhibits monoamine oxidase-B.33
- Figure 3 – Notable reactions by the Nobel Prize laureates Ryoji Noyori, William Standish Knowles and Karl Barry Sharpless: Noyori asymmetric hydrogenation with the chiral catalyst BINAP, development of Knowles experimental catalyst which substituted achiral phosphine ligands on a Wilkinson’s catalyst with chiral phosphine ligands and the Sharpless epoxidation with the chiral catalyst formed *in situ* with diethyl-tartrate, titanium tetra-isopropoxide and *tert*-butyl hydroperoxide. Figure adapted from Noyori and co-workers (1987),¹ Knowles and co-workers (1973)² and Sharpless and co-workers (1980).³34
- Figure 4 – First generation Grubbs catalyst used for olefin metathesis. The catalyst uses bulky ligands to achieve good selectivity. Figure adapted from Funk and co-workers (2006).⁴35
- Figure 5 – The two most common routes for the abiotic synthesis of chiral amines, asymmetric amination and reductive amination. Both routes require chiral catalysts.35
- Figure 6 – Reduction of imines with trichlorosilane activated with a proline organocatalyst in dichloromethane. Good yields and modest enantiomeric excess are achieved. Work performed by Kanemitsu and co-workers (2001).¹²36
- Figure 7 – Protein fitness landscapes of a wild type protein with different end points created by different evolution strategies: (a) iterative cycles of point mutations, (b) fast evolution with DNA shuffling and (c) saturation mutagenesis where degenerate codons are used to randomise the outcome of mutants. Figure adapted from Dalby (2011).¹⁸38
- Figure 8 – The current biocatalytic toolbox for accessing chiral amines with hydrolases, monoamine oxidases, transaminases, amine dehydrogenases and imine reductases.

Figure adapted from Höhne and co-workers (2009) ¹⁹ and Leipold and co-workers (2013). ²⁰	39
Figure 9 – Lipases are highly enantioselective and are often used in resolutions to afford a desired product and enantiopure starting material. Figure adapted from Takayama and co-workers (1999). ²⁴	40
Figure 10 – Examples of chiral amines that lipases have resolved (ChiPros® programme). Figure adapted from Karl and co-workers (2009). ²⁶	40
Figure 11 – Due to the highly enantioselective nature of monoamine oxidases, only a single enantiomer of the racemic amine will be chosen to oxidase to the corresponding amine. In combination with a non-selective reducing agent the imine will reduced to the racemic amine. This cycle will continue until the reactive enantiomer is completely consumed, leaving the opposite enantiomer in tact with 100% yield and enantiopure. Figure adapted from Ghislieri and co-workers (2013). ³⁰	41
Figure 12 – The existing abiotic route for the production of the bicyclic [3. 1. 0] proline moiety in Boceprevir. Figure adapted from Grunwald (2014) ³² and Li and co-workers (2012). ³¹	42
Figure 13 – The newly developed chemoenzymatic route for the production of the bicyclic [3. 1. 0] proline moiety in Boceprevir. Fewer steps are involved compared to the chemo route as well as improved yields, atom economy, lower water consumption and reduced process waste. Figure adapted from Li and co-workers (2012). ³¹	42
Figure 14 – a) A generic reaction with an (<i>R</i>)-selective transaminase in constant equilibrium with its substrates and products, where R _S and R _L are any small or large group respectively. b) Lonza in collaboration with the Bornscheuer group developed a method for removing the side product to shift the equilibrium to the desired product. Figure adapted from Höhne and co-workers (2008). ³⁸	44
Figure 15 – Asymmetric synthesis with amine dehydrogenases. As amine dehydrogenases have evolved from amino acid dehydrogenases, substrates are limited to amino acid derivatives. Figure adapted from Abrahamson and co-workers (2012). ³⁹	44
Figure 16 – Scheme of a generic prochiral imine being attacked either on its <i>re</i> - or <i>si</i> - face to form its respective chiral amines.	46
Figure 17 – Reduction of a 1-methyl-3,4-dihydroisoquinoline (MDQ) with an ATHases based on streptavidin. The ATHase resulting from incorporation of a biotinylated	

cofactor [Cp*Ir(biot- <i>p</i> -L)Cl] within streptavidin. Genetic optimisation (*) allows improvement of the catalytic performance of the artificial metalloenzyme. Figure adapted from Schwizer and co-workers (2013). ⁴⁴	47
Figure 18 – Retrosynthetic analysis (RSA) for the biosynthesis of tomaymycin in <i>S. achromogenes</i> . Core components of the diazepine ring are dihydroxyanthranilic acid and the pyrrole ring, which are joined by a non-ribosomal peptide synthetase containing the domains <i>tomA</i> and <i>tomB</i> . Figure adapted from Li and co-workers (2009) ⁴⁹	48
Figure 19 – RSA for the biosynthesis of the pyrrole ring for tomaymycin, which is biosynthesised in five steps from tyrosine by <i>tomI</i> , <i>tomH</i> , <i>tomK</i> , <i>tomN</i> and <i>tomJ</i> . Figure adapted from Li and co-workers (2009). ⁴⁹	48
Figure 20 – Schematic of the lysine degradation pathway. In the brain the pipecolic acid pathway takes precedence in the brain (left), where as in the liver the saccharopine pathway dominates (right). Figure adapted from Ijlst and co-workers (2000). ⁶⁶	50
Figure 21 – Biotransformations performed with ketimine reductase. The reductive amination reaction typically contained an α -keto acid-to-amine ratio of 1:6. Reaction rates were reported relative to the imine reduction of the piperidine substrate; rates for reductive amination were considerably lower. Figure adapted from Hallen and co-workers (2015). ⁶⁸	51
Figure 22 – Examples of siderophores, enterobactin and pyoverdine, highly complex iron chelating systems often biosynthesised by pathogens to scavenge iron from the environment or their host. Figure adapted from Pollack and co-workers (1970) ⁷⁰ and Wendenbaum and co-workers (1983) ⁷¹ for enterobactin and pyoverdine respectively.	52
Figure 23 – Simplified schematic of the catechol moieties of enterobactin chelating Fe(III). Figure adapted from Karpishin and co-workers (1992) ⁷³	52
Figure 24 – A simplified scheme for the biosynthesis of the siderophore yersiniabactin in <i>Yersinia enterocolitica</i> . A vital biosynthetic step involves the reduction of an imine bond in the central thiazole moiety by Irp3, an NADPH-dependent dehydrogenase. The substrate is covalently bound to the nonribosomal peptide synthetase high-molecular weight protein 1 (HMWP1) via a phosphopantetheinyl bond. Figure adapted from Meneely and co-workers (2012). ⁷⁶	53
Figure 25 – Scheme of the pteridine scavenging systems in <i>Leishmania major</i> with DHFR and PTR1. The pterin substrates are taken up into the cell by folate transporter (FT1)	

and biopterin transporter (BT1) allowing DHFR and PTR1 to metabolise folate and biopterin respectively. Figure adapted from Cunningham and co-workers (2001).⁷⁷ 54

Figure 26 – Simplified scheme of the pterin heterocycle for folate, dihydrofolate and tetrahydrofolate, illustrating points of hydrogenation with blue hydrogens.....55

Figure 27 – The chemical structure of methotrexate. Methotrexate is an analogue of folate therefore acts by inhibiting DHFR.55

Figure 28 – Reduction of β -carbolines using different strains of *Saccharomyces*. The best performing strain was from *S. bayanus* (Nouveaus). Figure adapted from Espinoza-Morga and co-workers (2010).⁸⁵57

Figure 29 – The epimerisation of (*S*)-reticuline to (*R*)-reticuline by the gene identified as STORR ((*S*)- to (*R*)-reticuline). Figure adapted from Winzer and co-workers (2015).⁸⁷58

Figure 30 – Schematic of the gene for STORR, where a P450 is fused to an oxidoreductase domain. Figure adapted from Winzer and co-workers (2015).⁸⁷58

Figure 31 – Simplified biosynthetic pathway of the production of morphine. Key intermediates (*R*)-reticuline and thebaine lead to morphine. Figure adapted from Kries and co-workers (2016).⁹⁰58

Figure 32 – Examples of opines. Saccharopine is the product of the metabolism of lysine. Nopaline is the product obtained from the condensation of glutamate and arginine. Octopine is the product obtained from the condensation of arginine and alanine...59

Figure 33 – Reductive amination of an amino acid with a keto acid by an OpDH, producing opines. Figure adapted from Vladimirov and co-workers (2014).⁹¹59

Figure 34 – 2-methyl-1-pyrroline (2MPN) is a 5 membered ring imine substrate which is often used to identify IRED activity.60

Figure 35 – The structure of SkIRED in complex with NADPH (PDB entry: 3ZHB). The structure composes of two monomers, subunit A in blue and subunit B in gold, dimerising as a homodimer. Each monomer has an N-terminal Rossmann fold domain for nucleotide binding highlighted in red and a C-terminal helical bundle which is involved with dimerising, with the putative active site formed in the cleft between the N- and C-terminal domains, highlighted in green. Secondary structure has been assigned for subunit A, with the N-terminal Rossmann fold with alternative β -sheets and α -helicies, starting at β 1, followed by α 1, through to α 7 and β 8, with

the C-terminal helical bundle starting at α 8 through to α 12, which dimerises with subunit B's helical bundle.....	61
Figure 36 – Proposed mechanism for the carbonyl reduction by HIBDH. Initially NADPH delivers the hydride to the carbonyl-carbon which reduces the carbonyl, followed by proton abstraction from the protonated catalytic Lys165. Figure adapted from Lokanath and co-workers (2005) ¹¹⁰	62
Figure 37 – Peptide chains of subunits A, in blue with the N- and C-termini labelled, and B, in gold, of <i>SkIRED</i> (PDB entry: 3ZHB) were superposed by secondary structure matching (SSM) on to subunits A and B of HIBDH (PDB entry: 2CVZ) which is not shown. Only the ligand, NADP(H), from <i>SkIRED</i> , in circles, and the secondary structure of <i>SkIRED</i> are shown for illustrative purposes only. The superimposition illustrates the putative catalytic Lys165 from HIBDH overlapping with Asp187 in <i>SkIRED</i> , with distances to the C4N of NADPH highlighted as 5Å and 8Å for HIBDH and <i>SkIRED</i> respectively.	63
Figure 38 – The proposed mechanism for <i>SkIRED</i> and other aspartate-IREDs was based on the mechanism for general NADPH-dependent dehydrogenases. Figure adapted from and Rodriguez-Mata co-workers (2013) ¹⁰⁸	64
Figure 39 – The proposed mechanism for <i>SSIRED</i> and other tyrosine-IREDs based on the mechanism for <i>SkIRED</i> is analogous to other NADPH-dependent dehydrogenases such as short chain alcohol dehydrogenases.	64
Figure 40 – Illustration of a typical PCR. Initially the double stranded template DNA is denatured into single strands. The reaction is cooled to allow complementary sequences to anneal by forming hydrogen bonds. The DNA polymerase is then able to recognise the primed regions and amplify the appropriate regions creating double strand fragments of the target DNA. The reaction is cycled 18 to 35 times.	71
Figure 41 - Scheme of Phase 1 of a TD-PCR, where the annealing temperature is reduced by 1-0.5°C per cycle for 10-15 cycles in the thermocycler.	73
Figure 42 – Summary of the LIC protocol using LIC-qualified T4 DNA polymerase. The left describes the linearisation protocol of the pET-YSBLIC3C vector. The right describes the treatment of the insert after PCR.....	76
Figure 43 – Summary of the In-Fusion cloning protocol. The left describes the linearisation protocol of any vector. The right describes the insert preparation by PCR.	79

Figure 44 – Two different strategies for SDM. The primer design strategies are different for both QuikChange and In-Fusion.....	80
Figure 45 – Scheme of the gradient used for purification of 6His-tagged proteins by Ni-affinity chromatography.....	86
Figure 46 – A scheme of a plot of velocity (reaction rate) versus substrate concentration that obeys Michaelis-Menten kinetics. V_{max} describes the maximum velocity of the reaction (substrate saturation), K_M (at $\frac{1}{2}V_{max}$) describes the enzymes affinity for a given substrate.	88
Figure 47 – Substrates used to assess IRED activity.....	90
Figure 48 – The substrate panel for reductive amination.....	95
Figure 49 – Protein sequence alignment of SSIRED, BcIRED and NhIRED, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Protic tyrosine for tyrosine-IREDs highlighted in red and yellow. Protein sequence alignment was performed using Clustal Omega. ¹⁴⁹ SSIRED shares 83.7% and 67.7 % sequence similarity with BcIRED and NhIRED respectively.....	110
Figure 50 – The DNA sequence in fasta format for BcIRED, codon optimised for <i>E. coli</i> .	111
Figure 51 – The DNA sequence in fasta format for NhIRED, codon optimised for <i>E. coli</i> .	111
Figure 52 – PCR products of BcIRED and NhIRED analysed by fragment analysis electrophoresis on a 1% agarose gel.	112
Figure 53 – Fragment analysis by gel electrophoresis for analysing the double digest of pET-YSB LIC3C plasmids with genes for BcIRED and NhIRED inserted.....	112
Figure 54 – The expression test for BcIRED, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of BcIRED is 36.5 kDa, including the 6His-tag and HRV3C precision site.	113
Figure 55 – The expression test for NhIRED, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of NhIRED is 33 kDa, including the 6His-tag and HRV3C precision site.....	114
Figure 56 – Chromatogram of the purification of BcIRED by SEC. Peak 2 elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa..	114
Figure 57 – Chromatogram of the purification of NhIRED by SEC. Peak 2 elutes between 50-60 mL, correlates to a species with an approximate molecular weight of 60 kDa.	115

Figure 58 – SDS-PAGE analysis of SEC purified <i>Bcl</i> RED and <i>Nhl</i> RED. Bands correspond to main peak elutions, peak 2, for both <i>Bcl</i> RED and <i>Nhl</i> RED from Figure 56 and Figure 57 respectively. Both proteins were purified by Ni-affinity prior to SEC.....	115
Figure 59 – The Michaelis-Menten fitted data for the kinetic assay with <i>Bcl</i> RED and the imine substrate 1 (2MPN).....	116
Figure 60 - The Michaelis-Menten fitted data for the kinetic assay with <i>Nhl</i> RED and the imine substrate 1 (2MPN).....	116
Figure 61 – Crystals of <i>Bcl</i> RED, at a protein concentration of 50 mg mL ⁻¹ , obtained in a 300 nL crystallisation drop in INDEX G12 from crystal trial screening.ik.....	118
Figure 62 – Optimised, using hanging drop format, crystals of <i>Bcl</i> RED in conditions containing 0.25 M MgCl ₂ , 0.1 M HEPES pH 7.5 and 30% (w/v) PEG 3350 obtained in a 2 µL crystallisation drop with a 1 mL mother liquor volume.....	119
Figure 63 – The crystal structure of <i>Bcl</i> RED-NADPH complex with subunits A and B coloured in blue and gold respectively. The electron density for NADPH in subunit A is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining NADPH.....	120
Figure 64 – The primary structure of <i>Bcl</i> RED with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵²	121
Figure 65 – Crystals of <i>Nhl</i> RED, at a protein concentration of 60 mg mL ⁻¹ , obtained from a 300 nL crystallisation drop in INDEX G10 from crystal trial screening.	122
Figure 66 – Optimised, using hanging drop format, crystals of <i>Nhl</i> RED in conditions containing 0.25 M MgCl ₂ , 0.1 M BisTris pH 5.5, 18% (w/v) PEG 3350 and 1% (w/v) OBDG obtained in a 2 µL crystallisation drop.....	123
Figure 68 – The primary structure of <i>Nhl</i> RED with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵²	125
Figure 69 – The curve obtained for the oxidative kinetic assay for <i>Nhl</i> RED with increasing concentrations of OBDG, where the reaction rate has been plotted against increasing concentrations of OBDG.....	126
Figure 70 – Protein sequence alignment of SSIREN, SRIREN and AoIREN, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interested are highlighted in red and yellow. Protein sequence alignment was performed using Clustal Omega. ¹⁴⁹ AoIREN shares 78.3% and 65.0% sequence similarity with SSIREN and SRIREN respectively.	131

Figure 71 – The expression test for AoIRED, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of AoIRED is 33 kDa, including the 6His-tag and thrombin cleavage site.	132
Figure 72 – Chromatogram of the purification of AoIRED by SEC. Peak 2 elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa..	133
Figure 73 – SDS-PAGE analysis of nickel and SEC purified AoIRED. Ni is the sample obtained from the pooled purest fractions from nickel affinity chromatography. SEC corresponds to a sample of peak 2 from Figure 72.	134
Figure 74 – Prochiral imines reduced by AoIRED. Performed by Aleku and co-workers (personal communication). n.d. = not determined. Substrate 1a had lack of baseline to accurately determine <i>e.e.</i> substrate 14 is achiral.....	135
Figure 75 – Enantioselective switch of AoIRED towards the imine substrate 6 (MDQ). Whole cells, over expressing AoIRED, cell lysate and freshly purified AoIRED yielded (<i>S</i>)-MTQ, whereas purified AoIRED that had been stored at different periods yielded (<i>R</i>)-MTQ. Enantiomeric excess is represented by percentage selectivity. Negative selectivity represents (<i>S</i>)-selectivity, whereas positive selectivity represents (<i>R</i>)-selectivity. Figure adapted from Aleku and co-workers (2016). ¹⁰⁷	136
Figure 76 – Native gel analysis of purified AoIRED upon different lengths of storage at 4°C. This figure has also been published in the supporting information of Aleku and co-workers (2016). ¹⁰⁷	137
Figure 77 – Crystals of <i>apo</i> -AoIRED, at a protein concentration of 20 mg mL ⁻¹ , obtained in a 300 nL crystallisation drop in CSS II B6 with 0.1 M Tris pH 9.0 from crystal trial screening.	138
Figure 79 – The primary structure of apo-AoIRED with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵² Dashed lines represents areas of missing density that could not be modelled.....	140
Figure 80 – The primary structure of AoIRED-NADPH complex with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵² Dashed lines represents areas of missing density that could not be modelled.....	141
Figure 81 – The active site of AoIRED, an Asn-IRED, in complex with NADPH with a water molecule within hydrogen bonding distance of Asn171. The distance observed between C4N-NADPH and OD1-Asn171 is 8Å. The electron density for NADPH is the $mF_o - DF_c$ map contoured to 3 σ obtained from refinement prior to modelling and	

refining NADPH. The carbon atoms of NADPH have been highlighted in yellow and in circles for illustrative purposes.....	143
Figure 82 – Substrates of interest with <i>AoIRED</i> for structural investigations.	143
Figure 83 – Binding sites obtained for products of 4, 5 and 12, obtained by soaking <i>apo</i> -crystals. The electron density for products of 4, 5 and 12 is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining products of 4, 5 and 12. NADPH is represented in yellow circles and ligands 4, 5 and 12 are represented in red circles for illustrative purposes. Binding sites of products 4 and 12 are bound towards the front of the active site at site S1, the residues that surround both products are hydrophobic. One of the two potential binding sites, site S4, of product 5 clashes with that of the NADPH binding site, site S2. The binding site for product 5 at site S3 is situated towards the rear of the active site and similar to products bound at site S1, are all too far from the C4N of NADPH to interpret the mechanism.	145
Figure 84 – The dimeric model of the <i>AoIRED</i> (<i>R</i>)-MTQ complex. Subunits A and B are in blue and gold respectively. The electron density for (<i>R</i>)-MTQ is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining (<i>R</i>)-MTQ.....	146
Figure 85 – The hydrophobic binding pocket for (<i>R</i>)-MTQ. Subunits A and B are in blue and gold respectively. (<i>R</i>)-MTQ is shown in red circles for illustrative purposes. The electron density for (<i>R</i>)-MTQ is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining (<i>R</i>)-MTQ.	147
Figure 86 – Comparison between <i>apo</i> - <i>AoIRED</i> (coral) and the <i>AoIRED</i> -NADPH complex (ice blue). The region between residues 207 to 247 is highlighted.....	148
Figure 87 – Crystals of <i>AoIRED</i> in complex with NADPH and MTQ, at a protein concentration of 25 mg mL ⁻¹ in buffer B (3.3), obtained in a 300 nL crystallisation drop in INDEX (with 1 mM NADPH and 1 mM MTQ) G3 from crystal trial screening.	149
Figure 88 – Optimised crystals of <i>AoIRED</i> in complex with NADPH and MTQ using sitting drop format with a 2 μ L crystallisation drop, in conditions containing 0.25 M lithium sulphate, 0.1 M BisTris pH 6.5, 25% (w/v) PEG 3350, 1 mM NADPH and 1 mM MTQ.	150

Figure 89 – Comparison of dimer length between previously obtained crystal structures of <i>AoIRED</i> . The NADPH-complex and the ternary complex are noticeably more closed compared to the <i>apo</i> - and binary xenobiotic ligand complexes.	151
Figure 90 – The primary structure of <i>AoIRED</i> -‘closed’-(<i>R</i>)-MTQ complex with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵² Dashed lines represents areas of missing density that could not be modelled.	152
Figure 91 – The <i>AoIRED</i> -‘closed’-(<i>R</i>)-MTQ complex. Subunits A and B are in blue and gold respectively. NADPH and (<i>R</i>)-MTQ is in yellow and red circles respectively for illustrative purposes only. The electron density for the ligands is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining NADPH and (<i>R</i>)-MTQ.	153
Figure 92 – Comparison between the dssp of ‘open’ and ‘closed’ forms of <i>AoIRED</i> between residues 191 to 279.	157
Figure 93 – The dimeric ‘open’ model of <i>AoIRED</i> with the binding sites for ligands 4, 5, 12, (<i>R</i>)-MTQ and NADPH. Subunits A and B are in blue and gold respectively. NADPH is shown in yellow circles for illustrative purposes only. Ligand binding sites for ligands 4, 5, 12 and (<i>R</i>)-MTQ are shown in red bloboids for illustrative purposes only. Sites S1 is where ligands 4 and 12 were bound. Site S2 is where NADPH was bound. Sites S3 and S4 are where ligand 5 were bound. Site (<i>R</i>)-MTQ is where (<i>R</i>)-MTQ was bound.	158
Figure 94 – Reductive amination performed with large quantities of IRED requiring up to 50 times excess amine donor source to drive the equilibrium towards imine formation.....	161
Figure 95 – Scheme for the production of (<i>R</i>)-rasagiline with <i>AspRedAm</i> from starting materials 1-indanone, 9, and propargylamine, c. Performed by Aleku and co-workers (personal communication).	162
Figure 96 – Sequence alignment of <i>AoIRED</i> and <i>AspRedAm</i> , where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega. ¹⁴⁹ <i>AspRedAm</i> shares 64.2% sequence similarity with <i>AoIRED</i>	163
Figure 97 – SDS-PAGE analysis from purification of <i>AspRedAm</i> by nickel-affinity chromatography and SEC. Ni is the sample obtained from purification by nickel-affinity. SEC is the sample obtained from purification by SEC. The expected	

molecular weight of <i>AspRedAm</i> is 33 kDa, including the 6 <i>His</i> -tag and thrombin cleavage site.	163
Figure 98 – Chromatogram of the purification of <i>AspRedAm</i> by SEC. Main peak elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa.....	164
Figure 99 – Optimised crystals of <i>AspRedAm</i> , using a 2 μ L crystallisation drop in sitting drop format, obtained in conditions containing 50 mM Tris pH 8.5, 23% (w/v) PEG 3350, 5% (v/v) butan-1,4-diol and 6% (w/v) xylose.....	167
Figure 100 – Crystals of <i>AspRedAm</i> , at a protein concentration of 17 mg mL ⁻¹ in buffer B (3.3), obtained a 300 nL crystallisation drop in CSS II + 0.1 M MES pH 6.0 + 0.5 mM NADPH A6 from crystal trial screening.....	168
Figure 101 – The crystal structure of <i>AspRedAm</i> -NADPH complex. Subunits A and B are in blue and gold respectively.....	169
Figure 102 – The primary structure of <i>AspRedAm</i> -NADPH complex with secondary structure assignment, created using DSSP ¹⁵⁰⁻¹⁵¹ and represented using ALINE. ¹⁵² Dashed lines represents areas of missing density that could not be modelled.	170
Figure 103 – The binding site of NADPH in <i>AspRedAm</i> . Subunits A and B are in blue and gold respectively. NADPH is in yellow circles for illustrative purposes only. Key interactions include Arg33 contributing π -stacking with the adenine ring of NADPH and Met 13 contributing π -stacking to the nicotinamide ring of NADPH.....	171
Figure 104 – Scheme of the reaction of interest: the production of (<i>R</i>)-rasagiline from 1-indanone and propargylamine with <i>AspRedAm</i>	171
Figure 105 – Crystals of <i>AspRedAm</i> , at a protein concentration of 17 mg mL ⁻¹ in buffer B (3.3), obtained in a 300 nL crystallisation drop in INDEX + 0.5 mM NADPH + 0.5 mM <i>rac</i> -rasagiline G3.....	172
Figure 106 – Optimised crystals of <i>AspRedAm</i> in conditions containing: (left) 0.2 M lithium sulphate, 0.1 M BisTris pH 6.5, 27% (w/v) PEG 3350, 10 mM NADPH and 10 mM <i>rac</i> -rasagiline, and (right) 0.2 M lithium sulphate, 0.1 M BisTris pH 6.5, 27% (w/v) PEG 3350, 5 mM NADP ⁺ and 5 mM <i>rac</i> -rasagiline.....	173
Figure 107 – Crystal (Left), highlighted in red (Right), of <i>AspRedAm</i> , at a protein concentration of 18 mg mL ⁻¹ , obtained in INDEX + 0.5 mM NADP ⁺ + 0.5 mM 1-indanone + 0.5 mM propargylamine G11.....	174

- Figure 108 – The crystal structure of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. a) Top down secondary structure view. b) Top down surface view, the central cavity measures a distance of 46 Å. c) Side on secondary structure view. d) Side on simplified bloboid crystal contacts view where the gold models represent the asymmetric unit and the ice blue represents the nearby crystal packing.....175
- Figure 109 – The primary structure of *AspRedAm*-ternary complex with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.176
- Figure 110 – The dimeric model of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. Subunits A and B are in blue and gold respectively.....176
- Figure 111 – Active site of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. Subunits A and B are in blue and gold respectively. The ligands NADP(H) and (*R*)-rasagiline are in yellow and red circles respectively for illustrative purposes only. The electron density for (*R*)-rasagiline is the omit map ($mF_o - DF_c$) contoured to $\pm 3\sigma$ obtained from refinement prior to modelling and refining (*R*)-rasagiline. Left – View of active site with surrounding residues. Right – View of active site with distances within bonding distances shown.178
- Figure 112 – Chart displaying relative activity of amine/carbonyl pairs in reductive amination reactions with suggested amine-to-carbonyl ratios for biotransformations. Compounds presented in the plot area are representative examples of products obtained in biotransformations.....181
- Figure 113 – IRED superfamily analysis. Grey sequences represent potential IREDs of bacterial origin. Purple sequences represent potential IREDs of fungal origin. Sequence analysis performed by Scheller and co-workers (personal communication).184
- Figure 114 – Sequence alignment between the fungal RedAms, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹185
- Figure 115 – Proposed mechanism for reductive amination with RedAms. a) Activation of the carbonyl by Tyr177 and deprotonation of the ammonium substrate by Asp169. b) Nucleophilic attack of the carbonyl by the amine. c) Imine reduction by NADPH. Residue numbering is based on the sequence numbering of *AspRedAm*.186

Figure 116 – Sequence alignment between IREDs and RedAms, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. The DSSP is of the ternary complex of <i>AspRedAm</i> . Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega. ¹⁴⁹ <i>AspRedAm</i> shares 69.9%, 64.2% and 61.1% sequence homology with <i>NhIRED</i> , <i>AoIRED</i> and <i>StrepRedAm</i> respectively.....	189
Figure 117 – The protein sequence for <i>StrepRedAm</i> in fasta format.....	190
Figure 118 – The DNA sequence for <i>StrepRedAm</i> , codon optimised for <i>E. coli</i> , in fasta format.....	190
Figure 119 – Fragment analysis of double digest by restriction enzyme, confirming the presence of positive gene insertions.....	191
Figure 120 – The expression test for <i>StrepRedAm</i> , where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of <i>StrepRedAm</i> is 33.5 kDa, including the 6 <i>His</i> -tag and HRV3C protease precision site.	192
Figure 121 – The chromatogram for the anion exchange purification for <i>StrepRedAm</i> . Main peak elution occurs at approximately 500 mM NaCl.	193
Figure 122 – The chromatogram of the SEC for <i>StrepRedAm</i> . The peak eluting between 50-60 mL indicates a species eluting with a molecular weight of approximately 60 kDa.	193
Figure 123 – SDS-PAGE analysis for the purification of <i>StrepRedAm</i> by nickel affinity, anion exchange (Q) and SEC. Samples Q1, Q2 and Q3 correspond to samples taken from the anion exchange purification in Figure 120. Over the course of the purification, degradation was increased. After SEC the purity of the protein is not sufficient for crystallographic and kinetic assay studies.	194
Figure 124 – The chromatogram for the nickel purification of <i>StrepRedAm</i> with protease inhibitors. The peak area highlighted in blue was the fractions taken for further purification.	194
Figure 125 – The chromatogram of SEC for <i>StrepRedAm</i> . The main peak eluting between 50-60 mL indicates a species eluting with a molecular weight of approximately 60 kDa. The peak area highlighted in blue was the fractions taken for further characterisation.....	195

- Figure 126 – SDS-PAGE analysis of the purification of *StrepRedAm*. Main Ni is a sample from the main peak of the nickel purification. Purest Ni are the fractions towards the end of the peak elution which were taken for further purification. SEC is the protein obtained after SEC.195
- Figure 128 – Chiral reductive amination biotransformations performed with *StrepRedAm* at carbonyl-to-amine ratios of 1:1, 1:4, 1:20 and 1:50. [a] Absolute configuration not yet assigned. [b] Absolute configuration assigned based on N-methyl analogues. .198
- Figure 129 – Reactions attempted in the crystallisation of *StrepRedAm*, in an attempt to trap a ternary complex.199
- Figure 130 – The phylogenetic tree for recently investigated IREDs and RedAms, including the IRED library from Roche.¹⁰³ The phylogenetic tree analysis was performed with POWER¹⁵⁴ and displayed with iTOL.¹⁵⁵ The phylogenetic tree shows the RedAms clustered together neighbouring IREDs that may exhibit RedAm activity. IREDs further away from the RedAm cluster such as the previously investigated *SkIRED*, *SRiRED*, *BcIRED*, *SSIRED* and *NhIRED* are unlikely to exhibit RedAm activity.200
- Figure 131 – Direct comparison of reductive amination biotransformations performed with *StrepRedAm* and *AspRedAm*. [a] Absolute configuration not yet assigned.202

List of Tables

Table 1 – A standard reaction setup for KOD Hot Start Polymerase. Adapted from KOD Hot Start DNA Polymerase user manual. ¹¹¹	69
Table 2 – Typical thermocycling conditions for KOD Hot Start Polymerase. Adapted from KOD Hot Start DNA Polymerase user manual. ¹¹¹	69
Table 3 – A standard reaction setup for Pfu DNA Polymerase. Adapted from Pfu DNA Polymerase user manual. ¹¹²	69
Table 4 – Typical thermocycling conditions for Pfu DNA Polymerase. Adapted from Pfu DNA Polymerase user manual. ¹¹²	70
Table 5 - Example of extending a standard PCR protocol to a TD-PCR, to be used on a 1000 bp fragment with KOD Hot Start DNA Polymerase. Table adapted from Korbie and co-workers (2008). ¹¹³	72
Table 6 – Components required for linearisation of the pET-YSB LIC3C vector by <i>BseRI</i>	74
Table 7 – Components required for the digest of the pET-YSB LIC3C vector to generate sticky-ends.....	75
Table 8 – Components required for the digest of insert to generate the reciprocal sticky-ends required to anneal to the treated pET-YSB LIC3C vector.	75
Table 9 – Components required for the annealing of the treated insert and treated vector. This reaction will generate a nicked plasmid.....	75
Table 10 – The parameters required for primers used in In-Fusion cloning. Adapted from In-Fusion® HD Cloning Kit user manual. ¹¹⁹	77
Table 11 – The components required for the In-Fusion reaction (annealing and linearised the vector to the insert). Adapted from In-Fusion® HD Cloning Kit user manual. ¹¹⁹ ..	78
Table 12 – Components required for the template digestion reaction by <i>DpnI</i>	81
Table 13 – Standard protocol for heat shock transformation.....	82
Table 14 – Protocol for performing a double digest.	82
Table 15 – Components required for SDS-PAGE.	84
Table 16 – Buffers components used in purification.....	85

Table 17 – Protocol used for purifying 6His-tagged protein by Ni-affinity chromatography.	85
Table 18 – Typical components used in a reaction for assessing kinetics parameters.	89
Table 19 – Typical reaction components used for biotransformations.....	90
Table 20 – Protocol for analysing biotransformations for substrates 1-6. Hex:iPrOH:DEA is hexane:isopropanol:diethylamine.....	92
Table 21 – Reaction components for reductive amination.	95
Table 22 – GC methods for analysing reductive amination substrates and products. Products with ketone 7 are non-chiral. Absolute configuration of 8e is based on 8a. Absolute configuration for products with ketone 10 have not yet been assigned. ...	97
Table 23 – HPLC methods for analysing reductive amination substrates and products.	99
Table 25 – The kinetic parameters for <i>BcIRED</i> and <i>NhIRED</i> for the imine substrate 1 (2MPN).	117
Table 26 – Biotransformations performed with <i>BcIRED</i> and <i>NhIRED</i> with imine substrates 1-6.....	117
Table 27 – Crystal data collection and refinement statistics for <i>BcIRED</i> -NADPH complex and <i>NhIRED</i> -OBDH complex. Numbers in brackets refer to the data in the highest resolution shell.	127
Table 29 – Comparison of biotransformations for <i>SSIRED</i> , <i>BcIRED</i> and <i>NhIRED</i> with substrates 1, 2 and 6. Biotransformations with substrates 3, 4 and 5 were not determined for <i>SSIRED</i> . [a] Biotransformation data for <i>SSIRED</i> were taken from Leipold and co-workers (2013). ²⁰	129
Table 30 – Kinetic parameters for selected <i>AoIRED</i> substrates.	138
Table 31 – Crystal data collection and refinement statistics for all <i>AoIRED</i> data collection. Numbers in brackets refer to the data in the highest resolution shell.....	155
Table 33 – Comparison of kinetic parameters between <i>SRIRED</i> and <i>AoIRED</i> for imine substrates 1, 6 and 14. [a] kinetics constants taken from Hussain and co-workers (2015). ⁴³ [b] kinetic constants provided by Aleku and co-workers.....	156
Table 34 – Biotransformations for variants of <i>AoIRED</i> . Substrate numbers refer to those in Figure 74. Data is presented as: conversion (%), <i>e.e</i> (%) and absolute configuration.	158

Table 35 – Biotransformation and kinetic data for substrate 1a (2MPN) for variants of <i>AoIRED</i>	159
Table 36 – Relative activity of carbonyl acceptors in <i>AspRedAm</i> catalysed reductive aminations. Specific activities (i) cyclohexanone/propargylamine = 6.68 U/mg (100%); (ii) cyclohexanone/methylamine = 2.23 U/mg (100%). Reaction conditions: 15 mM carbonyl acceptor, 60 mM amine nucleophile (1:4 ketone-to-amine ratio), 0.3 mM NADPH and 100 mM borate buffer pH 9.0.....	165
Table 37 – Relative activity of amines in <i>AspRedAm</i> catalysed reductive amination. Specific activities: (i) cyclohexanone/propargylamine = 6.68 U/mg (100%); (ii) 4-phenyl-2-butanone/propargylamine = 0.54 U/mg (100%). Reaction conditions 15mM carbonyl acceptor: 60mM amine nucleophile (1:4 ketone-to-amine ratio), 0.3mM NADPH and 100 mM borate buffer pH 9.	166
Table 38 – Crystal data collection and refinement statistics for all <i>AspRedAm</i> data collection. Numbers in brackets refer to the data in the highest resolution shell. ...	179
Table 39 – Kinetic data of <i>AspRedAm</i> variants.....	180
Table 40 – <i>AspRedAm</i> catalysed reductive amination of carbonyl compounds at low molar ratios.....	182
Table 41 – Comparison of reductive amination activities between <i>AspRedAm</i> and <i>AoIRED</i>	183
Table 42 – Primers and PCR conditions for the amplification of <i>StrepRedAm</i> to be annealed by In-Fusion® into the pET-YSBLIC3C vector.....	191
Table 43 – List of primers.	205

Acknowledgements

I would like to thank everyone at K-block and M-block who have made my time at YSBL very enjoyable and memorable. I would like to especially thank Sam Hart and Johan Turkenburg for all the X-ray robot testing I requested and for staying up late into the early mornings for the long shift data collections so that I didn't have to. Huge thanks go to the technical support staff Sally Lewis, Simon Grist, Juliet Borgia and Louise Haige, who are not your everyday technicians, they are very much the Lab Heroes keeping things in running order which is by no means an easy task.

Much of the work produced in this thesis could not have been done without my friends and collaborators, special thanks go to Nick Turner and his group at the Manchester Institute of Biotechnology, especially Godwin Aleku, Shahed Hussain, Scott France, Friedemann Leipold and Juan Mangas. It has been a pleasure working with them.

Of course I must thank my family for the support they have given me these last 8 years or so at York. Even though they have no idea what I do, they have been very supportive towards my decisions.

During my time at YSBL I have gained some invaluable friends: Annika, Chantel, Claudia, Lilly, Anibal, Muhi, Amina, Marc, Laziana, Zak, Wayne, Glyn, Sandra, Christian, Rui-Gang, Marek, Gideon and the list could go on.

Also I must thank my girlfriend and partner, Sherry, who has supported me in keeping me fed, fat, alive and happy since 2012.

Lastly, I am hugely in debt to Chantel Jensen who mentored my BSc project and Gideon Grogan who took me for the BSc project in 2012 and took a chance with me to do an MSc by Research which could potentially have been a big waste of time and effort on his behalf, however, luckily things turned out well. My

life would probably be very different if I did not do my BSc project with Gideon. If things had not gone that way, I'd probably have been managing a farm in China that has recently gone under. No words can express how grateful I am for meeting these two people. Not only have I gained additional degrees and experience, but I have also gained a career and now many future opportunities; things that can never be or will be taken for granted.

Author's Declaration

Dr Shahed Hussain synthesised and provided the following substrates:

- 2-methyl-1-piperidine
- 2-*n*-propyl-1-piperidine
- 2-phenyl-1-piperidine
- 2-phenyl-1-piperidine
- 2-*p*-methoxyphenyl-1-piperidine

Scott P. France synthesised and provided the following substrates:

- Dibenzazepine
- Sertraline "ketone"
- Sertraline "imine"
- *rac*-Sertraline "amine"

Dr Anibal Cuetos-Fernandez synthesised and provided the following substrate:

- *rac*-Rasagiline
- *rac-p*-fluoro-Selegiline

Godwin Aleku provided the plasmids for *AoIRED* and *AspRedAm*, as well as providing GC and HPLC chromatography support.

The work presented in this thesis has been performed by myself unless stated otherwise.

This thesis has not been presented for any other award at this or any other institution.

Parts of the work presented in this thesis has been published in the following journals:

H Man, E Wells, S Hussain, F Leipold, S Hart, JP Turkenburg, NJ Turner and G Grogan, *Structure, Activity and Stereoselectivity of NADPH-Dependent Oxidoreductases Catalysing the S-Selective Reduction of the Imine Substrate 2-Methylpyrroline*, ChemBioChem (2015), 16, 7, 1052-1059.

GA Aleku*, **H Man***, SP France, F Leipold, S Hussain, L Toca-Gonzalez, R Marchington, S Hart, JP Turkenburg, G Grogan and NJ Turner, *Stereoselectivity and Structural Characterization of an Imine Reductase (IREd) from Amycolatopsis orientalis*, ACS Catalysis (2016), 6, 6, 3880-3889.

GA Aleku*, H Man*, SP France, J Mangas-Sanchez, M Sharma, F Leipold, S Hussain, SL Montgomery, G Grogan and NJ Turner, *A Reductive Aminase Enzyme from Aspergillus oryzae*, (submitted, under revision).

1. Introduction

1.1. Abiotic Chiral Amine Synthesis

A chiral amine is a molecule with an amine functional group that has point chirality with its stereogenic centre at the carbon directly bonded to the nitrogen atom (**Figure 1**).

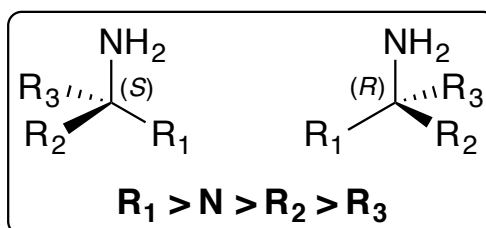


Figure 1 – Generic chiral amine structure with its Cahn-Ingold-Prelog (CIP) priority assigned as $R_1 > N > R_2 > R_3$. Where “>” means greater than.

Chiral amines have been used as versatile synthetic building blocks and as drug compounds themselves (**Figure 2**). Besides the synthetic versatility of amines, they have also been known to be biologically active.

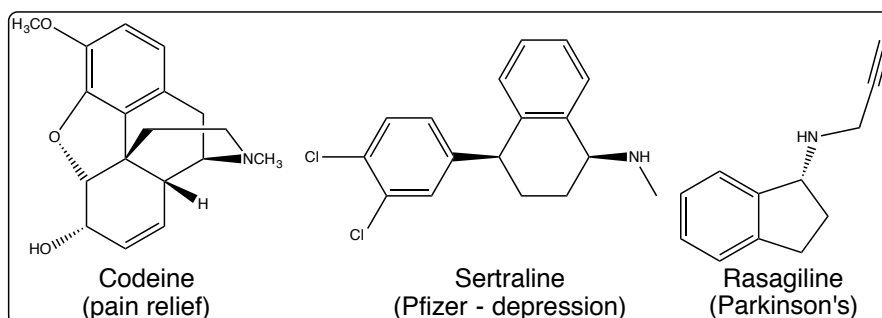


Figure 2 – Examples of chiral amines that are biologically active drugs. Codeine is an over-the-counter drug that is commonly used for pain relief. Sertraline is an antidepressant classed as a selective serotonin reuptake inhibitor. Rasagiline is used in the treatment of Parkinson's disease and irreversibly inhibits monoamine oxidase-B.

Despite advances in the development of asymmetric catalysis for the synthesis of chiral compounds, asymmetric synthesis is still a challenge. In 2001, Ryoji Noyori and William Standish Knowles shared the Nobel Prize in chemistry for

their work on asymmetric hydrogenation; with the other half of the Nobel Prize awarded to Karl Barry Sharpless for his work on asymmetric oxidations. Work performed by these Nobel Prize laureates rapidly advanced the field of organic chemistry especially for the production of chiral compounds (**Figure 3**).

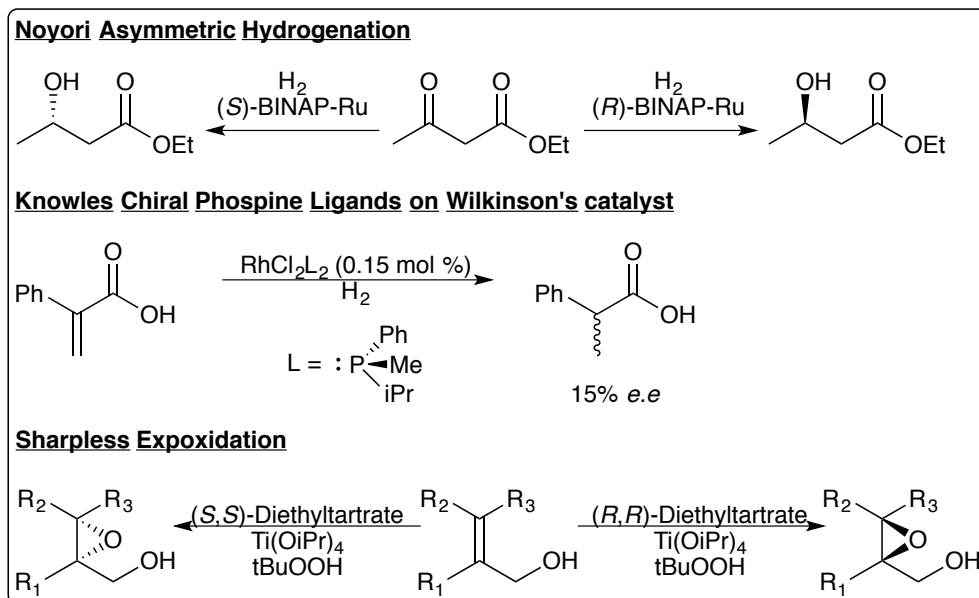


Figure 3 – Notable reactions by the Nobel Prize laureates Ryoji Noyori, William Standish Knowles and Karl Barry Sharpless: Noyori asymmetric hydrogenation with the chiral catalyst BINAP, development of Knowles experimental catalyst which substituted achiral phosphine ligands on a Wilkinson's catalyst with chiral phosphine ligands and the Sharpless epoxidation with the chiral catalyst formed *in situ* with diethyl-tartrate, titanium tetra-isopropoxide and *tert*-butyl hydroperoxide. Figure adapted from Noyori and co-workers (1987),¹ Knowles and co-workers (1973)² and Sharpless and co-workers (1980).³

Asymmetric catalysts are predominantly organometallic catalysts that contain a central metal ion. As such, due to the coordination sphere of the central metal ion of the asymmetric catalyst, chiral auxiliary ligands such as triphenylphosphine or tricyclohexylphosphine are forced to bind with constrained geometry, such as square planar or octahedral (**Figure 4**). Typically, bulky phosphine ligands are used to bind to the central metal ion as seen in Grubbs catalyst.⁴ Therefore, in combination with the metal ion's geometry and bulky ligands, enantioselectivity arises from substrates coordinating in a strict fashion. This is because the bulky ligands have created a chiral environment where steric hindrance and favourable interfacial interactions between the substrate and chiral catalyst, such as π -interactions or hydrogen bonding, take precedence and direct the coordination of the substrates. Thus, only one face of a prochiral asymmetric

substrate is exposed and presented for attack, which in turn yields enantioselectivity.

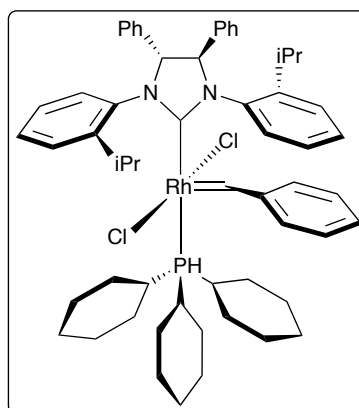


Figure 4 – First generation Grubbs catalyst used for olefin metathesis. The catalyst uses bulky ligands to achieve good selectivity. Figure adapted from Funk and co-workers (2006).⁴

Ketones and imines behave similarly upon attack of a nucleophile. The same chiral catalyst that can be used to asymmetrically hydrogenate ketones or keto-esters can also be used to asymmetrically hydrogenate imines.⁵⁻⁶ However, the two most common routes to synthesising chiral amines are through asymmetric amination⁷ and reductive amination⁸ (**Figure 5**).

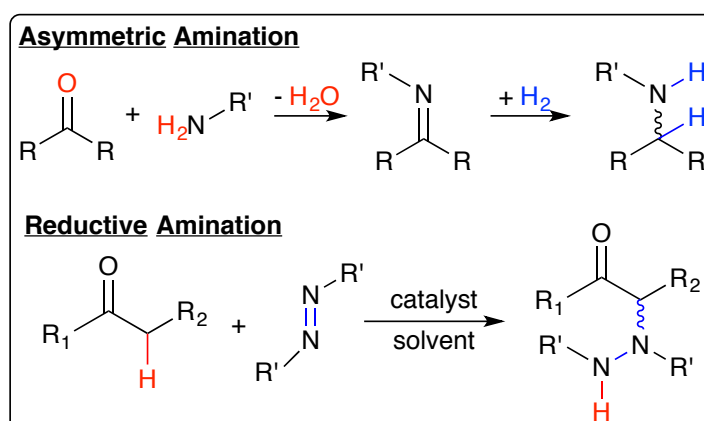


Figure 5 – The two most common routes for the abiotic synthesis of chiral amines, asymmetric amination and reductive amination. Both routes require chiral catalysts.

Despite high yields and enantiomeric purity, abiotic reactions can take up to 48 h to complete and may require heavy metals, air-sensitive chiral catalysts, toxic

reagents and solvents which may involve tedious work-up reactions resulting in a significant loss in yield.

Non-green conditions are a concern for synthetic chemists as well as the rising price of expensive organometallic chiral catalysts. Precious metals such as rhodium, platinum and palladium are often at the core of asymmetric organometallic catalysts, however the world supply of these precious metals is running out and will inevitably be exhausted, thus causing a rise in price for these asymmetric catalysts. Alternatives are being developed including, most recently proline organocatalysts, however these catalysts are not as efficient or selective as the organometallic catalysts containing precious metals.⁹ Proline based organocatalysts are continually being improved,¹⁰ for example in the reduction of imines¹¹⁻¹² (**Figure 6**); however the reaction conditions still require large quantity of solvent and halosilanes are used which will convolute product recovery. Some efforts are being made to substitute precious metals to more abundant metals such as nickel,¹³ however these catalysts are not as efficient or selective as their predecessor where yields and enantiomeric purity can vastly vary depending on the solvent system used,¹⁴ therefore, alternatives are required.

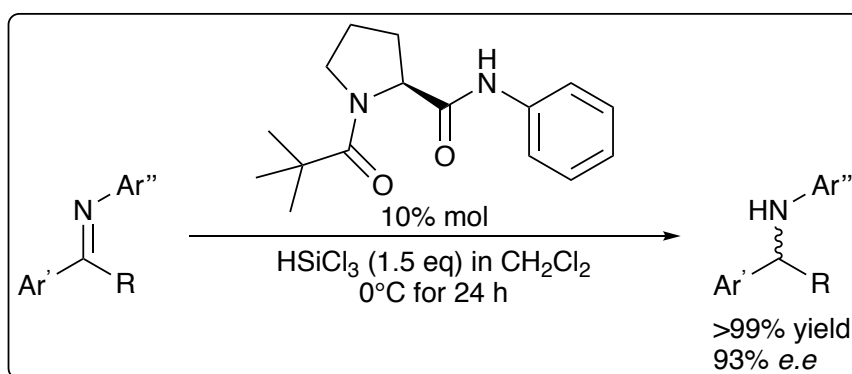


Figure 6 – Reduction of imines with trichlorosilane activated with a proline organocatalyst in dichloromethane. Good yields and modest enantiomeric excess are achieved. Work performed by Kanemitsu and co-workers (2001).¹²

1.2. Biocatalysts: Enzymes as catalysts

Enzymes have evolved to be very efficient catalysts with exquisite regio-, chemo- and enantioselectivity and the ability to work on specific substrates. To chemists using enzymes¹⁵, biocatalysts appear as an attractive alternative to abiotic catalysts for organic chemistry, as they are able to operate at ambient temperatures and pressures, and perform functional group transformations more efficiently with fewer steps, under more environmentally friendly conditions. More recently, biocatalysts have been the preferred catalyst over chemocatalysts in the pharmaceutical industry¹⁶ due to potentially higher enantioselectivity, and more, importantly easier product purification.

Biocatalysts are often used as purified enzymes or as whole cell biocatalyst where the gene is overexpressed in a host that has been genetically modified to be benign such as *E. coli*. Both of these can be supplied as a freeze-dried powder. Each has their own advantages, however a synthetic chemist will prefer a purified enzyme powder. This is as a synthesis laboratory may not have the correct safety procedures to work with genetically modified organisms; a minimum of biosafety level 1 is required, which includes knowledge of safely disposing and decontamination of biological material. Whole cell biocatalysts are capable of protecting the enzyme from solvents and bi-phasic layers that are typically used in organic synthesis. Purified enzymes do not have the longevity that whole cells provide and the enzyme would need to be prepared and purified where a synthetic chemist may not have the equipment or expertise to do so. However, there is less chance of by-products forming, as there are other enzymes present in a whole cell system and there are peptidases in the cell that will break down the over-expressed non-native enzyme.

1.2.1. Applying Molecular Biology to Biocatalysis

Due to advances in molecular biology and genome mining, techniques from biology have become standardised and easier to access, so it is no longer necessary to isolate the desired protein from the native source, which can be

difficult, time consuming and potential hazardous especially if the cultured microorganism is a pathogen. Today, the standard protocol is to obtain the synthetic gene, which can be codon-optimised to optimise expression for any desired expression host. These genes can be cloned or subcloned with a suitable affinity purification tag into a suitable expression vector and heterologously expressed in a host such as *Escherichia coli* (*E. coli*). This is as *E. coli* is capable of doubling its cell number by cell division¹⁷ within 20 min, typically at 37 °C, in a rich growth media. It is possible to obtain a very high cell density very quickly in this way.

When an enzyme of desired function has been isolated and characterised, it is often the starting point for an evolution programme to improve the enzyme or evolve the enzyme to perform a new function.¹⁸ For industrial biocatalysis, it is important that an enzyme exhibits the following criteria: it is (thermo-) stable, solvent tolerant and is very active. To achieve the aforementioned criteria, multiple rounds of directed evolution are often required where very little of the original DNA sequence remains (**Figure 7**).

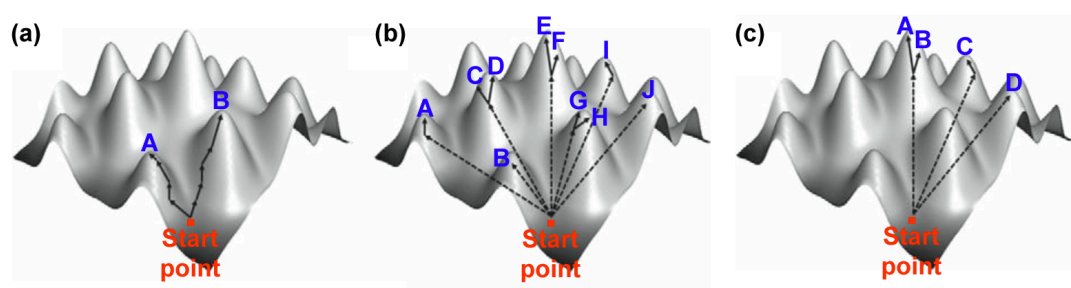


Figure 7 – Protein fitness landscapes of a wild type protein with different end points created by different evolution strategies: (a) iterative cycles of point mutations, (b) fast evolution with DNA shuffling and (c) saturation mutagenesis where degenerate codons are used to randomise the outcome of mutants. Figure adapted from Dalby (2011).¹⁸

1.2.2. Current Biocatalytic Toolbox for Producing Chiral Amines

There are many different classes of enzymes capable of producing chiral amines (**Figure 8**), such as hydrolases, monoamine oxidases and transaminases, however, most can only do so through kinetic resolutions, dynamic kinetic resolutions or using co-substrates respectively which may convolute product separation for the desired amine product.

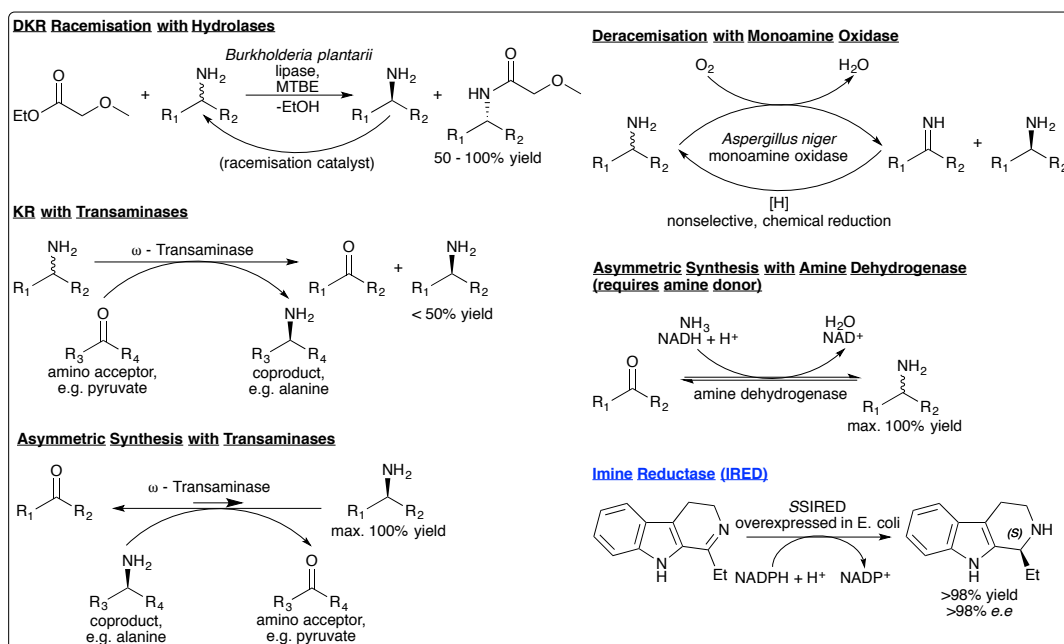


Figure 8 – The current biocatalytic toolbox for accessing chiral amines with hydrolases, monoamine oxidases, transaminases, amine dehydrogenases and imine reductases. Figure adapted from Höhne and co-workers (2009)¹⁹ and Leipold and co-workers (2013).²⁰

Biocatalysts have many industrial and commercial applications besides industrial synthetic chemistry. Biocatalysts are widely used in biologically active laundry detergents and in the dairy industry; in both these examples lipases are used. Lipases are capable of hydrolysing esters²¹ in a high water content environment. In low water systems where a solvent is the aqueous medium, lipases can also perform acylations on alcohols²² and amines.²³ Lipases are also incredibly enantioselective, however this is also a potential disadvantage for synthetic chemistry, for example if a racemate is presented to the lipases, it can only hydrolyse or acylate one of the two enantiomers²⁴ yielding a theoretical maximum of 50% yield rather than achieving complete conversion.

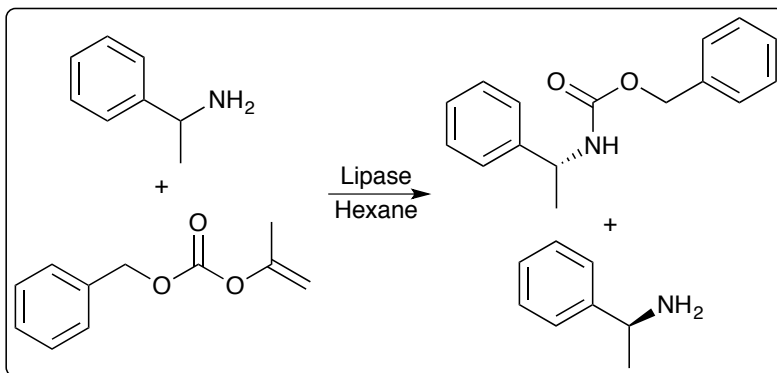


Figure 9 – Lipases are highly enantioselective and are often used in resolutions to afford a desired product and enantiopure starting material. Figure adapted from Takayama and co-workers (1999).²⁴

Despite the limitations of lipases, they are still used for resolutions (**Figure 9**) due to their versatility and, when combined with another enzyme or chemocatalyst²⁵ that can racemise the residual starting material, a dynamic kinetic resolution can be achieved. BASF use lipases to resolve many substituted phenylethylamines and their derivatives (**Figure 10**), including bicyclic amines, aliphatic amines, cycloaliphatic amines, aryl alkyl amines and other functional amines, to provide tonne scale synthesis of optically active amines as part of the BASF ChiPros® programme.²⁶

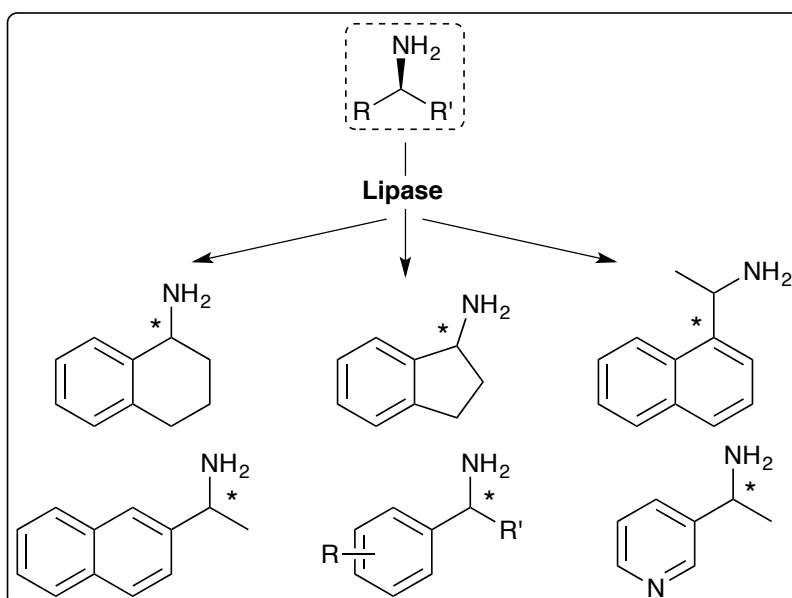


Figure 10 – Examples of chiral amines that lipases have resolved (ChiPros® programme). Figure adapted from Karl and co-workers (2009).²⁶

Monoamine oxidases are highly enantioselective flavin-dependent enzymes that are capable of oxidising one enantiomer to the imine and, in certain cases where the imine has not been stabilised by conjugation, the imine will dissociate to its corresponding ketone and amine (usually ammonia).²⁷ Rather than allowing the imine to dissociate, if a non-selective reducing agent is introduced, the imine will be reduced to its racemic amine starting material²⁸⁻³⁰. The recycling of the imine (**Figure 11**) will continue until a single enantiomer is left which the monoamine oxidase is unable to accept as a substrate due to its selectivity.

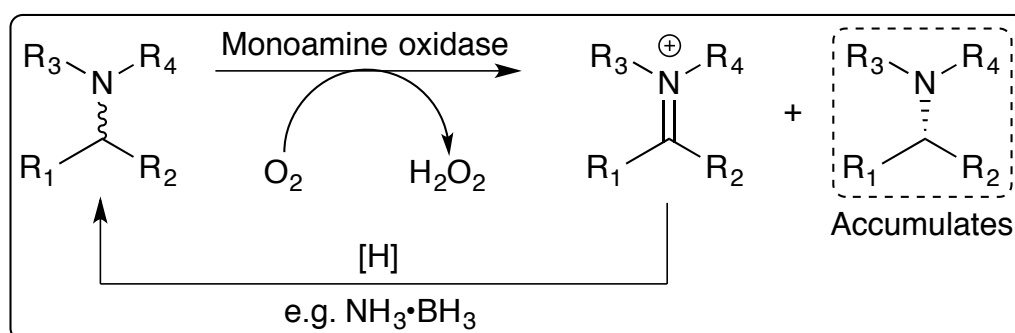


Figure 11 – Due to the highly enantioselective nature of monoamine oxidases, only a single enantiomer of the racemic amine will be chosen to oxidise to the corresponding amine. In combination with a non-selective reducing agent the imine will be reduced to the racemic amine. This cycle will continue until the reactive enantiomer is completely consumed, leaving the opposite enantiomer in tact with 100% yield and enantiopure. Figure adapted from Ghislieri and co-workers (2013).³⁰

In 2012 Merck and Codexis collaborated to develop a more efficient chemoenzymatic route for the production of the bicyclic [3.1.0] proline moiety in Boceprevir³¹, a drug for the treatment for hepatitis C, to replace the existing chemo route (**Figure 12**).

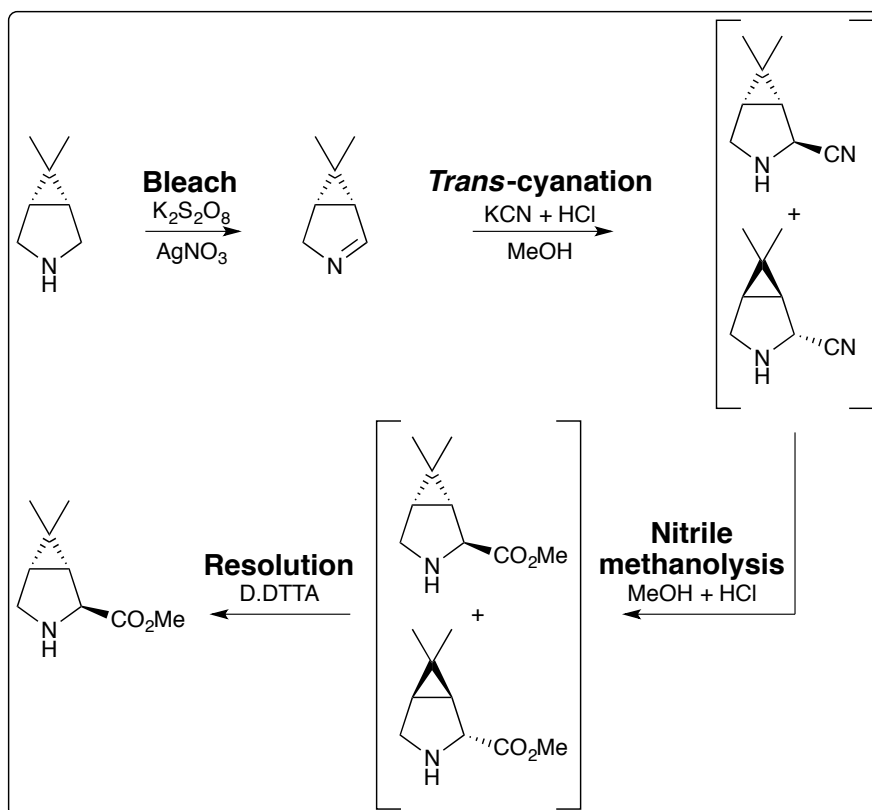


Figure 12 – The existing abiotic route for the production of the bicyclic [3. 1. 0] proline moiety in Boceprevir. Figure adapted from Grunwald (2014)³² and Li and co-workers (2012).³¹

The newly developed chemoenzymatic route (Figure 13) increased the yield by 150% and reduced water consumption and process waste by 60.7% and 63.1% respectively.

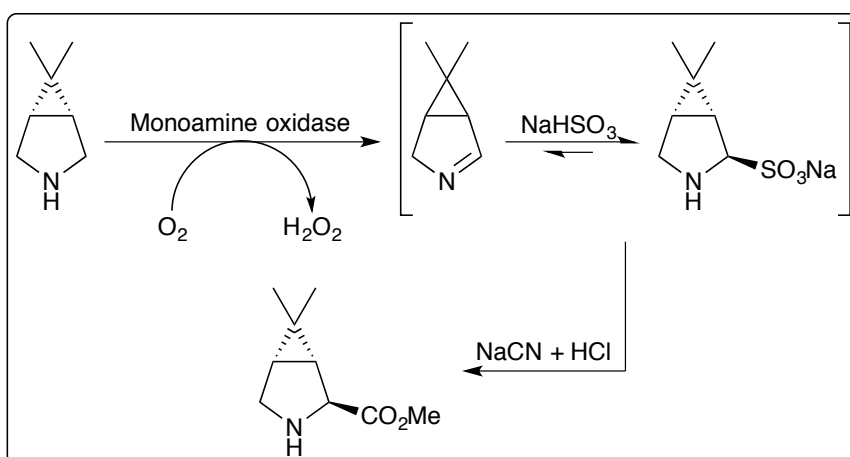


Figure 13 – The newly developed chemoenzymatic route for the production of the bicyclic [3. 1. 0] proline moiety in Boceprevir. Fewer steps are involved compared to the chemo route as well as improved yields, atom economy, lower water consumption and reduced process waste. Figure adapted from Li and co-workers (2012).³¹

Monoamine oxidases are therefore incredibly useful tools with broad substrate specificity for deracemisation reactions, to obtain 100% theoretical yields with high enantiopurity.

Transaminases, also known as aminotransferases, are enzymes capable of transferring an amino group from an amine donor to a desired ketone utilising the cofactor pyridoxal phosphate (PLP). Transaminases can be generally assigned as either α -transaminases or ω -transaminases, which is based on the relative position of the amino group to be transferred to that of the carboxyl group of the substrate.³³ α -transaminases requires a carboxylic group in the α -position to the carbonyl or amine functional group, which in turn, only allows the formation of α -amino acids.³⁴ ω -transaminases in principle are capable of transferring amino groups from any amine donor onto any carbonyl receptor,³⁵ which makes them much more useful for synthetic chemistry. Lonza have played a big part in using biocatalysts for the production of fine chemicals such as nicotinamide³⁶ and L-carnitine.³⁷ Lonza are interested in using transaminases to produce enantiopure amines, however a common issue with transaminases is that the reaction is reversible, therefore the amine product will be used to aminate the keto product, yielding an equilibrium where only 50% of the desired product could ever be obtained (**Figure 14a**). As such, in collaboration with the Bornscheuer group at the University of Greifswald,³⁸ Lonza attempted to shift the equilibrium by including pyruvate decarboxylase to remove the unwanted by- product (**Figure 14b**). The concept was successful and yielded 1-*N*-Boc-(3*S*)-aminopyrrolidine in 80% yield and 99% *ee*.

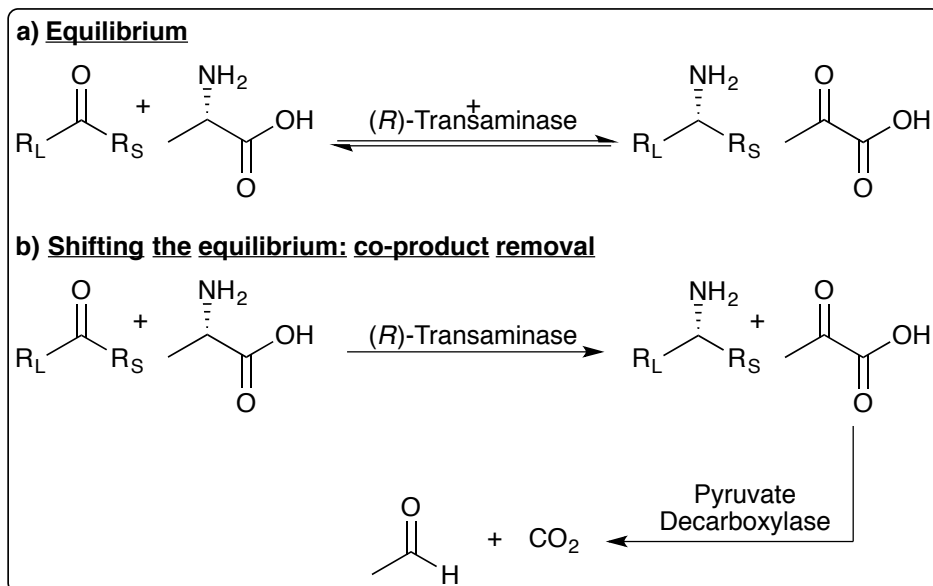


Figure 14 – a) A generic reaction with an (*R*)-selective transaminase in constant equilibrium with its substrates and products, where R_S and R_L are any small or large group respectively. b) Lonza in collaboration with the Bornscheuer group developed a method for removing the side product to shift the equilibrium to the desired product. Figure adapted from Höhne and co-workers (2008).³⁸

However, despite developing reactions to remove side products from reactions with transaminases, they often suffer from substrate or product inhibition, which is still an inherent issue with transaminases.

Amine dehydrogenases are NADH-dependent dehydrogenases that are capable of producing chiral amines with the addition of ammonia to a ketone of interest by reductive amination (**Figure 15**). However, the majority of the amine dehydrogenases are evolved *in vitro* from amino acid dehydrogenases,³⁹⁻⁴¹ and as such their substrate scope appears very limited to substrates that exhibit similar scaffold to their native amino acids.

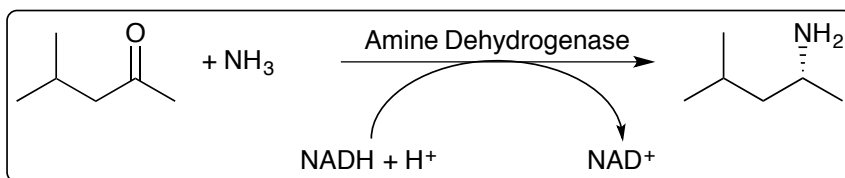


Figure 15 – Asymmetric synthesis with amine dehydrogenases. As amine dehydrogenases have evolved from amino acid dehydrogenases, substrates are limited to amino acid derivatives. Figure adapted from Abrahamson and co-workers (2012).³⁹

A natural amine dehydrogenase from *Streptomyces virginiae* IFO 12827 was characterised to have a very broad substrate specificity, however it lacked enantioselectivity.⁴²

A further possibility for the enzymatic production of amines is offered by imine reductases, which currently provide the simplest method for producing chiral amines through the reduction of prochiral imines with the nicotinamide cofactor, NADPH, achieving theoretical yields of >99% with an optical purity of >99%.

1.3. Imine Reductases (IREDs): Function and Activity

Imine reductases (IREDs) are chemoselective enzymes that are capable of reducing prochiral (**Figure 16**) imines to produce chiral amine products.⁴³

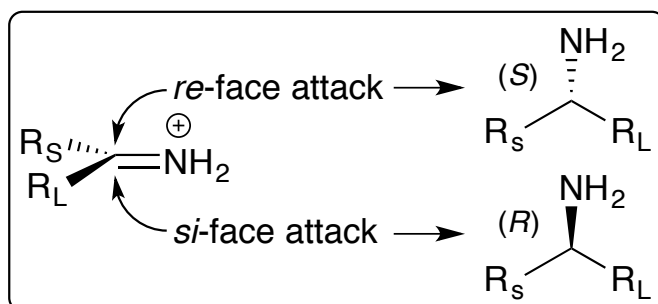


Figure 16 – Scheme of a generic prochiral imine being attacked either on its *re*- or *si*-face to form its respective chiral amines.

1.3.1. Metalloenzymes used for the reduction of imines

Prior to 2012, investigations on IREDs has only just begun, so a lot of effort was put into creating artificial metalloenzymes that were capable of reducing imines. These artificial imine reductases, created by the Ward group, were termed ATHase⁴⁴⁻⁴⁷. These ATHases were created using Biotin-Streptavidin Technology, where an organometallic catalyst is attached to the carboxylic acid end of biotin, which is subsequently bound to streptavidin with an incredibly high affinity. The organometallic end is ideally exposed such that it can coordinate to the imine substrates to allow imine reduction (**Figure 17**). These ATHases appear promising, proving that enzymes that could produce chiral amines through the reduction of prochiral imines. However both the biotin cofactor and the organometallic catalyst are relatively expensive. Therefore ATHases are not readily applicable to industrial biocatalysis.

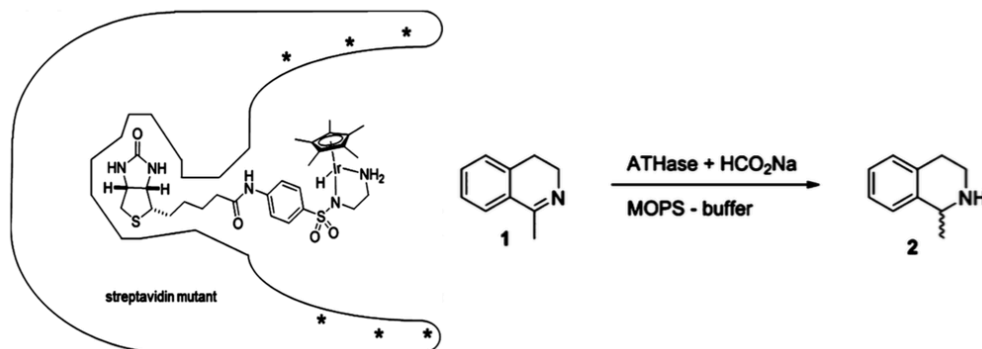


Figure 17 – Reduction of a 1-methyl-3,4-dihydroisoquinoline (MDQ) with an ATHases based on streptavidin. The ATHase resulting from incorporation of a biotinylated cofactor [Cp*Ir(biot-*p*-L)Cl] within streptavidin. Genetic optimisation (*) allows improvement of the catalytic performance of the artificial metalloenzyme. Figure adapted from Schwizer and co-workers (2013).⁴⁴

1.3.2. IREDs with known function

In the literature, there have been several examples of enzymes with imine reduction capabilities as their natural function. Examples of these imine-reducing enzymes will be discussed below.

1.3.2.1. F420-dependent IREDs

Streptomyces achromogenes produces tomaymycin, which is a pyrrolobenzodiazepine with weak antibiotic properties⁴⁸. The biosynthetic gene cluster for tomaymycin was identified and characterised by Li and co-workers in 2009⁴⁹. It was suggested that the nonribosomal peptide synthetases, *tomA* and *tomB*, catalysed the formation of the diazepine ring by joining dihydroxyanthranilic acid with pyrrole (Figure 18). The pyrrole was synthesised from tyrosine by enzymes encoded by five genes and the final biosynthetic step for the pyrrole moiety was proposed to be the reduction of an imine by the gene identified as *tomJ* (Figure 19).

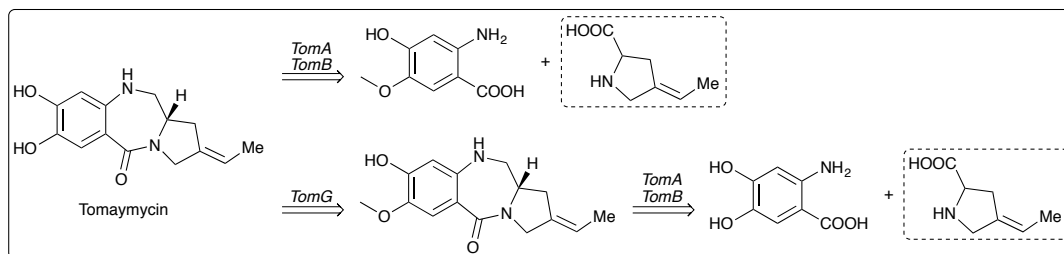


Figure 18 – Retrosynthetic analysis (RSA) for the biosynthesis of tomaymycin in *S. achromogenes*. Core components of the diazepine ring are dihydroxyanthranilic acid and the pyrrole ring, which are joined by a non-ribosomal peptide synthetase containing the domains *tomA* and *tomB*. Figure adapted from Li and co-workers (2009)⁴⁹.

TomJ was suggested to be a dehydrogenase that was dependent on the deazaflavin cofactor, F_{420} , which is mostly found in archaeal methanogens and in actinomycetes⁵⁰. Despite the F_{420} cofactor being structurally similar to FAD and FMN flavin cofactors, redox potential for the F_{420} cofactor⁵¹ is -380 mV, which is more similar to NAD(P)H (-320 to -350 mV)⁵², rather than -230 mV of FMN⁵³ and -160 mV of FAD, however the redox potential for FAD is heavily influenced by the environment it is bound to and can change by ± 100 mV.⁵⁴⁻⁵⁵ Due to the similarity in redox potentials, the F_{420} cofactor is suggested to behave more similarly to NAD(P)H rather than FAD or FMN.⁵⁶⁻⁵⁷

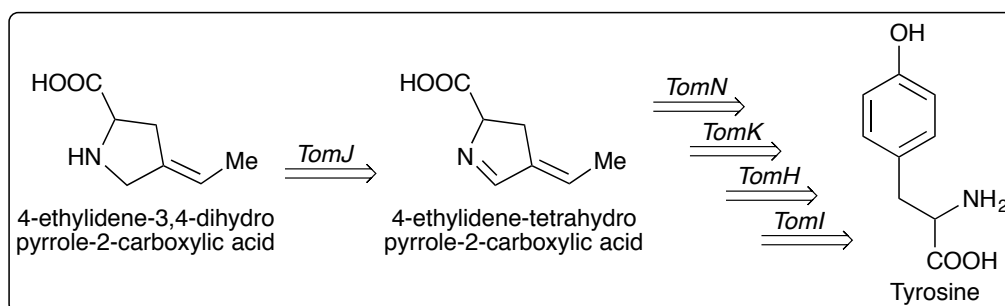


Figure 19 – RSA for the biosynthesis of the pyrrole ring for tomaymycin, which is biosynthesised in five steps from tyrosine by *tomI*, *tomH*, *tomK*, *tomN* and *tomJ*. Figure adapted from Li and co-workers (2009).⁴⁹

Streptosporangium sibiricum is a producer of the antibiotic sibiromycin and a homologue of *tomJ* was identified as *sibT*, an F_{420} -dependent dehydrogenase that also catalyses the reduction of an imine in the pyrrole moiety of sibiromycin⁵⁸. *TomJ* and *sibT* have the potential to be exploited for biocatalysis, however both have not yet been fully characterised nor heterologously expressed.

Heterologous expression may be difficult in hosts that do not produce the F₄₂₀ cofactor, however, if expressed in a host unable to produce the F₄₂₀ cofactor, some F₄₂₀-dependent reductases exhibit cofactor promiscuity and are capable of utilising flavin cofactors, such as FMN⁵⁹. Without further characterisations, *tomJ* and *sibT* are currently only putative imine reductases.

1.3.2.2. Ketimine reductase

Crystallins are the highly abundant soluble proteins of the crystalline vertebrate eye lens that fill the fibre cells of the lens.⁶⁰⁻⁶¹ The human gene *CRYM* encodes for the NADPH-dependent protein μ -crystallin. μ -crystallin has other roles besides in lens biology; μ -crystallin also plays a role in thyroid hormone signalling⁶² and regulation, as such it is also known as an NADP-regulated thyroid-hormone-binding protein (THBP).

μ -crystallin is a key enzyme in the pipecolate pathway, which is the main lysine degradation pathway (**Figure 20**) in the brain; saccharopine pathway dominates in other parts of the body. Hallen and co-workers in 2015 had established that thyroid hormones regulate the pipecolate pathway that implies a reciprocal relationship exists where the bioavailability of thyroid hormones is determined by the activity of μ -crystallin.⁶² As a THBP, unusual expression levels of μ -crystallin have been observed in several neurological disorders. These include increased vulnerability of neurons to mutant huntingtin protein (Huntington's disease).⁶³ Expression of μ -crystallin has also been observed to be down-regulated in schizophrenia⁶⁴ and significantly overexpressed after administration of antidepressants, which implies μ -crystallin could also provide a role in depression.⁶⁵

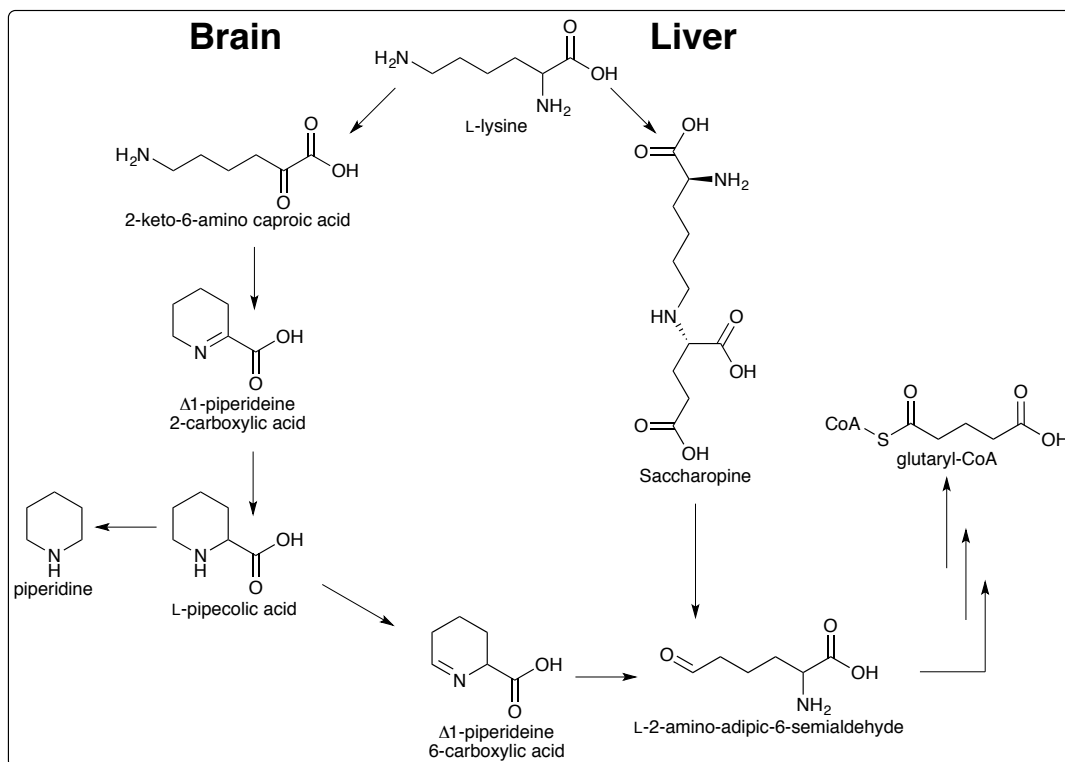


Figure 20 – Schematic of the lysine degradation pathway. In the brain the pipecolic acid pathway takes precedence in the brain (left), where as in the liver the saccharopine pathway dominates (right). Figure adapted from Ijlst and co-workers (2000).⁶⁶

μ -crystallin is also known as ketimine reductase, an NADPH dependent dehydrogenase that has mostly been characterised because of its physiological interest. Ketimine reductase has shown to be the main cytosolic thyroid hormone binding protein⁶⁷ that not only exhibited IRED activity but also reductive amination activity.⁶⁸

A small panel of substrates (**Figure 21**) was screened against ketimine reductase, however these were chosen out of their biochemical and physiological context. As such, ketimine reductase requires further investigation, but has the potential to be as a biocatalyst in synthetic chemistry for the formation of chiral amines.

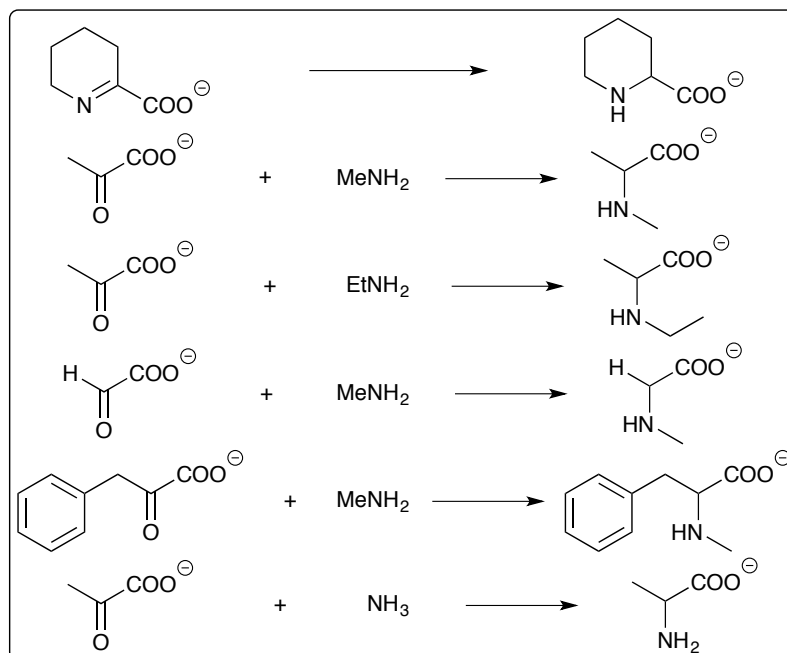


Figure 21 – Biotransformations performed with ketimine reductase. The reductive amination reaction typically contained an α -keto acid-to-amine ratio of 1:6. Reaction rates were reported relative to the imine reduction of the piperidine substrate; rates for reductive amination were considerably lower. Figure adapted from Hallen and co-workers (2015).⁶⁸

1.3.2.3. Thiazolinyl Imine Reductase

Many metalloproteins require iron either structurally or catalytically. Iron is also required in many metabolic pathways, including processes within electron transport, DNA synthesis and the citric acid cycle. Bacteria typically require 0.3 – 1.8 μM of iron for optimal growth and concentrations of iron in soil and in a mammalian host are typically $<0.1 \mu\text{M}$ and $10^{-9} \mu\text{M}$ respectively.⁶⁹ In addition to iron being scarce, iron as Fe(III) is very insoluble, which results in it often being biologically inaccessible. As such, pathogenic bacteria have evolved to scavenge iron through utilising complex chelating systems (**Figure 22**).

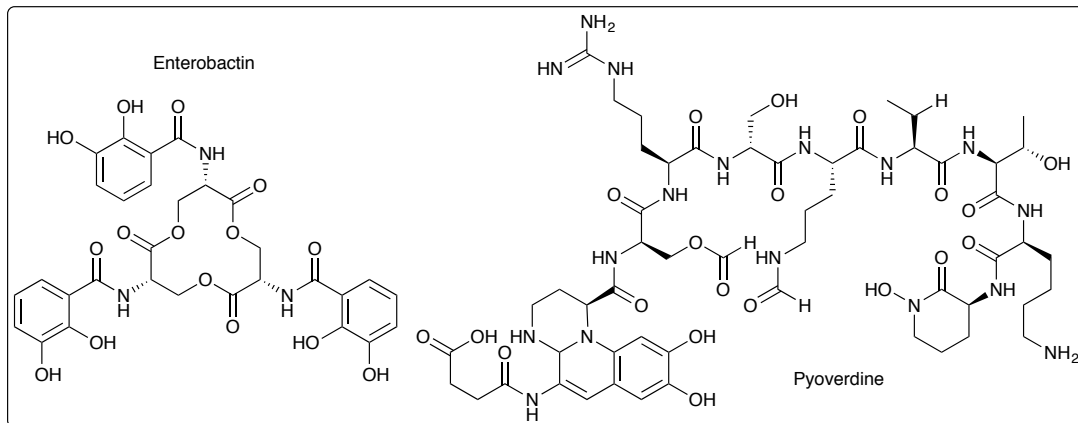


Figure 22 – Examples of siderophores, enterobactin and pyoverdine, highly complex iron chelating systems often biosynthesised by pathogens to scavenge iron from the environment or their host. Figure adapted from Pollack and co-workers (1970)⁷⁰ and Wendenbaum and co-workers (1983)⁷¹ for enterobactin and pyoverdine respectively.

These complex chelating systems are siderophores,⁷² which are molecules that with very high affinity for Fe(III) (**Figure 23**) and are often secreted outside the bacterial cell to scavenge Fe(III) which are then transported back into the cell via uptake.

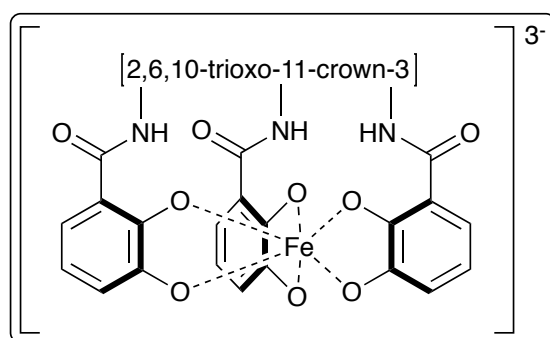


Figure 23 – Simplified schematic of the catechol moieties of enterobactin chelating Fe(III). Figure adapted from Karpishin and co-workers (1992)⁷³.

The siderophore yersiniabactin is biosynthesised by *Yersinia enterocolitica*. The biosynthesis of yersiniabactin requires two large nonribosomal peptide synthetases: high-molecular-weight protein 1 (HMWP1) and high-molecular-weight protein 2 (HMWP2), a protein of unknown function (Irp3), a thioesterase (Irp4) and a salicyl-AMP ligase (Irp5).⁷⁴⁻⁷⁵

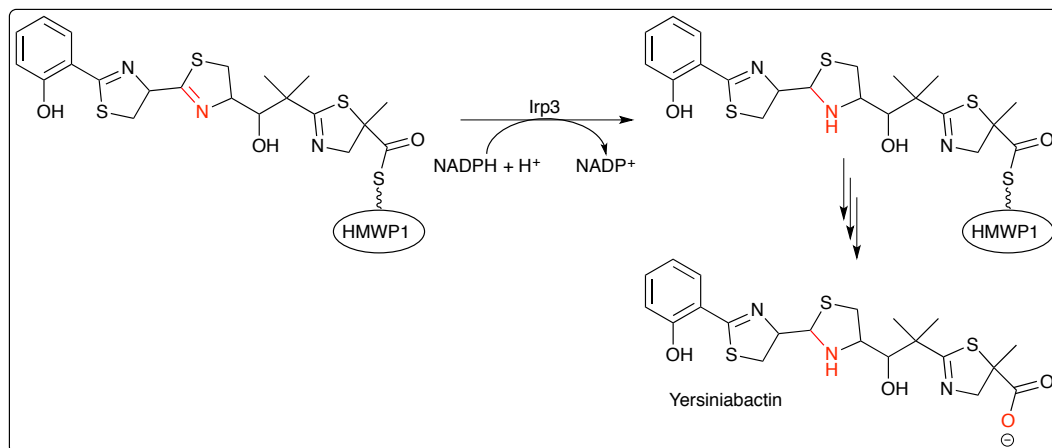


Figure 24 – A simplified scheme for the biosynthesis of the siderophore yersiniabactin in *Yersinia enterocolitica*. A vital biosynthetic step involves the reduction of an imine bond in the central thiazole moiety by Irp3, an NADPH-dependent dehydrogenase. The substrate is covalently bound to the nonribosomal peptide synthetase high-molecular weight protein 1 (HMWP1) via a phosphopantetheinyl bond. Figure adapted from Meneely and co-workers (2012).⁷⁶

The function of Irp3 was not known until recently, when in 2012 Meneely and co-workers identified Irp3 as an IRED.⁷⁶ Irp3 is an NADPH-dependent IRED that reduces the imine bond in the central thiazole moiety of yersiniabactin (**Figure 24**).

Irp3 could potentially be useful for biocatalysis due to its ability to accept very large substrates and its ability to reduce heterocyclic substrates, however its activity could possibly be dependent on the carrier protein HMWP1 delivering substrates to Irp3. As of now, the substrate scope and selectivity of Irp3 have not yet been explored.

1.3.2.4. Dihydrofolate Reductase and Pteridine reductase

Pteridine reductase (PTR1) and dihydrofolate reductase (DHFR) are known IREDs that have been thoroughly investigated and characterised in biochemical studies⁷⁷. Both PTR1 and DHFR are involved in salvaging pterin-like substrates in *Leishmania major* (**Figure 25**).

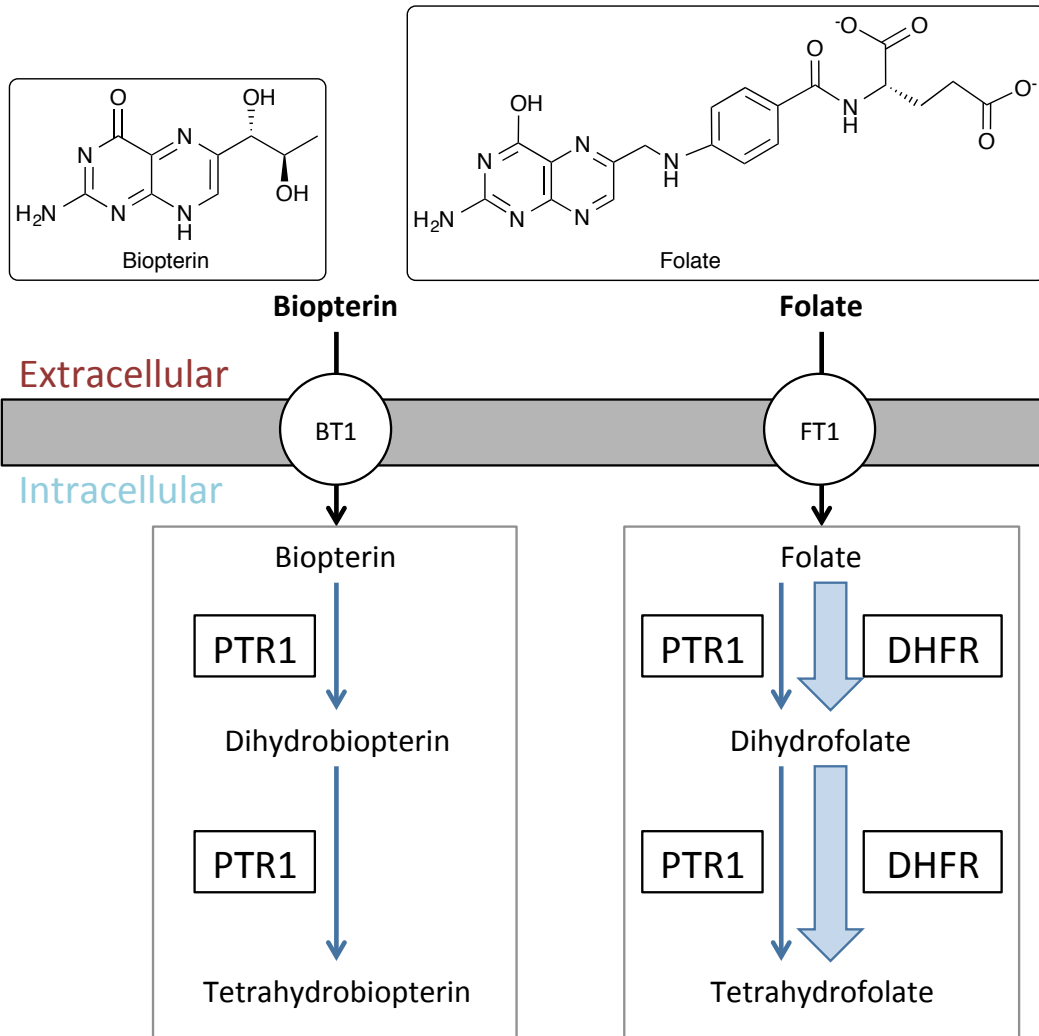


Figure 25 – Scheme of the pteridine scavenging systems in *Leishmania major* with DHFR and PTR1. The pterin substrates are taken up into the cell by folate transporter (FT1) and biopterin transporter (BT1) allowing DHFR and PTR1 to metabolise folate and biopterin respectively. Figure adapted from Cunningham and co-workers (2001).⁷⁷

DHFR and PTR1 salvages pterins by reducing folate to dihydrofolate then further reducing dihydrofolate to tetrahydrofolate (**Figure 26**). Folate is metabolised by DHFR into the nucleoside thymidine and is also necessary for the biosynthesis of purines and pyrimidines each of which is essential to the biosynthesis of DNA, RNA and proteins.⁷⁸

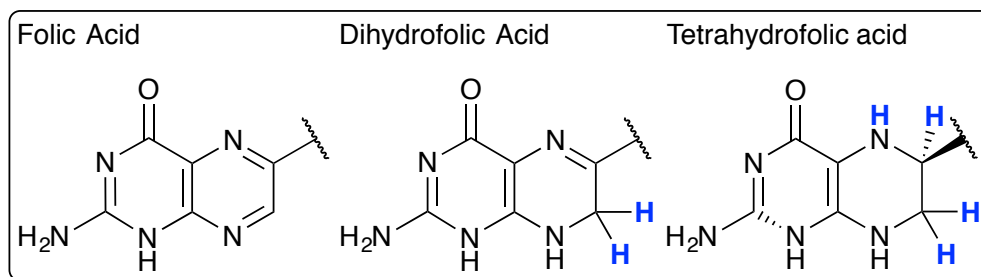


Figure 26 – Simplified scheme of the pterin heterocycle for folate, dihydrofolate and tetrahydrofolate, illustrating points of hydrogenation with blue hydrogens.

DHFR is often studied with the anti-cancer drug methotrexate (**Figure 27**). Methotrexate is also used for the treatment of autoimmune diseases, ectopic pregnancies and for the induction of medical abortions. However in DHFR, methotrexate acts by inhibits DHFR thereby inhibiting the metabolism of folic acid⁷⁸⁻⁷⁹ and therefore the biosynthesis of DNA, RNA and proteins.

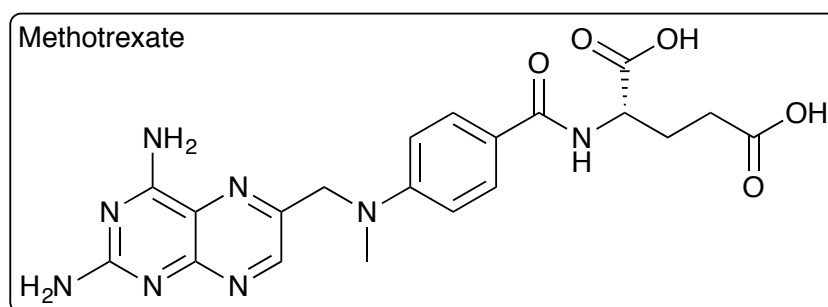


Figure 27 – The chemical structure of methotrexate. Methotrexate is an analogue of folate therefore acts by inhibiting DHFR.

DHFR and PTR1 are highly selective imine reductases for their natural substrates folate and biopterin. In 1990, the synthetic route for producing tetrahydrofolic acid from dihydrofolic acid using glucose and glucose dehydrogenase as a NADPH cofactor recycling system was patented.⁸⁰ This is as tetrahydrofolic acid is a useful intermediate for L-(-)-leucovorin. However, the substrate scopes for DHFR and PTR1 have not been explored outside of physiological context.

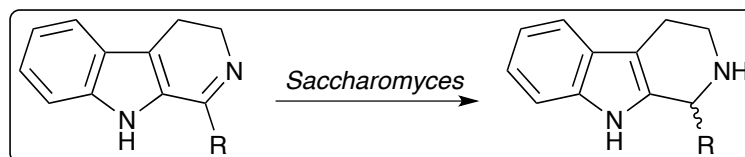
1.3.2.5. Eukaryotic IRED activity

Synthetic organic chemists are more widely adopting the use of biocatalysts, as they are being recognised as being superior catalysts in some respects to organometallic catalysts, and the production of chiral compounds is still very much a challenge. Many different strains of yeasts have been employed in a wide variety of functional group transformations;⁸¹⁻⁸⁴ for example Baker's yeast, *Saccharomyces cerevisiae*, is often used for the reduction of prochiral carbonyls to form chiral alcohols.

In 2010, the Santos group at the University of Talca employed eight different commercially available strains of *Saccharomyces* to reduce β -carboline imines with various substituents attached (**Figure 28**). The yeast that gave highest conversions and enantioselectivity was from *S. bayanus* (Nouveaux).⁸⁵

More recently in 2013, the Santos group used the cell-free extract from red Californian earthworms (*Eisenia foetida*) to reduce the same panel of β -carboline imines as previously mentioned⁸⁵. It was reported that the cell-free extract of *E. foetida* was enantiocomplementary to the whole cell biocatalyst *S. bayanus* (Nouveaux) as well as a more effective biocatalyst in that conversions and selectivity was higher than using the whole cell biocatalyst *S. bayanus* (Nouveaux).⁸⁶

The Santos group has identified strains of *Streptomyces* that exhibit IRED activity. However the genes responsible have not yet been identified, therefore the enzyme cannot be further characterised or engineered for improved activities and substrate spectrum.



Entry	R	% ee	Absolute Configuration	Yield (%)
a	Me	94	(S)-	68
b	Et	93	(S)-	66
c	CH ₂ Cl	92	(S)-	68
d	CH ₂ I	50	(S)-	45
e	iPr	84	(S)-	62
f	4-Pent	85	(S)-	53
g	C ₁₁ H ₂₃	74	(S)-	58
h	C ₁₅ H ₃₁	65	(R)-	62
i	C ₁₇ H ₃₃	62	(R)-	60
j	C ₁₇ H ₃₁	64	(R)-	61
k	Ph	88	(R)-	52
l	2-(OH)Ph	97	(R)-	48
m	Bn	79	(R)-	62
n	Nicotinic	75	(R)-	57

Figure 28 – Reduction of β -carbolines using different strains of *Saccharomyces*. The best performing strain was from *S. bayanus* (Nouveaus). Figure adapted from Espinoza-Morga and co-workers (2010).⁸⁵

1.3.2.6. Conversion of (S)-reticuline to (R)-reticuline in the opium poppy

Morphine, an opiate alkaloid drug, first discovered and isolated from the opium poppy (*Papaver somniferum* L.) more than 200 years ago by Friedrich Sertürner, is still one of the most potent pain relief drugs used today. Starting from tyrosine, 7 biosynthetic steps are required to produce a key intermediate (S)-reticuline, followed by a further 9 biosynthetic steps to produce morphine. Until recently, one of the key steps in the biosynthesis of morphine, the epimerisation of (S)-reticuline to (R)-reticuline, was unknown (Figure 29). In 2015, three independent groups⁸⁷⁻⁸⁹ discovered the last key step in the biosynthesis of morphine.

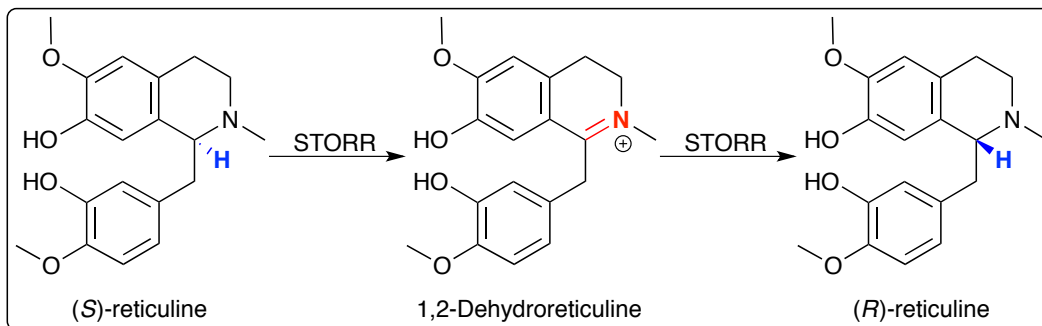


Figure 29 – The epimerisation of (*S*)-reticuline to (*R*)-reticuline by the gene identified as STORR ((*S*)- to (*R*)-reticuline). Figure adapted from Winzer and co-workers (2015).⁸⁷

The key step, the epimerisation of (*S*)-reticuline, was performed by a P450-NADPH-oxidoreductase fusion protein (**Figure 30**), which has been named STORR ((*S*)- to (*R*)-reticuline). After the epimerisation, only (*R*)-reticuline is further metabolised to morphine (**Figure 31**).



Figure 30 – Schematic of the gene for STORR, where a P450 is fused to an oxidoreductase domain. Figure adapted from Winzer and co-workers (2015).⁸⁷

STORR initially oxidises the amine (*S*)-reticuline to the iminium species 1, 2 - dehydroreticuline, which is then selectively reduced to the amine (*R*)-reticuline (**Figure 29**) using NADPH. As such, STORR is an IRED.

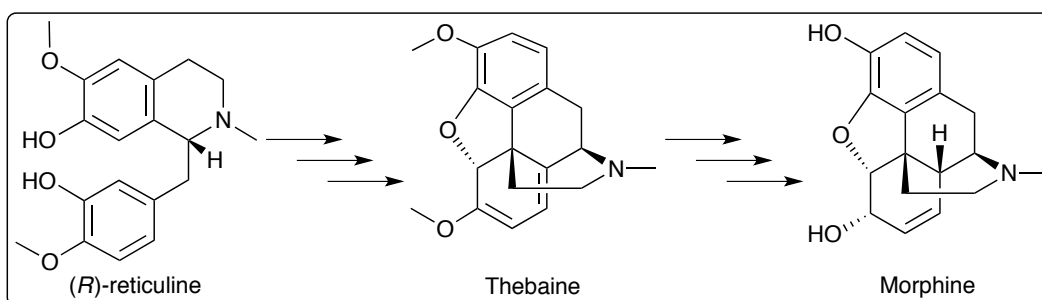


Figure 31 – Simplified biosynthetic pathway of the production of morphine. Key intermediates (*R*)-reticuline and thebaine lead to morphine. Figure adapted from Kries and co-workers (2016).⁹⁰

STORR would be interesting to structurally characterise due its attractive reactivity. STORR could be engineered and exploited to be a useful racemisation tool for biocatalysis. However, full length STORR can only be expressed in yeast, which as an engineering platform, is relatively slow compared to *E. coli*.

1.3.2.7. Opine Dehydrogenases

Opines (**Figure 32**) are typically the condensation products of amino acids with keto acids or sugars⁹¹. Opines have a variety of physiological roles; in invertebrates they are formed in muscle tissue upon anaerobic stress similar to lactic acid in mammals, whereas in parasitic bacteria such as *Agrobacteria*, opines are nutrients which are the cause of crown gall tumour in plants⁹¹⁻⁹².

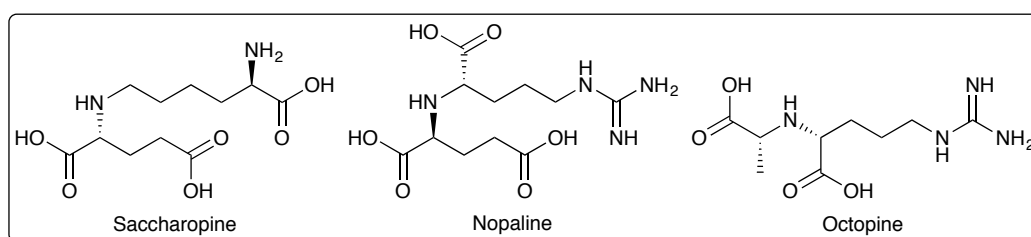


Figure 32 – Examples of opines. Saccharopine is the product of the metabolism of lysine. Nopaline is the product obtained from the condensation of glutamate and arginine. Octopine is the product obtained from the condensation of arginine and alanine.

Opine dehydrogenases (OpDHs) are NADPH-dependent enzymes that catalyse the reductive amination of amino acids with keto acids for the production of opines (**Figure 33**).

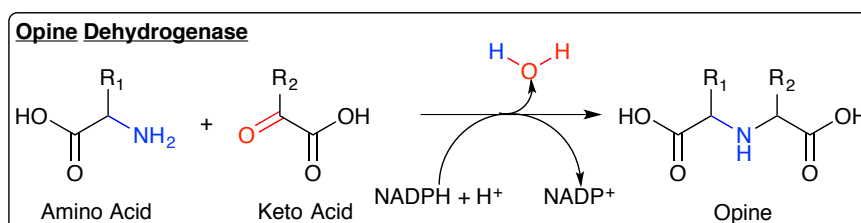


Figure 33 – Reductive amination of an amino acid with a keto acid by an OpDH, producing opines. Figure adapted from Vladimirov and co-workers (2014).⁹¹

Reports have shown OpDHs to be efficient catalysts for the transformation of physiological substrates, that being amino acids with keto acids.⁹²⁻⁹⁶ More recently in 2013, Codexis patented OpDHs,⁹⁷ which have been engineered to accept a much wider range of substrates that are not limited to amino acids and keto acids. These engineered OpDHs are much more useful as a tool in biocatalysis.

1.3.2.8. EC 1.5.1.48 Imine reductases (IREDs): Structure, Function and Activity

In the recent years, IREDs have proven to be promising biocatalysts for the production of chiral amines by the reduction of prochiral imines.^{20, 43, 98-107} Using only NADPH as the cofactor, theoretical yields of 100%, with 100% stereoselectivity can be achieved.

The first investigations on biocatalytic IREDs were performed by Mitsukura and co-workers,⁹⁸⁻¹⁰⁰ who had identified two enantiocomplementary NADPH-dependent dehydrogenases from *Streptomyces* sp. that were capable of reducing the prochiral imine 2-methyl-1-pyrroline (2MPN) (**Figure 34**) to the (*S*)- and (*R*)-amine. These were subsequently named SSIRED and SRIRED respectively.

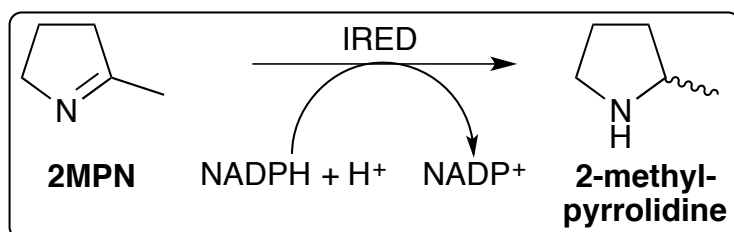


Figure 34 – 2-methyl-1-pyrroline (2MPN) is a 5 membered ring imine substrate which is often used to identify IRED activity.

Following the initial work performed from Mitsukura and co-workers, many groups began investigating homologues of SSIRED and SRIRED, and the substrate scopes of IREDs. It was initially thought that these IREDs had a very narrow substrate scope, however this was mostly due to the inherent instability of non-cyclic imines and the lack of commercially available imines. The substrate 2MPN has been the model substrate for IRED investigations, however, certain IREDs are more active towards six-membered rings^{20, 43} and even fused ring systems over

five-membered rings, meaning that investigations based solely on 2MPN have made the IREDs appear to be poor biocatalysts.¹⁰⁸

The first crystal structure of an IRED was from *Streptomyces kanamyceticus*, *SkIRED*, (PDB entry: 3ZHB).¹⁰⁸ It was identified as a homologue to *SRiRED* in which it shares 50% sequence similarity with a further 19% strongly similar residues. Through biotransformations, *SkIRED* was shown to be (*R*)-selective towards the substrate 2MPN.

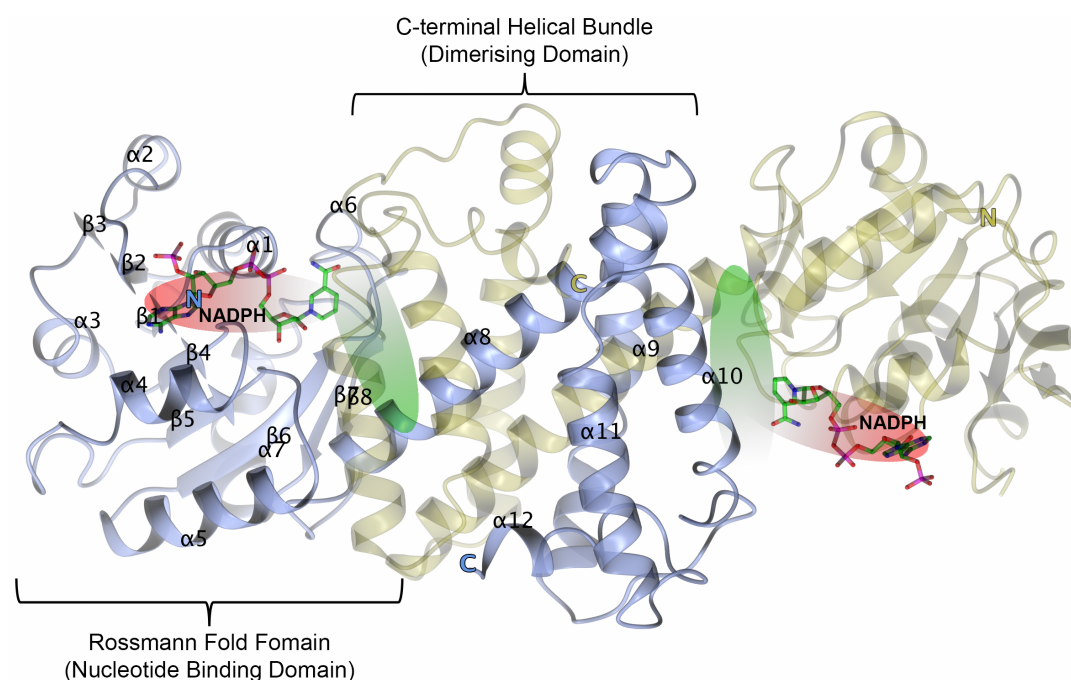


Figure 35 – The structure of *SkIRED* in complex with NADPH (PDB entry: 3ZHB). The structure composes of two monomers, subunit A in blue and subunit B in gold, dimerising as a homodimer. Each monomer has an N-terminal Rossmann fold domain for nucleotide binding highlighted in red and a C-terminal helical bundle which is involved with dimerising, with the putative active site formed in the cleft between the N- and C-terminal domains, highlighted in green. Secondary structure has been assigned for subunit A, with the N-terminal Rossmann fold with alternative β -sheets and α -helicies, starting at $\beta 1$, followed by $\alpha 1$, through to $\alpha 7$ and $\beta 8$, with the C-terminal helical bundle starting at $\alpha 8$ through to $\alpha 12$, which dimerises with subunit B's helical bundle.

The structure is composed of a homodimer (**Figure 35**). Each monomer consists of an N-terminal Rossmann Fold (nucleotide binding) domain, similar to other Rossmann Fold domains of other dehydrogenases such as alcohol dehydrogenases, which is used for the binding of the NADPH cofactor, and a C-terminal helical bundle, which is involved, with domain sharing, with a reciprocal monomer from the adjacent monomer to form the dimer. The nicotinamide ring

of NADPH protrudes into the cleft formed between the N-terminal domain and the C-terminal domain of the reciprocal monomer; the cleft is thought to be where the active site is situated. The putative active site is lined with hydrophobic residues.

Once the first structure of *SkIRED* had been solved,¹⁰⁸ structural comparisons using the DALI server¹⁰⁹ had identified that *SkIRED* was structurally similar to a carbonyl reductase, 3-hydroxyisobutyrate dehydrogenase (HIBDH). In 2005, Lokanath and co-workers proposed a mechanism for HIBDH¹¹⁰ in which the substrate 2-methyl-3-oxopropanoic acid is reduced by NADPH followed by proton abstraction from a putative catalytic lysine, Lys165 (**Figure 36**).

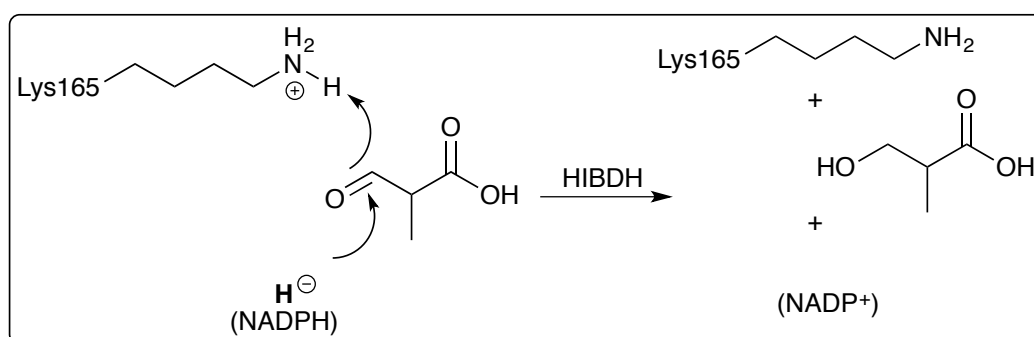


Figure 36 – Proposed mechanism for the carbonyl reduction by HIBDH. Initially NADPH delivers the hydride to the carbonyl-carbon which reduces the carbonyl, followed by proton abstraction from the protonated catalytic Lys165. Figure adapted from Lokanath and co-workers (2005)¹¹⁰.

Upon superposing *SkIRED* with HIBDH an aspartate residue (Asp187) overlaps with Lys165 in HIBDH (**Figure 37**). Mutations to *SkIRED* were made in an attempt to identify the active site. The mutation D187A was made and it was shown that the mutation to alanine abolished activity.

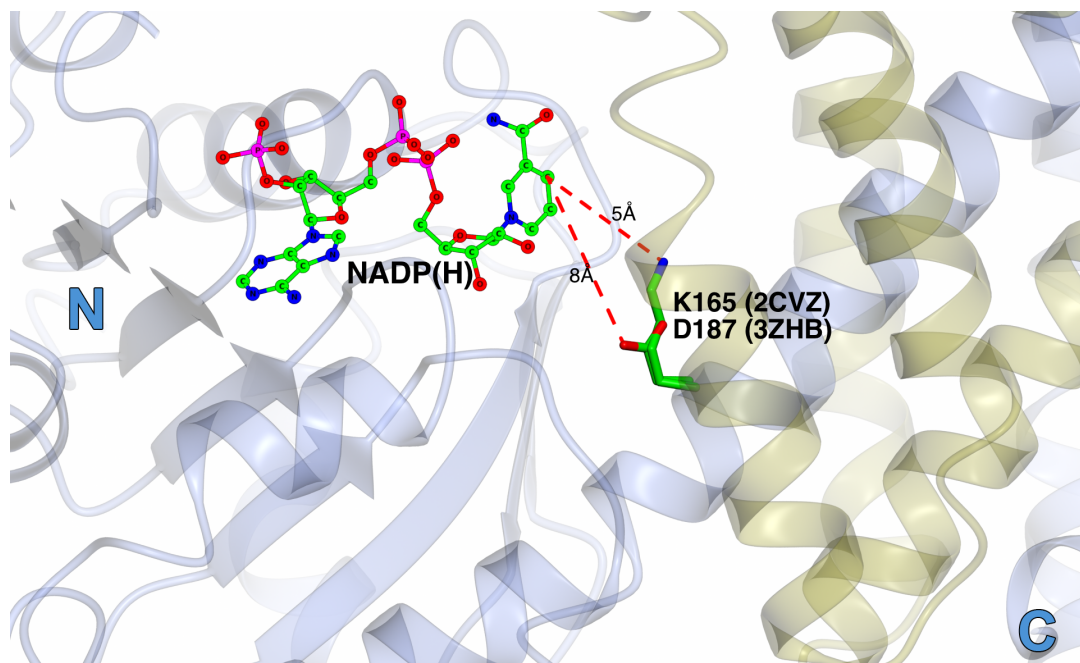


Figure 37 – Peptide chains of subunits A, in blue with the N- and C-termini labelled, and B, in gold, of *SkIRED* (PDB entry: 3ZHB) were superposed by secondary structure matching (SSM) on to subunits A and B of HIBDH (PDB entry: 2CVZ) which is not shown. Only the ligand, NADP(H), from *SkIRED*, in circles, and the secondary structure of *SkIRED* are shown for illustrative purposes only. The superimposition illustrates the putative catalytic Lys165 from HIBDH overlapping with Asp187 in *SkIRED*, with distances to the C4N of NADPH highlighted as 5 Å and 8 Å for HIBDH and *SkIRED* respectively.

The mechanism that was proposed for *SkIRED* was similar to HIBDH based on a generic NADPH-dependent dehydrogenase mechanism; where a hydride is delivered from the C4N position of NADPH to the electron deficient carbon of the imine bond, followed by proton abstraction from Asp187 (**Figure 38**). However the distance between the C4N of NADPH and the hydroxyl of D187 in *SkIRED* was an average overall all subunits of 8 Å, whereas the distance between the C4N of NADPH and the amine of lysine in HIBDH was 5 Å. This could mean that the aspartate in *SkIRED* is too far away for direct protonation and that imine reduction could be water assisted.

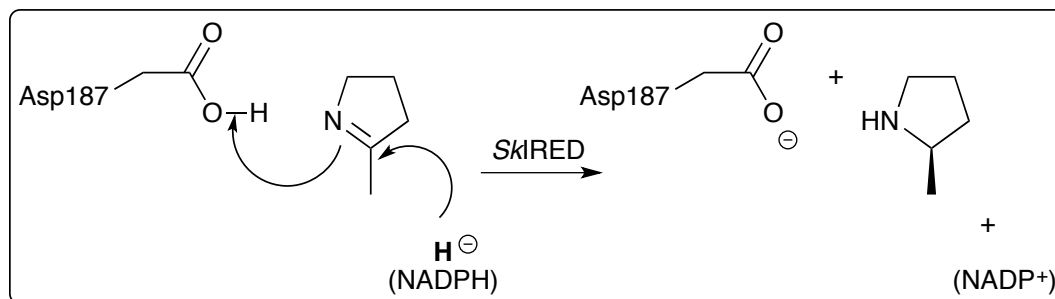


Figure 38 – The proposed mechanism for *SkIRED* and other aspartate-IREDs was based on the mechanism for general NADPH-dependent dehydrogenases. Figure adapted from and Rodriguez-Mata co-workers (2013)¹⁰⁸.

The crystal structure for the (*S*)-selective IRED *SSIRED* was solved by Huber and co-workers¹⁰¹. The overall secondary structure displayed high homology to *SkIRED*. Superposing *SkIRED* with *SSIRED* revealed a tyrosine residue (Tyr169) where Asp187 is found in *SkIRED*. The mutation Y169F was made for *SSIRED* by Elizabeth Wells (personal communication, in our group) and the mutation to Y169F abolished activity. A mechanism for *SSIRED* was proposed that was analogous to *HIBDH* and (**Figure 39**); after hydride delivery from the C4N of NADPH to the carbon of the imine bond, a proton is abstracted from a neighbouring tyrosine. The distance observed between the C4N of NADPH to the hydroxyl of Tyr169 in *SSIRED* is approximately 5 Å, which is significantly closer to that of *SkIRED*. Therefore, Tyr169 in *SSIRED* is more likely to be a direct proton donor, compared to Asp187 in *SkIRED*.

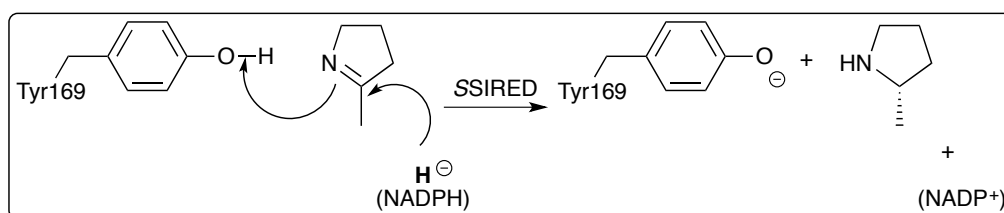


Figure 39 – The proposed mechanism for *SSIRED* and other tyrosine-IREDs based on the mechanism for *SkIRED* is analogous to other NADPH-dependent dehydrogenases such as short chain alcohol dehydrogenases.

The Stuttgart IRED database, created by the Hauer group, contains potential IRED sequences, where the majority of IREDs are of bacterial origin and only a handful of sequences are of eukaryotic fungal origin^{102, 105}. The database contains sequences of putative IREDs that are not limited to being aspartate or

tyrosine-IREDs. The database increased the number of potential IRED homologues available to investigated, with sequences that are 30% similar to SSIRED and SRIRED, in the hope of finding IREDs with more diverse substrate specificity.

2. Aim of the project

The aim of the PhD is to determine the active site of IREDs and thereby enable inference on the mechanism. Prior to starting the PhD, the mechanism of IREDs were unclear. By performing structural studies of IREDs by single crystal X-ray crystallography to obtain high-resolution crystal structures, the active site could be identified. The identification of the active site would allow for structure based evolution programmes to further evolve IREDs for better enantioselectivity, substrate specificity, potentially develop new functionality, and enhance reaction rates to be suitably utilised on an industrial scale. The individual project aims are as follows:

1) Identify (S)-selective IRED homologues and obtain the crystal structures

At the time of starting the PhD, a high-resolution crystal structure of an (S)-selective IRED had not yet been published. The first aim was to identify close homologues of SSIREN, order the synthetic genes, clone, express and purify the IREDs. Once purified, the IREDs will be characterised by biotransformations, kinetics assays, thermal shift assays and structurally characterised by single crystal X-ray crystallography to determine structural differences between (S)- and (R)-selective IREDs.

2) Obtain crystal structures of the IRED from *Amycolatopsis orientalis* (AoIRED) in complex with xenobiotic ligand

The gene for AoIRED was obtained from G. Aleku, Manchester Institute of Biotechnology (Manchester, UK). AoIRED had been characterised as a superior IRED compared with previously investigated IREDs. The second aim was to structurally characterise AoIRED in an attempt to identify the differences in active sites with previously structurally characterised IREDs. Ligand complexes will be pursued using xenobiotic ligands forming a series of binary and ternary complexes in an attempt to gain mechanistic insight.

3) Obtain ligand complexes of the reductive aminase (RedAm) from *Aspergillus oryzae* (AspRedAm)

The gene for AspRedAm was obtained from G. Aleku. AspRedAm was characterised not only to be an outstanding IRED, but it was also characterised to be able to catalyse reductive amination utilising a substrate carbonyl-to-amine ratio of 1:1. Not only was the functionality and activity of AspRedAm remarkable, AspRedAm also displayed an incredibly wide substrate specificity for ketone acceptors and amines donors, not limited ammonia. However, AspRedAm was relatively unstable once isolated as pure protein. As such, the third aim was to structurally characterise AspRedAm by single crystal X-ray crystallography to determine structural and mechanistic differences between IREDs and RedAms. By obtaining a series of ternary and quaternary xenobiotic ligand complexes, it would hopefully identify catalytic residues that differ from standard IREDs such as SSIREd and SRIREd, which could aid in identifying more stable RedAms.

4) Identification and characterisation of bacterial RedAms

Despite AspRedAm being an incredibly diverse biocatalyst, the only drawback was that it was relatively unstable and often aggregates within 24 h. This may be due to the fact that AspRedAm is from a eukaryote and may require glycosylation to stabilise it. Ongoing collaborative work between M. Sharma (in our group) and J. Mangas at the Manchester Institute of Biotechnology (Manchester, UK), had characterised all available fungal IREDs and demonstrated that they were also all RedAms. Therefore the fourth aim was to explore the possibility of a bacterial RedAm with the potential of being more stable than AspRedAm. A more stable RedAm would be an incredibly valuable biocatalyst for synthetic chemistry and industrial biocatalysis.

Imine reduction and reductive amination are powerful methods for obtaining chiral amines. Therefore, by identifying and characterising synthetically useful IREDs and RedAms; these enzymes can be used for as atom efficient biocatalysts for the production of chiral amines.

3. Methods

3.1. Cloning

3.1.1. Primer design

Primer design is vital for successful polymerase chain reaction (PCR). The main reason for PCRs failing is usually due to bad primer design where the primer is unable to anneal to the target gene or vector.

Primers for cloning were either designed manually (In-Fusion® and Gibson Assembly®) or designed using HiTel Primer Design (T4 based LIC). Primers for mutagenesis were either designed manually (inverse PCR, In-Fusion® and Gibson Assembly®) or designed using Primer X and Agilent QuikChange primer design (QuikChange).

HiTel Primer Design: <http://bioltfws1.york.ac.uk/cgi-bin/primers.cgi>

Primer X: <http://www.bioinformatics.org/primerx/>

Agilent Technologies: <http://genomics.agilent.com/primerDesignProgram.jsp>

The list of primers can be found in **Table 44 (6.1)**.

3.1.2. Polymerase Chain Reaction (PCR)

As each polymerase differs, protocols were followed according to the manufacturer's user manual (**Table 1** and **Table 2**).¹¹¹ Typically for fragments smaller than 3000 bp, KOD Hot Start DNA Polymerase (Merck Millipore) was used. For fragments larger than 3000 bp where a high fidelity DNA polymerase was required, Pfu DNA polymerase (ThermoFisher Scientific) was used (**Table 3** and **Table 4**).¹¹²

Table 1 – A standard reaction setup for KOD Hot Start Polymerase. Adapted from KOD Hot Start DNA Polymerase user manual.¹¹¹

Component	Volume (μL)	Concentration
10X buffer	5	1X
25 mM MgSO ₄	3	1.5 mM
dNTPs (2 mM each)	5	0.2 mM (each)
Forward Primer (10 μM)	1.5	0.3 μM
Reverse Primer (10 μM)	1.5	0.3 μM
Template (10 ng/μL)	1	0.2 (ng/μL)
KOD polymerase	1	0.02 U/μL
ddH ₂ O	Varied	N/A
Reaction Volume	50	

Table 2 – Typical thermocycling conditions for KOD Hot Start Polymerase. Adapted from KOD Hot Start DNA Polymerase user manual.¹¹¹

Step	Target Size			
	< 500 bp	500-1000 bp	1000-3000 bp	> 3000bp
1. Polymerase activation	95°C for 2 min			
2. Denature	95°C for 20 s			
3. Annealing	Lowest Primer T _m °C for 10 s			
4. Extension	70°C for 10 s/kb	70°C for 15 s/kb	70°C for 20 s/kb	70°C for 25 s/kb
5. Repeat steps 2-4	18-35 cycles			
6. Final Extension	70°C for 5-15 min			
7. Cool	10°C ∞ hold			

Table 3 – A standard reaction setup for Pfu DNA Polymerase. Adapted from Pfu DNA Polymerase user manual.¹¹²

Component	Volume (μL)	Concentration
10X buffer with MgSO ₄	5	1X
dNTPs (2 mM each)	5	0.2 mM (each)
Forward Primer (10 μM)	1.5	0.3 μM
Reverse Primer (10 μM)	1.5	0.3 μM
Template (10 ng/μL)	1	0.2 (ng/μL)
Pfu polymerase	1	0.02 U/μL
ddH ₂ O	Varied	N/A
Reaction Volume	50	

Table 4 – Typical thermocycling conditions for Pfu DNA Polymerase. Adapted from Pfu DNA Polymerase user manual.¹¹²

Step	Target Size	
	< 2000 bp	> 2000 bp
1. Polymerase activation	95°C for 3 min	
2. Denature	95°C for 30 s	
3. Annealing	Lowest Primer T _m -5°C for 30 s	
4. Extension	72°C for 2 min/kb ; 72°C for 3 min/kb	
5. Repeat steps 2-4	18-35 cycles	
6. Final Extension	70°C for 5-15 min	
7. Cool	10°C ∞ hold	

Three-step PCRs are typically used (**Figure 40**) where the temperature used for denature, annealing and extension are different. However, in cases where the primers have high GC-content or are poorly designed, a two-step PCR can be used where the annealing and extension temperatures are the same.

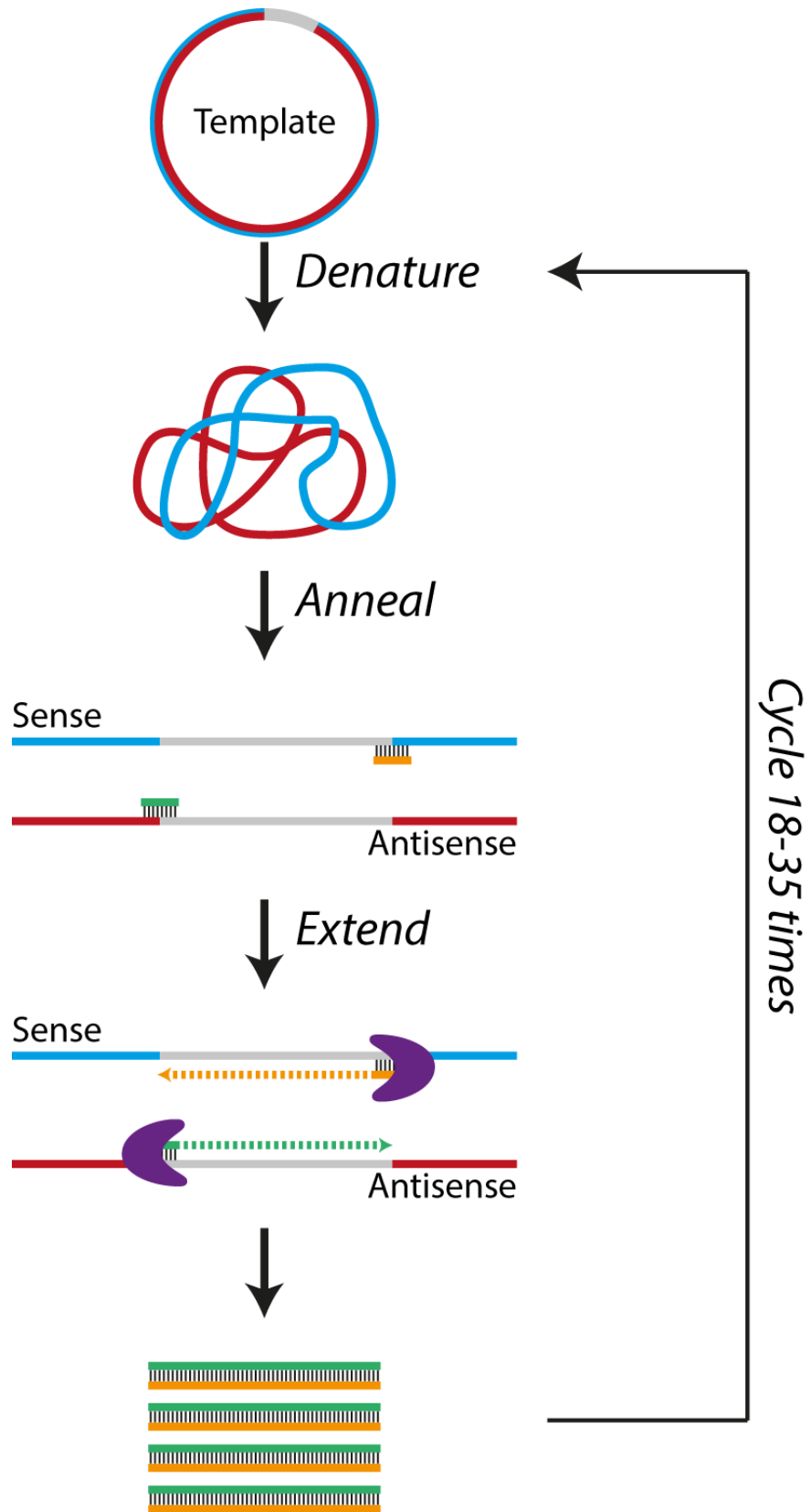


Figure 40 – Illustration of a typical PCR. Initially the double stranded template DNA is denatured into single strands. The reaction is cooled to allow complementary sequences to anneal by forming hydrogen bonds. The DNA polymerase is then able to recognise the primed regions and amplify the appropriate regions creating double strand fragments of the target DNA. The reaction is cycled 18 to 35 times.

3.1.3. Touchdown PCR

The most common formula used to calculate the melting temperature (T_m) of primers is very much a simplified estimate (**Equation 1**). The equation does not factor in complex parameters, such as kinetics of DNA duplex formation (primer-primer and primer-template hybridisations), which is determined by the sequence of nucleotides, ssDNA concentration, solvents and buffer. Therefore the T_m obtained from **Equation 1** is often inaccurate. A consequence of inaccurate T_m can lead to unsuccessful PCR's as the annealing temperature is often based on the lowest primer T_m .

$$T_m = 4(G + C) + 2(A + T)$$

Equation 1 – The simplified equation for the calculation of the primers' melting temperature (T_m)

One way to overcome the guesswork of the T_m is to perform a touchdown PCR (TD-PCR's).¹¹³ TD-PCR allows for a simple and rapid method of optimising PCR's without performing many gradient PCR's with separate reactions. The reaction setup follows the user manual of the polymerase of choice, however the thermal cycling parameters are adapted. The PCR is split into three phases (**Table 5**).

Table 5 - Example of extending a standard PCR protocol to a TD-PCR, to be used on a 1000 bp fragment with KOD Hot Start DNA Polymerase. Table adapted from Korbie and co-workers (2008).¹¹³

	Step	Temperature	Time
Phase 1	1. Polymerase Activation	95°C	2 min
	2. Denature	95°C	20 s
	3. Anneal	$T_m + 10^\circ\text{C}^*$	10 s
	4. Extension	72°C	15 s
	Repeat steps 2-4	10-15 cycles	
Phase 2	5. Denature	95°C	20 s
	6. Anneal	T_m or $(T_m - 5^\circ\text{C})$	10 s
	7. Extension	72°C	15 s
	Repeat steps 5-7	20-25 cycles	
Termination	8. Final Extension	72°C	5 min
	9. Cool	10°C	∞

The initial 10 cycles of a PCR are the most crucial for generating the desired fragment. The first phase of the TD-PCR targets the T_m 10°C higher than calculated from **Equation 1**. At higher temperatures it is more difficult to anneal less specific

ssDNA, therefore the chances of ssDNA with high homology to correctly anneal is significantly greater. In the first phase, the annealing temperature (T_a) is reduced between 1-0.5°C per cycle for 10-15 cycles (**Figure 41**). The purpose of Phase 1 is to exploit the exponential nature of PCR in that, once the primers have correctly primed and amplified the desired product, further priming to the newly amplified product is more likely to succeed as the temperature is reduced with unspecific priming greatly reduced. Once Phase 1 has generated sufficient desired product, the purpose of Phase 2 is to mass amplify more desired product by reducing the T_a and allow for easy DNA priming and annealing by which point any chance of unspecific priming and annealing has been mostly eliminated. The final phase, Termination, is to complete and polish any incomplete amplification, and finally to stop any enzymatic activity by cooling.

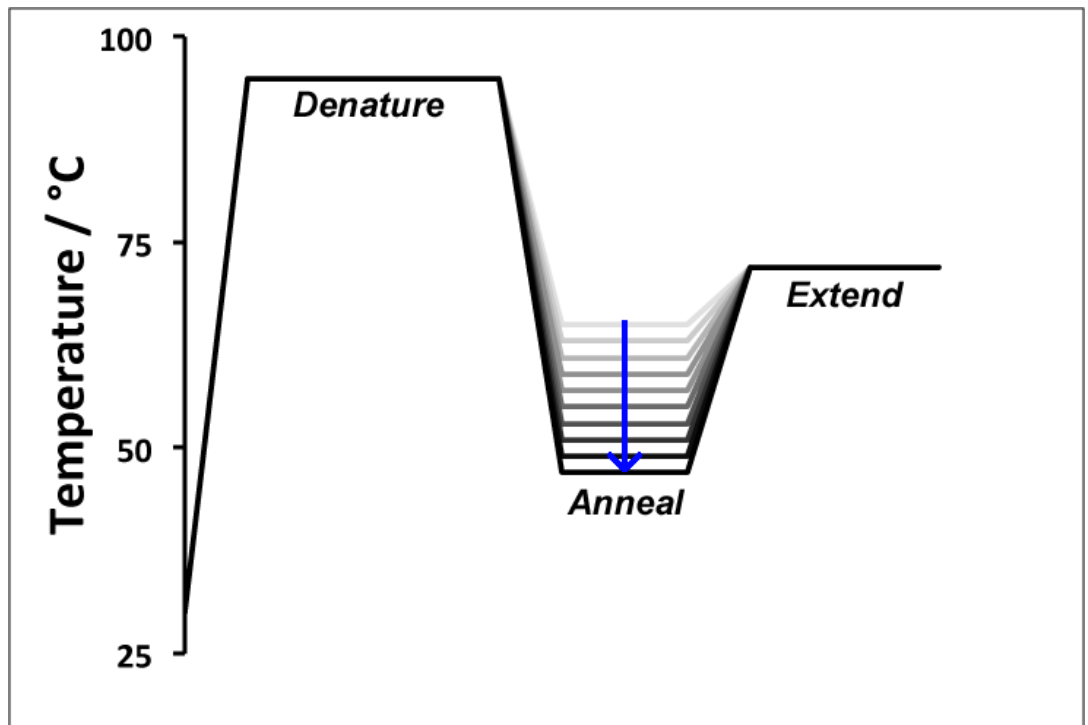


Figure 41 - Scheme of Phase 1 of a TD-PCR, where the annealing temperature is reduced by 1-0.5°C per cycle for 10-15 cycles in the thermocycler.

3.1.4. Fragment analysis by agarose gel electrophoresis

Fragment analysis by agarose gel electrophoresis is a common method for both analysis and purification of DNA fragments. A 1% (w/v) agarose gel was prepared with 1x TAE buffer and used for all fragment analysis. SYBR®Safe DNA

stain was used in a 1:10,000 dilution added to the molten agarose. Agarose gels were run by electrophoresis at 110 V for 50 min with 1x TAE buffer to allow for good resolution between DNA fragments. The 1 kb DNA ladder (NEB) was used as an internal standard and ladder. Agarose gels are visualised using a Safe Imager™ Blue-Light transilluminator. Correct DNA fragments were excised and purified following GenElute™ Gel Extraction Kit (Sigma-Aldrich).

3.1.5. Ligation independent cloning (LIC) qualified T4 DNA polymerase cloning

A wide range of vectors adapted for ligation independent cloning (LIC) was developed by the Structural Proteomics in Europe (SPINE) consortium. The purpose was to develop vectors for high-throughput cloning to easily incorporate genes with generic primers into vectors with different purification and solubility tags.¹¹⁴⁻¹¹⁵ The York members of SPINE adapted the multiple cloning site (MCS) of pET-28a, a vector that fuses an N-terminal hexa-histidine (*6His*) tag to the gene of interest, for LIC; one vector with a cleavable *6His*, pET-YSB LIC3C vector, and one vector with a non-cleavable *6His* tag, pET-YSB LIC.¹¹⁶

Only the pET-YSB LIC3C was used for the preparations of N-terminal *6His* tagged protein. The pET-YSB LIC3C vector was first linearised by *BseRI* (**Table 6**), and subsequently purified by gel extraction. The linearised vector was then digested by LIC-qualified T4 DNA polymerase (**Table 7**) and was cleaned up by gel extraction. After purification, the vector was ready to be used in annealing reactions.

Table 6 – Components required for linearisation of the pET-YSB LIC3C vector by *BseRI*.

Component	Quantity
Vector DNA	50 µg
<i>BseRI</i>	50 µL
10x CutSmart® buffer	100 µL
ddH ₂ O	to 1 mL
Incubate at 37°C for 1 h 50 min	

Table 7 – Components required for the digest of the pET-YSBLIC3C vector to generate sticky-ends.

Component	Quantity
<i>BseRI</i> linearised vector	4.0 pmol
10x T4 buffer	40 µL
dTTP (25 mM)	40 µL
DTT (100 mM)	20 µL
LIC-qualified T4 DNA polymerase	8.0 µL
ddH ₂ O	to 400 µL
1) Incubate at 22°C for 30 min	
2) Inactivate at 75°C for 20 min	

Primers with the appropriate generic overhands adapted for the pET-YSBLIC3C vector were designed using the HiTel Primer design (3.1.1). Following PCR (3.1.2) and product clean up by gel extraction; the insert was digested by LIC-qualified T4 DNA polymerase to generate the stick-ends (Table 8), which allows for easy annealing (Table 9) to the prepped vector.

Table 8 – Components required for the digest of insert to generate the reciprocal sticky-ends required to anneal to the treated pET-YSBLIC3C vector.

Component	Quantity
Insert	0.2 pmol
10x T4 buffer	2 µL
dATP	2 µL
DTT	1 µL
LIC-qualified T4 DNA polymerase	0.4 µL
ddH ₂ O	to 20 µL
1) Incubate at 22°C for 30 min	
2) Inactivate at 75°C for 20 min	

Table 9 – Components required for the annealing of the treated insert and treated vector. This reaction will generate a nicked plasmid.

Component	Quantity
Digested insert reaction	2 µL
T4 treated vector (~50 ng/µL)	1 µL
1) Incubate at 22°C for 1 h	
2) Add 1 µL EDTA (25 mM)	
3) Incubate at 22°C for 1 h	

After annealing, the nicked plasmid was transformed by *E. coli* to repair and replicate the plasmid. The protocol is summarised in Figure 42.

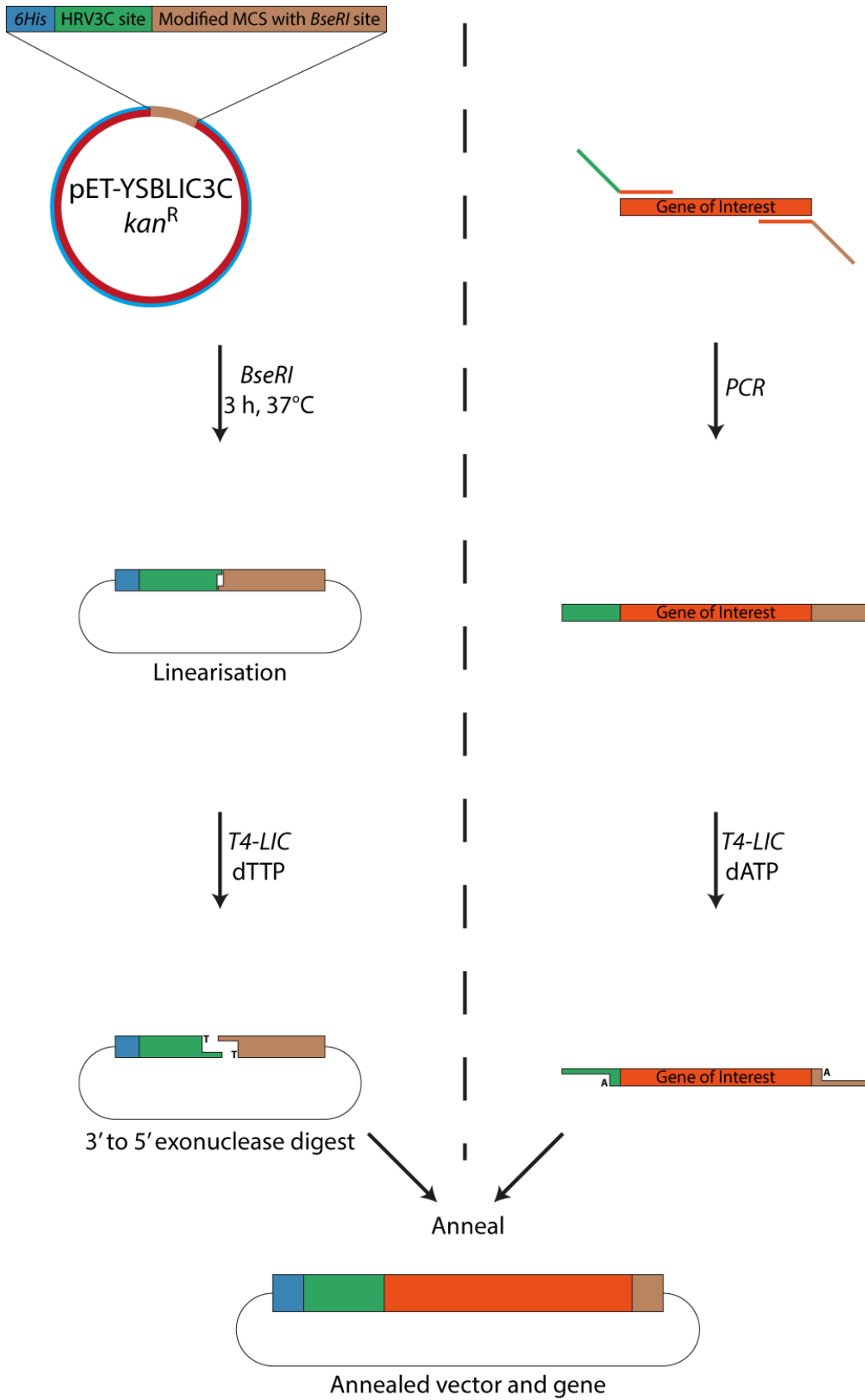


Figure 42 – Summary of the LIC protocol using LIC-qualified T4 DNA polymerase. The left describes the linearisation protocol of the pET-YSB LIC3C vector. The right describes the treatment of the insert after PCR.

3.1.6. In-Fusion® Cloning

Sequence and ligation independent cloning (SLIC) methods are alternatives to the LIC method described in 3.1.5. Two commercially available SLIC kits are In-Fusion® (Clontech® Laboratories, Inc.) and Gibson Assembly® (NEB). Both In-Fusion® and Gibson Assembly® are based on the poxvirus DNA polymerase, which has 3' to 5' proofreading to recognise the crucial 15 bp sequence homology and has 3' to 5' exonuclease activity.¹¹⁷⁻¹¹⁹ The main advantage of the aforementioned SLIC kits over the LIC-qualified T4 DNA polymerase based LIC is that the kits require much fewer steps. The linearisation of the pET-YSBLIC3C vectpr with *BseRI* will yield a maximum of 75% linearised vector followed by further losses from sequential gel extractions. Therefore LIC-qualified T4 DNA polymerase based LIC is very inefficient.

Using the SLIC kits, the vector of choice first needs to be linearised at the MCS either by inverse PCR or enzymatic restriction digest. Linearising by inverse PCR may result in a higher yield, however enzymatic restriction digest with one or two enzymes may give a cleaner digest. After linearisation, the linearised vector is purified by gel extraction and is ready for annealing reactions.

Primer design and quality is crucial for the success of the In-Fusion reaction. Primers for In-Fusion cloning are designed with the parameters outlined in **Table 10**.

Table 10 – The parameters required for primers used in In-Fusion cloning. Adapted from In-Fusion® HD Cloning Kit user manual.¹¹⁹

- | |
|---|
| <ol style="list-style-type: none">1) 5' ends of primers contain 15 bp homology to the target DNA (vector or another insert).2) 3' ends of the primer should:<ol style="list-style-type: none">a) be gene specificb) be 18-25 bp in length excluding the 15 bp homology regionc) have a GC-content between 40-60%d) have Tm's of 61°C ± 3°C and primer pairs with ≤ 4°C differencee) not contain identical runs of nucleotides3) Avoid complementary primers4) Desalted oligonucleotides are generally sufficient |
|---|

After the insert was amplified with the appropriate primers, the In-Fusion reaction (**Table 11**) can take place. At which point, the In-Fusion enzyme digests the linearised vector's overhangs to match the 5' 15 bp homologous regions of the

amplified insert and anneals the homologous regions. After the In-Fusion reaction has taken place, the nicked plasmid is ready to be repaired and recovered in *E. coli*.

Table 11 – The components required for the In-Fusion reaction (annealing and linearised the vector to the insert). Adapted from In-Fusion® HD Cloning Kit user manual.¹¹⁹

Component	Reaction
Purified PCR fragment	10-100 ng ^(a)
Linearised vector	50-200 ng ^(b)
5x In-Fusion HD Enzyme Premix	1 µL
ddH ₂ O	to 5 µL
Incubate at 50°C for 15 min	
(a) <0.5 kb: 10-50 ng, 0.5-10 kb: 50-100 ng	
(b) <10 kb: 50-100 ng, >10 kb: 50-200 ng	

The In-Fusion cloning protocol is summarised in **Figure 43**.

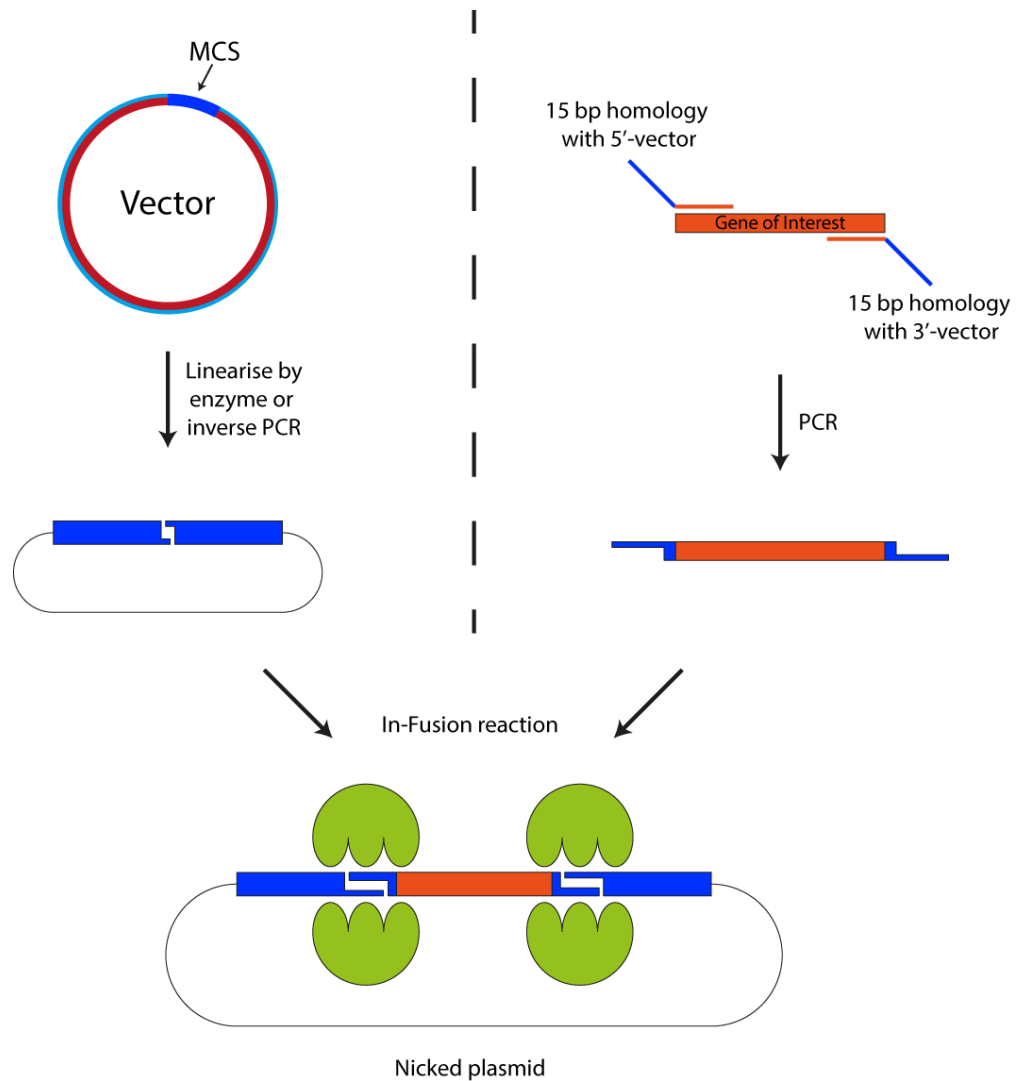
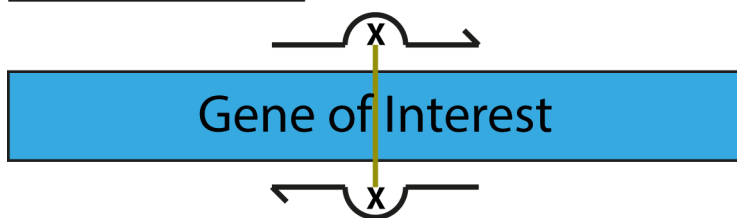


Figure 43 – Summary of the In-Fusion cloning protocol. The left describes the linearisation protocol of any vector. The right describes the insert preparation by PCR.

3.1.7. Site Directed Mutagenesis

Site directed mutagenesis (SDM) is a method to evolve an enzyme, which uses mutant primers and PCR to amplify a plasmid. SDM can be used to either enhance an enzymes' activity or eliminate activity for mechanistic studies. There are two common methods for performing SDM, QuikChange and In-Fusion; both of which employ inverse PCR. However, the primer design for both methods is different (**Figure 44**). As such, the PCR protocol should follow the manufacturer's user manual.

QuikChange



In-Fusion

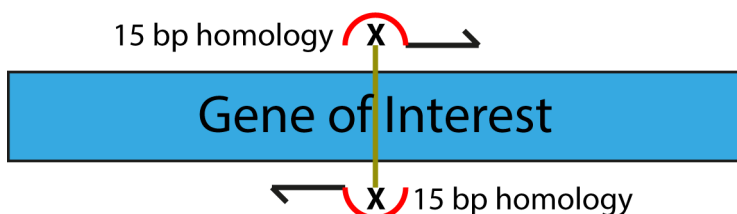


Figure 44 – Two different strategies for SDM. The primer design strategies are different for both QuikChange and In-Fusion.

The QuikChange method requires long (~40 bp) overlapping primers with the site of mutation in the centre of the primers. These mutant primers should allow for annealing to the wild type template and a high-fidelity DNA polymerase should be used to ensure there are no further mutations other than the ones introduced *via* primers.

The In-Fusion SDM primers are often shorter than those used with QuikChange. The mutation is contained in the 5' end of the primer along with the 15 bp homology region. However rather than the 5' region being vector specific, as in standard insert PCR, the 5' region is specific to the region to be mutated.

In both cases, after the inverse PCR is complete, the wild type template needs to be removed from the PCR mixture. *E. coli* encodes for the gene containing deoxyadenosine methylase (DAM) which recognises the sequence 5'-GATC-3' and methylates the adenine in that sequence.¹²⁰⁻¹²¹ Therefore DNA that has been produced by *E. coli* has been marked by methylation. The wild type template produced by *E. coli* and can be digested by *DpnI* (**Table 12**). *DpnI* recognises the sequence 5'-G(Me-A)TC-3' and cleaves on the 3'-side of the methylated adenine. DNA produced by polymerases through PCR is deficient in methylation and can therefore be differentiated from wild type template.

Table 12 – Components required for the template digestion reaction by *DpnI*.

Component	Quantity
PCR reaction	20-50 μ L
<i>DpnI</i>	1 U
Incubate at 37°C for 3 h	

In the case of QuikChange based SDM, after the template has been digested, the mutant dsDNA is ready to be repaired and recovered in *E. coli*. In the case of In-Fusion SDM, the mutant dsDNA must be annealed using the In-Fusion enzyme, as described in **Table 11**, prior to repair and recovery in *E. coli*.

Both QuikChange based SDM and In-Fusion SDM have advantages and disadvantages. QuikChange based SDM is cheaper in that it does not need an enzyme kit, however the primers are complementary, therefore the chance of the mutant primers annealing to each other forming primer-dimer complexes are significantly higher. In-Fusion SDM is more likely to have a higher chance of success as the primers only have the 15 bp homology, therefore the chance of primer-template duplex formation is significantly higher, however, the In-Fusion® HD Cloning Kit is expensive.

3.1.8. Bacterial Transformation

A transformation describes the uptake of a plasmid by *E. coli*. There are a wide variety of commercially available strains of *E. coli*. The strains used in this thesis were NovaBlue Singles™ (Novagen) for cloning, to repair and recover nicked plasmids, and BL21(DE3) (NEB) for protein expression.

LB-Agar plates were prepared in advance with the appropriate antibiotics, here kanamycin (30-50 μ g mL⁻¹, final concentration) unless otherwise specified. Standard heat shock transformation (**Table 13**) was used throughout this thesis. After heat shock transformation, the LB-agar plates were incubated at 37°C overnight.

Table 13 – Standard protocol for heat shock transformation.

Heat Shock Transformation
1) Mix 1 μ L plasmid (upto 5 μ L if recovering nicked plasmid) with 20 μ L competent cells in a 1.5 mL eppendorf tube
2) Incubate on ice for 30 min
3) Heat shock for 45 s at 42°C
4) Incubate on ice for 5 min
5) Add 1 mL super optimal broth prewarmed to 37°C
6) Incubate at 37°C for 1 h with shaking in an orbital shaker 180-300 rpm
7) Pellet cells by centrifugation (5000 x g, for 2 min)
8) Discard 800 μ L media
9) Resuspend pelleted cells in remaining 200 μ L media
10) Plate cells on to LB-agar with appropriate antibiotic

3.1.9. Double Digest

To confirm the presence of an insert of the desired weight has been correctly annealed into the vector, a double digest (**Table 14**) is performed with restriction digest enzymes with sites usually at the 5'-end and 3'-end of the insert. For example, for an insert cloned into the pET-YSBLIC3C vector, enzymes *NcoI* and *NdeI* would be used for a double digest.

Table 14 – Protocol for performing a double digest.

Component	Volume (μL)
Plasmid (50-300 ng/ μ L)	7
5'-end restriction digest enzyme	1
3'-end restriction digest enzyme	1
CutSmart Buffer (10x)	1
Incubate 3 h at 37°C	

After performing the double digest, fragments are analysed by gel electrophoresis. The analysis should result in expected bands at weights corresponding to the linearised vector (~5.0 kbp) and insert (~1.0-1.5 kbp). Occasionally the gene may contain a cut site that is recognised by one of the restriction digest enzymes, as such, smaller fragments would be expected.

3.2. Gene Expression

Prior to expressing genes, a pre-culture is required. Colonies containing the desired plasmid were picked from the LB-agar plate and used to inoculate 5 mL autoclaved LB media containing the appropriate antibiotic in a 50 mL tube. This pre-culture was incubated at 37°C overnight, shaking in an orbital shaker at 250 rpm.

The pre-culture, at a 1:100 dilution, was used to inoculate a larger volume of autoclaved LB media in baffled Erlenmeyer flasks. The flasks were then incubated at 37°C shaking in an orbital shaker at 250-300 rpm, until the optical density at 600 nm (OD_{600}) reached 0.6-1.0 Abs. At this point the culture was induced with 1 mM final concentration isopropyl- β -D-1-thiogalactopyranoside (IPTG). After induction, the culture was incubated at 16°C for up to 18 h shaking at 250-300 rpm in an orbital shaker.

After 18 h of protein expression, the cells were harvested by centrifugation (5000 x g for 20 min) with the residual media discarded. Cells were resuspended in water and pelleted again by centrifugation (10,000 x g for 20 min) into 50 mL falcon tubes with the residual water discarded. Cell pellets were flash frozen in liquid nitrogen and stored at -20°C.

3.2.1. Protein analysis by sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE) and native-PAGE

A common method for analysing protein samples and purity is by sodium dodecyl sulphate polyacrylamide electrophoresis (SDS-PAGE) either on 10% or 12% polyacrylamide gels (**Table 15**), which were run at 200 V for 50 min. The low molecular weight ladder (BioRad) was used as internal standard and ladder. Loading buffer dye is added to the protein sample to a final concentration of 1x and is denatured at 94°C for 5 min prior to loading the samples onto the gel.

Table 15 – Components required for SDS-PAGE.

2x Loading Buffer Dye
60 mM Tris-HCl pH 6.8
10% (v/v) Glycerol
2% (w/v) SDS
0.05% (w/v) bromophenol blue
0.715 mM β -mercaptoethanol
Load 20 μ L denatured samples onto gel
1x Running Buffer
191.82 mM Glycine (electrophoresis grade)
24.76 mM Tris-base
Resolving Gel Buffer
1.5 M Tris-HCl pH 8.8
0.4% (w/v) SDS
Stacking Gel Buffer
0.5 M Tris-HCl pH 6.8
0.4% (w/v) SDS

Two SDS Gels		
Resolving Layer	10%	12%
ddH ₂ O (mL)	4	3.2
Resolving Gel Buffer (mL)	2.5	
Acrylamide (mL)	3.3	4.2
Mix with 100 μ L 10% APS and 8 μ L TEMED, then pour. Wait 5 min prior to pouring stacking layer.		
Stacking Layer		
ddH ₂ O (mL)	3.2	
Stacking Gel Buffer (mL)	1.5	
Acrylamide (mL)	0.5	
Mix with 50 μ L 10% APS, 8 μ L TEMED and 10 μ L 0.1% (w/v) bromophenol blue solution, then pour.		

The components in **Table 15** can be adapted to native-PAGE, where SDS and β -mercaptoethanol are replaced with water, bromophenol blue solution is not used in the stacking layer and most importantly, protein samples are not denatured. Native-PAGE gels were run at 150 V for 1 h 30 min. Both SDS-PAGE and native-PAGE gels were stained and visualised with Coomassie blue staining solution.

3.3. Protein Purification

Unless stated otherwise the buffers used were as follows in **Table 16**:

Table 16 – Buffers components used in purification.

Purpose	Buffer	Components
Resuspension Buffer	A	500 mM NaCl, 50 mM Tris-HCl pH 8.0, 20 mM imidazole and 10% (v/v) glycerol
Purification Buffer 1	B	500 mM NaCl and 50 mM Tris-HCl pH 8.0
Purification Buffer 2	C	500 mM NaCl, 50 mM Tris-HCl pH 8.0 and 500 mM imidazole

Cell pellets were resuspended with 50 mL buffer A for every 4 g of wet cell pellet. DNase I was added to the cell suspension to a final concentration of 10 $\mu\text{g mL}^{-1}$. One cOmplete™ EDTA-free protease inhibitor cocktail (Roche) tablet was added for every 50 mL of lysate. The cell suspension was lysed by cell disruption using a constant cell disruptor (Constant Systems Ltd) at 25 kpsi. The cell lysate was clarified by centrifugation (50,000 $\times g$ for 30 min).

The clarified lysate was loaded at 5 mL min⁻¹ using a peristaltic pump onto a 5 mL HisTrap FF Crude column (GE Healthcare Life Sciences) charged with Ni₂SO₄ and equilibrated with 5 column volumes (CVs) of buffer A. Once the lysate was loaded, the column was attached to a fast protein liquid chromatography (FPLC) purifier, such as an ÄKTA™ pure (GE Healthcare Life Sciences). A flow rate of 2.5 mL min⁻¹ was used throughout the purification, collecting 1.8 mL fractions in a 96-well deep well plate (VWR International), whilst monitoring the absorbance at 280 nm (A_{280}), using the protocol outlined in **Table 17** and **Figure 45**.

Table 17 – Protocol used for purifying 6His-tagged protein by Ni-affinity chromatography.

Ni-purification
1) 5 CV wash at 20 mM Imidazole
2) 5 CV wash at 50 mM Imidazole
3) 20 CV gradient to 500 mM Imidazole
4) 5 CV wash at 500 mM Imidazole

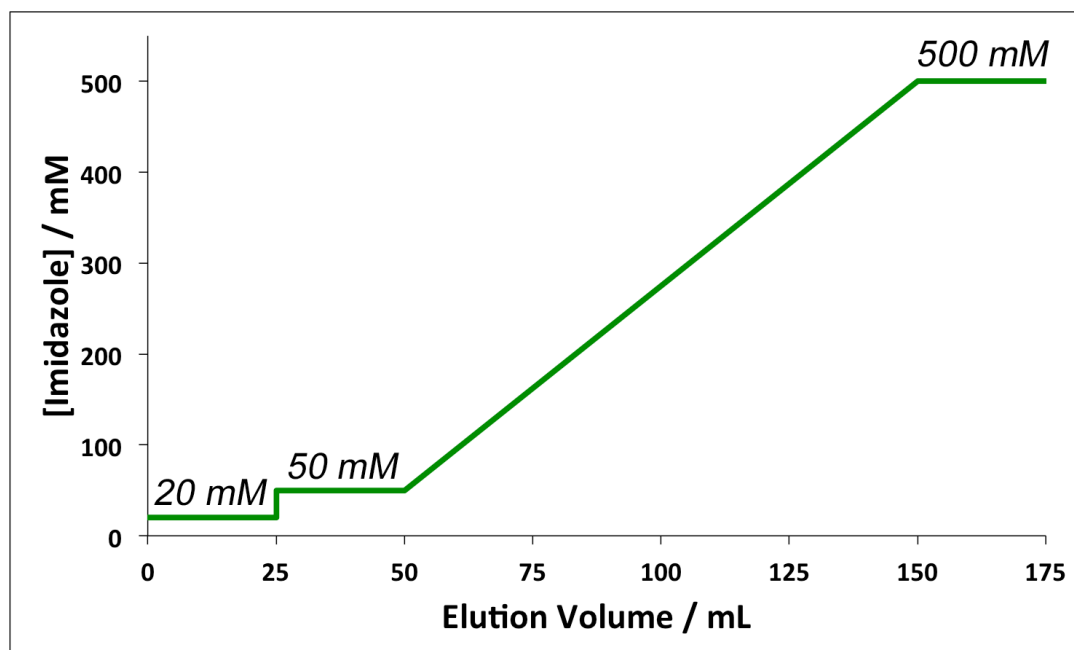


Figure 45 – Scheme of the gradient used for purification of 6His-tagged proteins by Ni-affinity chromatography.

The aromatic amino acids tryptophan, tyrosine and phenylalanine contribute to the proteins' absorbance at A_{280} and can provide an estimate of protein concentration. For a protein that expresses well, a large peak will appear at A_{280} during the 20 CV gradient elution. Fractions from the elution gradient are taken for purity analysis by SDS-PAGE (3.2.1). The purest fractions were taken for further purification by size exclusion chromatography (SEC).

The HiLoad 16/600 Superdex 75 PrepGrade (GE Healthcare Life Sciences) SEC column was used for all size exclusion purifications. Prior to purifying by SEC, the SEC column was equilibrated with 1 CV of buffer B. The fractions obtained from Ni-affinity purification were concentrated to a volume less than 2 mL. The protein was then loaded onto the SEC column by injecting from a 2 mL loop and purified by running 1 CV buffer B at a flow rate of 1 mL min^{-1} , collecting 1 mL fractions in a 96-well deep well plate, whilst monitoring the absorbance at 280 nm. Fractions were again assessed for purity by analysis by SDS-PAGE and native-PAGE (3.2.1) where necessary.

3.4. Assays

3.4.1. Kinetic Assay by UV-VIS

Many substrates or co-factors exhibit different absorption spectra with varying λ_{\max} values in their oxidised or reduced states. Exploiting this, kinetic assays are often performed on UV-spectrophotometers with standard methods of obtaining kinetic parameters using the Beer-Lambert Law (**Equation 2**). Spectrophotometers are used to monitor the change in absorbance (either continuously or at set intervals) over time.

$$dA = \varepsilon \cdot dc \cdot l$$

Equation 2 – Beer-Lambert's Law, where A is the absorbance, ε is the molar extinction coefficient, c is the concentration of the sample and l is the pathlength of the reaction cell.

The rate of a reaction can also be defined as velocity (**Equation 3**).

$$v = \frac{dc}{dt} = \frac{\left(\frac{dA}{\varepsilon \cdot l}\right)}{dt}$$

Equation 3 – The definition of velocity, v, is equal to the change in concentration, c, over the change in time, t. The change in concentration can also be expressed using the rearrangement of the Beer-Lambert law.

The rate of reaction can be plotted against a given substrate concentration (**Figure 46**), which allows for the Michaelis-Menten kinetics model to be used to model the enzyme kinetics. The Michaelis constant K_M is the substrate concentration at half the maximum velocity, $\frac{1}{2}V_{\max}$. Therefore a low value for the K_M indicates that a small concentration of substrate is required before saturation, V_{\max} ; the opposite argument for large K_M values is also true. As such, K_M is occasionally used to describe an enzymes affinity for a given substrate; lower values equate to higher affinity.

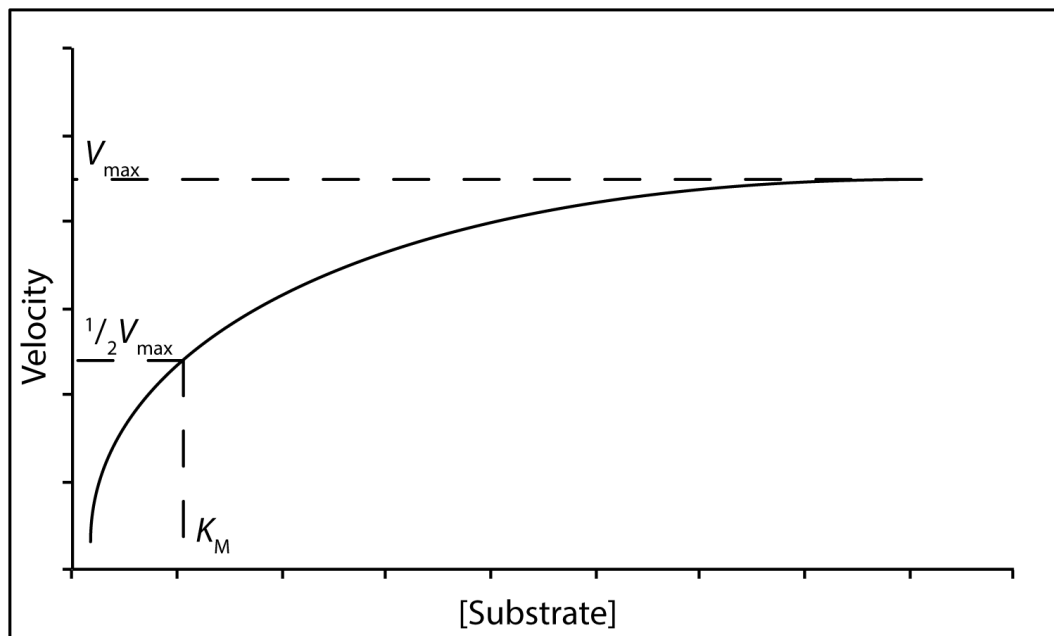


Figure 46 – A scheme of a plot of velocity (reaction rate) versus substrate concentration that obeys Michaelis-Menten kinetics. V_{max} describes the maximum velocity of the reaction (substrate saturation), K_M (at $\frac{1}{2}V_{max}$) describes the enzymes affinity for a given substrate.

Once V_{max} has been obtained, the turnover number (k_{cat}) can be determined. k_{cat} (**Equation 4**) is defined as the maximum number of chemical conversions of substrate molecules per second that a single catalytic site will turn over for a given enzyme (or catalyst) site concentration ($[E]_T$).

$$k_{cat} = \frac{V_{max}}{[E]_T}$$

Equation 4 – The definition of k_{cat} , the turnover number.

The ratio between k_{cat} and K_M (**Equation 5**) is often defined as the catalytic efficiency. High catalytic efficiency (where k_{cat} is greater than K_M) suggests a large proportion of bound substrate is converted before dissociation. Where as, low catalytic efficiency (where k_{cat} is smaller than K_M) is the opposite argument.

$$\frac{k_{cat}}{K_M}$$

Equation 5 – Catalytic efficiency, the ratio between k_{cat} and K_M .

A standard NADPH depletion assay was used to assess imine reduction parameters (**Table 18**). Imine stock solutions were made by serial dilution in ethanol. All reactions were performed in triplicates on a Cintra UV-spectrophotometer (GBC Scientific) at 340 nm ($\epsilon_{\text{NADPH}} = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$). After all components from **Table 18** were added into a quartz cuvette, the cuvette was inverted 6 times to gently mix the reaction. Each reaction was monitored for 5 min.

Table 18 – Typical components used in a reaction for assessing kinetics parameters.

<i>Imine Reduction</i>	
Component	Typical Quantity
Purified Enzyme	100 μL or 250 μL (5 mg mL^{-1} or 2.5 mg mL^{-1} respectively)
NADPH	0.1 mM final concentration
Imine Substrate	0.05 mM, 0.1 mM, 0.5 mM, 1.0 mM, 2.0 mM, 4.0 mM, 10.0 mM and 20 mM
Buffer	up to 1 mL (50 mM Tris-HCl pH 8.0)

The assay can be modified to accommodate an NADP^+ oxidation assay, simply by changing the cofactor to NADP^+ and the substrate to substrate of interest.

3.4.2. Biotransformation Assay with Substrates of Interest

Biotransformations describe the altering of a functional group transformation of a substrate by a microorganism or enzyme. After quenching the biotransformations and extracting the products and reactants, samples are often analysed by techniques such as thin layer chromatography (TLC), gas chromatography (GC), high performance liquid chromatography (HPLC), mass spectrometry (MS), nuclear magnetic resonance (NMR) and polarimetry or a combination of the techniques.

Substrates that were not soluble in water or buffer were typically made in dimethylformamide (DMF) to a stock concentration of 250 mM. Such that, when the stock solution was diluted to 5 mM for the biotransformation, the final concentration of DMF was 2 % (v/v).

To monitor imine reduction, a small panel of substrates (**Figure 47**) was screened to confirm IRED activity. Components (**Table 19**) were mixed in a 2 mL

Eppendorf tube. The reaction proceeded at 30°C in an orbital shaker at 250 rpm for up to 24 h. The reaction was subsequently quenched with 30 µL 10 M NaOH and extracted with 1 mL of dichloromethane (DCM) by vortexing for 1 min. The aqueous and solvent layers were separated by centrifugation (16,600 x *g* for 2 min). For HPLC and achiral GC samples, 500 µL of the extracted reaction was directly analysed. For chiral GC samples, 100 µL of the extracted reaction was acetylated with 30 µL triethylamine and 15 µL acetic anhydride in a tapered GC vial prior to analysis. The retention times and protocols are listed in **Table 19** and **Table 20** respectively.

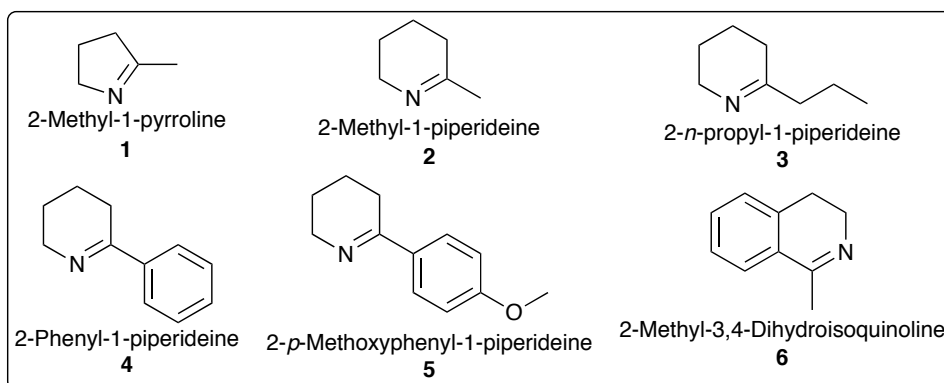


Figure 47 – Substrates used to assess IRED activity.

Table 19 – Typical reaction components used for biotransformations.

Component	Quantity
Enzyme	0.5 mg mL ⁻¹
NADPH	5 mM
Imine 1-6	5 mM
100 mM Sodium Phosphate pH 7.0 Buffer	to 1 mL
Perform biotransformation in 2 mL Eppendorf tube	

Table 20 – Protocol for analysing biotransformations for substrates 1-6. Hex:iPrOH:DEA is hexane:isopropanol:diethylamine.

Substrate	Retention Time / min		Analysis Method	Relative Response Factor ($A_{\text{imine}}/A_{\text{amine}}$)	Column	Protocol	
	Imine	Amine					
		(R)-					(S)-
1	33	38	36	GC	N/A	BGB-175 (BGB Analytik) hold 130°C for 30 min, 130°C - 160°C at 20°C min ⁻¹ , hold 160°C for 30 min	
2	10.7	10.9	10.8	GC	N/A	CP-ChiraSil-Dex CB (Agilent Technologies) 50°C - 200°C at 10°C min ⁻¹	
3	14.9	15.8	15.6	GC	N/A	CP-ChiraSil-Dex CB (Agilent Technologies) 90°C - 200°C at 4°C min ⁻¹	
4	18.1	25.1	24.8	GC	N/A	CP-ChiraSil-Dex CB (Agilent Technologies) 90°C - 200°C at 4°C min ⁻¹	
5	16.1	8.9	12.9	HPLC	19.1	CHIRALPAK® IC™ (Daicel) Hex:iPrOH:DEA = 90:10:0.1, flow rate at 1 mL min ⁻¹	
6	18.2	11.4	10.4	HPLC	9.1	CHIRALPAK® IC™ (Daicel) Hex:iPrOH:DEA = 90:10:0.1, flow rate at 1 mL min ⁻¹	

For reductive amination, initially substrates **7a**, **7c**, **8a** and **8c** were initially screened to assess RedAm activity. After confirming RedAm activity, biotransformations were performed on the substrates listed in **Figure 48**. The components required (**Table 21**) were mixed in a 2 mL eppendorf. Reductive amination biotransformations and quenching procedures were carried out using the same method as imine reduction biotransformations. The retention times and analysis methods are in **Table 22** and **Table 23**.

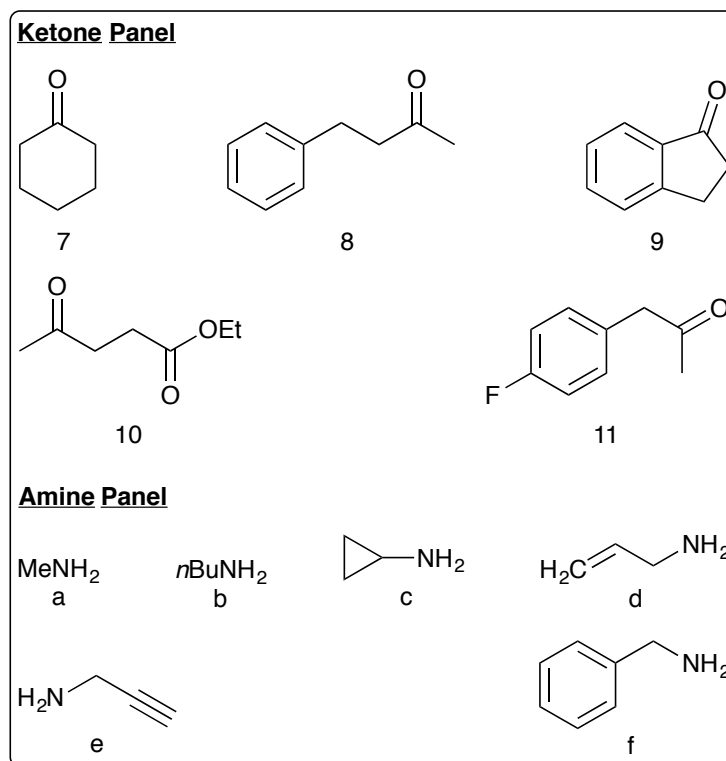


Figure 48 – The substrate panel for reductive amination.

Table 21 – Reaction components for reductive amination.

Component	Quantity
Enzyme	0.5 mg mL ⁻¹
NADPH	5 mM
Ketone 7-11	5 mM
Amine a-f	5 mM, 20 mM or 250 mM
100 mM Tris-HCl pH 9.0 Buffer	to 1 mL

Table 22 – GC methods for analysing reductive amination substrates and products. Products with ketone 7 are non-chiral. Absolute configuration of 8e is based on 8a. Absolute configuration for products with ketone 10 have not yet been assigned.

Substrate		Retention Time / min			Analysis	
		Ketone	Amine Product			
Ketone Acceptor	Amine Donor		Ketone	t _{R1}	t _{R2}	Column
7	a	5.9	6.3	N/A	CP-ChiraSil-DEX CB (Agilent)	90°C - 160°C at 0.5°C min ⁻¹ , hold 160°C for 10 min
7	d	10.1	13.4	N/A	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
7	e	10.1	14.9	N/A	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
7	f	10.1	24.3	N/A	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
8	e	17.7	24.5		CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
10	c	12.5	19.2	19.3	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
10	d	12.5	17.4	17.5	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
10	e	12.5	18.2	18.3	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹
11	e	16.3	20.3 (R)	20.5 (S)	CP-ChiraSil-DEX CB (Agilent)	50°C - 200°C at 10°C min ⁻¹

He flow rate: 1.2 mL min⁻¹; Injector Temperature: 200°C; and Dectector Temperature: 250°C

Table 23 – HPLC methods for analysing reductive amination substrates and products.

Substrate		Retention Time / min			Analysis		
		Ketone	Amine Product				
Ketone Acceptor	Amine Donor		Ketone	t _{R1}	t _{R2}	Solvent ratio	Relative Response Factor (A _{imine} /A _{amine})
8	a	8.1	6.6 (R)	7.2 (S)	99:0.5:0.1	1	CHIRALPAK® OD-H (CHIRAL)
8	b	8.5	5.3 (R)	5.6 (S)	99:0.1:0.1	1	CHIRALPAK® OD-H (CHIRAL)
9	e	23.5	6.6 (S)	7.4 (R)	98:2:0.1	2.3	CHIRALPAK® IC (CHIRAL)
11	a	7.2	8.7 (S)	9.6 (R)	98:2:0.1	1.3	CHIRALPAK® IC (CHIRAL)

Running solvent = n-hexane/iPrOH/diethylamine

3.5. Protein Crystallography

3.5.1. High Throughput Crystal Trial Screening

Once protein was purified to high purity and homogeneity it was subjected to crystallisation trials, in sitting-drop format, with commercially available sparse matrix crystallisation screens. 54 μ L of commercially available screens such as INDEX HT (Hampton Research),¹²²⁻¹²⁵ PACT *premier*[™] (Molecular Dimensions),¹²⁶ and Clear Strategy[™] Screen's 1 and 2 (Molecular Dimensions)¹²⁷ were dispensed into the reservoir of the MRC 96 well crystallisation plates[™] (Molecular Dimensions) using the Hydra[™]-96 Microdispenser (Robbins Scientific) liquid handling unit. Using the mosquito[®] LCP (TTP Labtech), crystallisation drops, typically 300 nL, containing a 1:1 ratio of pure protein:crystallisation solution were dispensed into the crystallisation wells of the MRC 96 well crystallisation plates[™]. Once the crystallisation drops are dispensed, the MRC 96 well crystallisation plates[™] were sealed using ThermalSeal RT[™] (EXCEL Scientific, Inc.) before significant evaporation that could potentially hinder crystallisation occurred. One of the benefits to using the MRC 96 well crystallisation plates[™] is that there are two wells available for crystallisation. Typically one well is used to screen the *apo*-protein and the second well is used to screen the protein-ligand complex.

3.5.2. Co-crystallisation

There are various methods of co-crystallisation,¹²⁸ however the ones only used in this thesis will be discussed below.

3.5.2.1. Incubation with ligand prior to screening

One of the most common methods of attempting to obtain a ligand complex is by incubating the ligand with the pure protein prior to subjecting the protein-ligand solution to crystal trials. This can be performed by incubating the

ligand with a dilute stock of protein typically up to 3 h on ice either with or without gentle shaking or rocking prior to concentrating the protein-ligand solution to the desired concentration. Another method is to add ligand to a concentrated protein stock, prior to setting up the crystallisation drops (3.5.1) in the MRC 96 well crystallisation plates™.

3.5.2.2. Ligand Soaking

Another method for obtaining a ligand complex is to obtain *apo*-protein crystals and then retrospectively attempt to add the ligand into the crystal *via* soaking. Crystals are often soaked in the conditions of the mother liquor with a cryoprotectant (e.g. 10-20% (v/v) ethylene glycol) added to minimise the formation of ice-rings, by removing water content, upon X-ray diffraction analysis. Cryoprotectants can occasionally influence the quality of the diffracting crystal, both positively by improving diffraction quality,¹²⁹ and negatively by impairing diffraction quality and even dissolving or cracking the crystal. Ligands can be added in addition to the cryoprotectant to both attempt to form a protein-ligand complex and minimise the formation of ice rings. If liquid soaks with cryoprotectants are mostly unsuccessful and yield low occupancy complexes; the ligand concentration in the cryoprotectant can be increased. However, this may impair crystal quality and diffraction quality. Lengths of soaking must be tested if protein-ligand interactions are weak, as such, soaking can be anywhere between 5 min, to 24 h and even up to one week prior to fishing and flash freezing the crystal. However this may create non-physiologically relevant crystallographic artefacts with ligands appearing in the crystal structure in places with no relevance.

Another method of soaking is to directly add the ligand compound (solid powder, liquid or oil) to the crystallisation drop that has formed the *apo*-crystal. Again the length of the soak must be trialled.

Soaking is only recommended if the enzyme does not require a large conformational change. As protein in the crystals are often still active, the large conformational change often causes the crystal to dissolve or crack.

3.5.2.3. Adding the ligand as an additive

This method was developed to provide an alternative method for co-crystallisation. It was inspired by the fact that components in crystallisation solution, for example buffers such as bis-tris propane, precipitants such as polyethyleneglycol (PEG) 3350 and salts such as acetate or phosphate are often seen in crystal structures providing crystal contacts. The presence of the components is so significant that clear density is seen in the crystal structure in both $2mF_o - DF_c$ and $mF_o - DF_c$ maps contoured to 1.0σ and $\pm 3.0 \sigma$ respectively.

Once a commercially available screen was dispensed into an MRC 96 well crystallisation plates™, the ligand was added into the reservoir, typically 54 μL , in nanolitre quantities using the mosquito® LCP, then sealed using Classic ThermalSeal® (EXCEL Scientific, Inc) and incubated (typically 1 h at room temperature), prior to setting up crystallisation drops (3.5.1), to allow for the ligand to diffuse and mix appropriately with the mother liquor. The temperature and length of incubation is dependent on ligand stability. Ligands are often only partially soluble in water and stock solutions are made up using water-miscible solvents. In general, the aim was to alter the final volume of the mother liquor by less than 1% to minimise altering the molarity of the crystallisation components. Some proteins are solvent-sensitive, by minimising the final concentration of solvent to less than 1% (v/v), the affect of solvent on the protein is less harsh compared to soaking ligands in cryoprotectants (3.5.2.2).

For example, 54 μL of commercially available screen (i.e. INDEX HT) are dispensed into an MRC 96 well crystallisation plates™ and the highest stock concentration of a given ligand in DMF is 250 mM. The desired final concentration of ligand in the mother liquor is 0.5 mM, which equates to 108 nL of ligand in the 54 μL mother liquor. This also equates to 0.2% (v/v) final concentration of DMF. The screen has now been altered, effectively a new screen has been generated: INDEX HT + 0.5 mM ligand. The mosquito® LCP capable of dispensing as little as 25 nL, therefore the higher the stock concentration, the better. 25 nL would equate to approximately 0.05% solvent contamination. However, certain ligands may require more solvent in the mother liquor to stay in solution and not precipitate. If the protein is not solvent-sensitive it is possible to go beyond 1% (v/v) final solvent concentration.

3.5.3. Crystal Optimisation

Due to the small crystallisation drops in crystal trail screening versus a relatively large mother liquor volume, a volumetric ratio of 1:180 crystallisation drop:mother liquor volume, vapour exchange is rapid. As such, crystals formed in the screening often form rapidly and are anisotropic. Therefore it is necessary to optimise crystals. One of the simplest and most effective methods is to slow down crystal growth. It is common to optimise crystals, by repeating the condition where the crystal hit was observed with a smaller ratio of crystallisation drop:mother liquor and by altering chemical parameters such as PEG, salt, precipitant concentration or pH. By reducing the PEG and precipitant concentration, the crystal will take longer to form, but the crystal packing will be less anisotropic and will provide better diffraction quality.

To further slow crystal growth, small quantities (up to 10% (v/v) or (w/v)) of additives can be added, such as: short chain alcohols (glycerol, ethylene glycol, 2-methylpentan-2,4-diol and butan-1,4-diol), amines (triethylamine, benzamidine and spermidine), water-miscible solvents (DMF, dimethyl sulfoxide and 1,4-dioxane), surfactants and detergents (*n*-octyl- β -D-glycoside and non-detergents sulfobetaines [NDSB's]).

Optimisations are often performed with larger crystallisation drops either in sitting-drop format with 48-well MRC Maxi Optimization plates (Molecular Dimensions) or in hanging-drop format with 24-well Linbro plates (Hampton Research) with siliconised glass cover slips (Hampton Research) sealed with Dow Corning® High Vacuum Grease (Dow Corning). Typical volumes used with the 48-well MRC Maxi Optimization plates are 2 μ L crystallisation drop (1:1 protein:mother liquor ratio) and 150 μ L reservoir solution. Typical volumes used with 24-well Linbro plates are up to 4 μ L crystallisation drop (1:1 protein:mother liquor ratio) and up to 1 mL reservoir solution.

3.6. Data Collection and Model Building

Initial diffraction analysis of crystals was performed in-house on a Rigaku Micromax-007HF fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. Crystals that diffracted to a resolution greater than 3 Å were sent for full dataset collection at the synchrotron. Full datasets were collected on macromolecular beamlines with Pilatus detectors at Diamond Light Source synchrotron, Didcot, Oxford. The Pilatus detector is a digital pixel based detector, therefore its sensitivity and readout times are more superior compared to CCD detectors, which allows for less beam transmission to be used for a shorter time during collection.

Grid scans and line scans were initially performed on crystals to screen and assess the best diffracting area and the space group by indexing the screening data. Typically 180-220° of data were collected, which usually provides near complete data in the outer shell and sufficient redundancy. With precious samples, data collection strategies can be formed with the initial screening data, such as fewer degrees of data collected with crystals of high symmetry.

3.6.1. Radiation damage

All data in this thesis were collected under a cryostream of liquid nitrogen, therefore collection occurred at temperatures of liquid nitrogen (-196°C). This aided in preserving the crystal in a cryoprotected state and also aids in retarding damage caused by being bombarded by electrons during data collection; this is also known as radiation damage.¹³⁰ Radiation damage can severely influence¹³¹⁻¹³⁴ the quality of data by decreasing the resolution limit and signal-to-noise ratio, $I/\sigma(I)$, increase atom movement (B factors) and overall reliability of the data (R_{merge}), and can increase mosaicity and cause unit cell expansion. The effects of radiation damage can also be seen whilst modelling the structure; as well as residues having an increased B factor value, decarboxylation and demethylation bond cleavage is common. As such, these structural changes can lead to misinterpretation of the structure. Metalloproteins and ligands are susceptible to

redox changes when exposed to X-rays and the major affect this is that the metal ion being in the undesired redox state is that coordination could be different or the ligand has been excited by the X-rays, which again leads to misinterpretation of the structure.

In combination to using advanced equipment at the synchrotron, such as Pilatus detectors and faster goniometers, it has become more common to use a lower percentage of beam transmission. This is because the Pilatus detectors are sensitive enough to not require 100% beam transmission to detect diffraction. The benefits of this are that the signal-to-noise on the diffraction analysis is significantly less and radiation damage is significantly reduced. Using less beam transmission allows for more accurate data collection on sensitive samples.

3.6.2. Helical data collection

Once a crystal has been screened for the best diffracting area, even when it is collected using a lowered beam transmission, it is still possible that the crystal will suffer from radiation damage. This is because the crystal will accumulate a large dose of radiation at a single point during the full dataset collection. For radiation-sensitive crystals this is an issue, however it is possible to collect crystal data in a 'helical' fashion;¹³⁵ this is essentially an extension of line scan. Two points of the crystal are defined as the start and end. Data is collected whilst the crystal is translated and rotated between the two points. This potentially allows for longer exposure per image for a higher signal-to-noise ratio, since a fresh part of the crystal is exposed for data collection every image, thereby minimising the total radiation dose per image and therefore minimising radiation damage.

3.6.3. Data processing: Indexing, Integration and Scaling

Datasets are automatically processed as a part of the data collection pipeline by the data processing software Xia2¹³⁶ at Diamond Light Source. In the majority of cases, the automatically processed data is sufficient to proceed with structure solution and model building.

In cases where the automatic data processing performed by Xia2 was not ideal, the data was initially visually screened for anomalies using Adxv¹³⁷; such as the presence of two lattices, in some cases the minor lattice was able to be processed out. The data was then indexed using mosfilm¹³⁸⁻¹³⁹ via the GUI imosflm¹⁴⁰, with space groups chosen where the penalty value for a given space group is less than 20. The data was then integrated using XDS¹⁴¹ via the XDS GUI with the parameters obtained from Adxv (resolution limits) and mosflm (space group and cell dimensions); typically the outershell resolution limit was set 0.5 Å greater than what is visually observed in Adxv as faint pixels not visible by eye can be detected by XDS. Once integrated, the data was scaled using Aimless¹⁴², which normalises reflection intensities. Aimless also runs the program pointless within its pipeline. Pointless checks for twinning, lattice centring (i.e. recentring for a given crystal, centred to C faces $C12_11$ to body centred $I12_11$) and checks for all possible Laue groups within the crystal class based on cell dimension restraints (i.e. for $P222$ could be $P2_12_12_1$).

3.6.4. Structure Solution, Model Building and Refinement

Aimless generated an MTZ file that contains all necessary data such as reflection data, symmetry operations and cell dimensions, which is then useable for structure solution and refinement. There are many different methods of obtaining a structure solution, however only molecular replacement was used in this thesis. Direct heavy atom solution methods are not discussed in this thesis.

Molecular replacement relies on the existence of a model that is structurally similar to the unsolved structure. The two most common software for performing molecular replacement are Molrep¹⁴³ and Phaser,¹⁴⁴ both of which have been used for structure solutions in this thesis. The approaches behind each software are very different; Molrep is Patterson derived, whereas Phaser uses maximum likelihood estimation.

Once a successful structure solution has been found, iterative cycles of model building in COOT¹⁴⁵ and refinement in Refmac5,¹⁴⁶ without TLS and jelly-body refinement, are performed until negligible change is observed with the models' reliability factors (R_{factor} and R_{free}). Water molecules are then modelled

using COOT build waters function into the structure, which is then refined before modelling ligands. It is common that xenobiotic ligands do not have known bond constraints (cif dictionary file), therefore the ligands are energy minimised using ProDrg¹⁴⁷ with a dictionary file with appropriate bond constraints generated. Structures are then validated with PROCHECK¹⁴⁸ prior to deposition.

4. Results

4.1. (S)-Selective homologues: *Bc*IRED and *Nh*IRED

In 2013, the structure of *Sk*IRED, a homologue of *Sr*IRED, was published and was also the first IRED structure.¹⁰⁸ However, *Sk*IRED was characterised as an (*R*)-selective IRED, with activity that was significantly worse compared to *Sr*IRED.

Elizabeth Wells (personal communication, in our group) attempted to crystallise *Ss*IRED, however the crystallisation had only resulted in partial success, yielding a low-resolution structure of an (*S*)-selective IRED. At the time of starting the project there was not yet a published crystal structure for an (*S*)-selective IRED. Therefore, two potential homologues of *Ss*IRED, from *Bacillus cereus* BAG3X2 (*Bc*IRED) and *Nocardiosis halophila* (*Nh*IRED) were chosen, based on sequence similarity (**Figure 49**), for further investigation.

```
NhIRED -----MTNTKAKKSPVTLIGLPGMQAMAGALLEAGYELTVVNRKAKAEALAERGA AVADSPAEALR
SsIRED -----MSKQSVTVIGLPGMQAMVNTFLDNGHEVTVWNRKAEALVARGAVLAPTVEDALS
BcIRED MLKSNSQNEKNGSETTNAVGNRKSVTVIGLPGMQAMADVFLYGYSVTVWNRSSKADQLVAKGAI RVSTVNEALA
      .:. **:*****. .:*. *:***** **:*. **: . : **

NhIRED AGGPVLLSLTEHAVMYRVLEGAESDLKGRITILNLGSDTPAASRAAAWAEGHGARYVTGGVMS PAPGIGSSSVFSFY
SsIRED ANELIVLSLTDYDAVYAILEPVTGSLSGKVIANLSSDTPDKAREAAKWAAKHGAKHLTGGVQVPPPLIGKPESSTYY
BcIRED ANELVILSLTDYVNMYSILEPVSLENLFGKVLVNLSSDTPKARKAAKWLEDRGARHITGGVQVPPSGIGKSES YTTY
      * :*****: .:* **: . * *: : ** ***** * ** * **::***** * ** . : **

NhIRED SGDRAAFEENKALLEVL TATDFRGEDPGLAQVYQILLDLFWTMTGYLHALAVARAEGVPVGTITPYLIEGN-DMA
SsIRED SGPKDVFDAHEDTLKVL TNADYRGEDAGLAAMYQAQMTIFWTTMLSYYQTLALGQANGVSAKELLPYATMMTSMMP
BcIRED SGDRVVFEAHRETLEVL TSSDYRGEDPGLAMLYQIQMDIFWTAMLSYLHALAIANANGITAEQFLPYASAMMSSLP
      ** : .*: . . ***: ***:** * ** : ** : ***: * . * ***: . .*: : . : **

NhIRED MFFEGTSAAVAEGRFPGDEDRISMDAASMEHVQVTSRDAGVDTALPEAVLSLFRRLDAGFAESSFARLVTLMDAEG
SsIRED HFLELYAQHVSDADYPGDVDRILAMGAASVDHVLH THQDAGVSTVLPAAVAEIFKAGMEKGF AENSFSSLIEVLK KPA
BcIRED KFVEFYTPRLDEGEHPGDVDRILAMGLASVEHVVHTTQEAGIDIALPATVLEVFRRGMKTGHASDSFTSLIEIFKNSD
      *.* : : . . *** ***: * ***:** * ***: . ** * * .*: * . . * . ***: * : : .

NhIRED ---
SsIRED V-
BcIRED IRS
```

Figure 49 – Protein sequence alignment of *Ss*IRED, *Bc*IRED and *Nh*IRED, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Protic tyrosine for tyrosine-IREDs highlighted in red and yellow. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹ *Ss*IRED shares 83.7% and 67.7 % sequence similarity with *Bc*IRED and *Nh*IRED respectively.

The synthetic genes for *Bc*IRED (**Figure 50**) and *Nh*IRED (**Figure 51**) were codon optimised for *E. coli* and ordered from GeneArt Gene Synthesis (ThermoFisher Scientific).

>BcIRED
 ATGCTGAAAAGCAATAGCCAGAACGAAAAAATGGTAGCGAAACCACCAATGCAGTGGGTAATCGTAAAAGCGTTACCGTTATT
 GGTCTGGGTCCGATGGGTGAGGCATGGCAGATGTTTTCTGGAAATATAGGTTATAGCGTGACCGTTGGAAATCGTACCAGCAGC
 AAAGCAGATCAGCTGGTTGCAAAAAGGTGCAATTCGTGTTAGCACCGTTAATGAAGCACTGGCAGCAAAATGAACTGGTTATTCTG
 AGCCTGACCGATTACAATGTGATGTATAGCATTCTGGAACCGTTAGCGAAAACCTGTTTGGTAAAGTTCTGGTGAATCTGAGC
 AGCGATACACCGGAAAAAGCACGTAAAGCAGCAAAATGGCTGGAAGATCGTGGTGCACGTCATATTACCGGTGGTGTTCAGGTT
 CCGCTAGCGGTATTGGTAAAAGCGAAAGCTATACCTATATAGCGGTGATCGTGTGTTTTGAAACCATCGTAAAACCGT
 GAAGTTCAGCAGCAGCAGATTATCGTGGTGAAGATCCGGGTCTGGCAATGCTGATATTATCAGATTGATGGATATCTTTTGG
 ACCGCAATGCTGAGCTATCTGCATGCACCTGGCAATGCAAAATGCAAAATGGTATTACCGCAGAACAGTTCTGCCGTATGCAAGC
 GCAATGATGAGCAGCCTGCCGAAATTTGTTGAATTTCTATACACCGCTCTGGATGAAGGCGAACATCCGGGTGATGTTGATCGT
 CTGGCCATGGGTCTGGCCAGCGTTGAACATGTTGTTTATACACCCCAAGAGCCGGTATGATATGCACTGCCTGCAACCGTT
 CTGGAAGTTTTCTGTCGGGTATGAAAACCGTTCATGCAAGCGATAGCTTTACCAGCCTGATTGAGATTTTCAAAAATAGCGAT
 ATCCGACGCTAA

Figure 50 – The DNA sequence in fasta format for BcIRED, codon optimised for E. coli.

>NhIRED
 ATGACCAACACCAAGCAAAAAAAGTCCGGTTACCTGATTGGTCTGGGTCCGATGGGTGAGCAATGGCAGGCGCACTGCTG
 GAAGCAGTTTATGAACTGACCGTTTGGAAATCGTACCAAAAGCCAAAGCCGAAAGCACTGGCAGAACGTGGTGCAGCAGTTGCAGAT
 AGCCTGCGGAAGCACTGCGTGCCGGTGGTCCGGTCTGCTGAGCCTGACCGAACATGCAGTTATGATCGTGTCTGGAAGT
 GCAGAAAAGCGATCTGAAAAGGTCTGACCAATCTGAATCTGGGTAGCGATACACCGCAGCAAGCCGTGCAGCAGCAGCATGGGCT
 GAAGGTGATGGTGCACGTTATGTTACCGTGGTGTATGAGTCCGGCACCGGTATTGGTAGCAGCAGCGTTTTTGTAGCTTTTAT
 AGCGGTGATCGTGCAGCCTTTGAAAGAAAATAAAGCCCTGCTGGAAGTTCTGACCGCAACCGATTCTCGTGGTGAAGATCCGGGT
 CTGGCACAGGTTTTATTATCAGATTCTGCTGGACCTGTTTTGGACCACCATGACAGGTTATCTGCATGCACTGGCCGTTGCACGT
 GCCGAAGGTGTTCCGGTTGGCACCATTACCCCGTATCTGATTGAAGGTAATGATATGGCCATGTTTTTTGAAAGGCACCAGCGCA
 GCCGTTGCGGAAGTCTGTTTTCCGGGTGATGAAGATCGTATTAGCATGGATGCAGCAAGCATGGAACATGTTGTTTCAGACCAGC
 CGTGATGCCGGTGTGATACCGCTCTGCCGGAAGCAGTTCTGAGCCTGTTTCGTGCTGGTCTGGATGCAGGTTTTTCCGCAAGC
 AGCTTTGCACGCTGGTGACCTGATGGATGCCGAAGGCTAA

Figure 51 – The DNA sequence in fasta format for NhIRED, codon optimised for E. coli.

The synthetic genes were amplified by PCR (3.1.2) using the primers listed in Table 43 (6.1) and the protocol in Table 24.

Table 24 – Protocol for amplifying BcIRED and NhIRED.

Reaction Mixture		PCR Protocol		
Component	Volume (µL)	Temperature (°C)	Time	Cycles
Template (50 ng/µL)	1	94	2 min	1
For Primer (20 µM)	1	94	30 s	
Rev Primer (20 µM)	1	50	30 s	35
dNTPs (2 mM each)	5	72	30 s	
MgSO ₄ (25 mM)	3	72	3 min	1
KOD Hot Start Polymerase Buffer (10 x)	5	10	∞	-
KOD Hot Start DNA Polymerase (1 U/µL)	1			
H ₂ O	33			

After amplification, the resulting genes were separated by gel electrophoresis on a 1% agarose gel and subsequently purified by gel extraction (Figure 52). The purified material was then prepared for annealing (Table 8) (3.1.5). Once annealed, the nicked plasmids were used to transform *E. coli*, NovaBlue Singles™ (Table 13 - 3.1.8), to repair the plasmids and subsequently replicate the plasmid for plasmid extraction by miniprep.

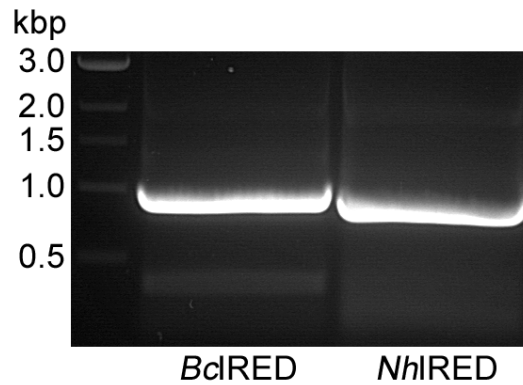


Figure 52 – PCR products of *BclIRED* and *NhlIRED* analysed by fragment analysis electrophoresis on a 1% agarose gel.

To confirm the presence of an insert of the correct weight annealed into the desired vector, a double digest was performed (**Table 14 - 3.1.9**). Fragment analysis confirmed inserts of the correct DNA band had been inserted (**Figure 53**). The gene encoding *BclIRED* contained an *NcoI* cut site 200 bp upstream of the 3'-end, as such a fragment smaller than 1 kbp was expected.

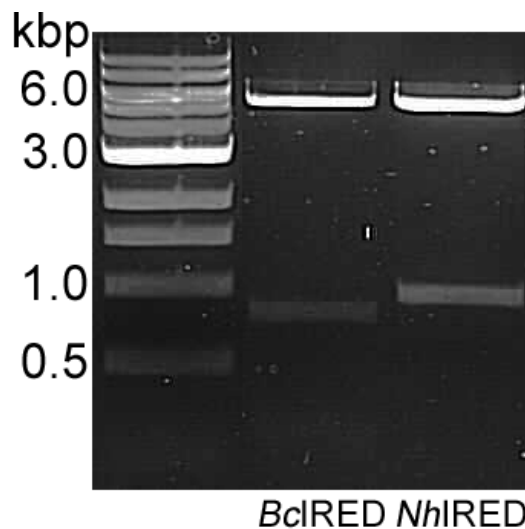


Figure 53 – Fragment analysis by gel electrophoresis for analysing the double digest of pET-YSBLIC3C plasmids with genes for *BclIRED* and *NhlIRED* inserted.

Plasmids that yielded correct fragments were subsequently sent for sequencing at GATC. Plasmids were sequenced with generic T7 primers for the forward read and generic pET-rp for the reverse read. Plasmids that contained the gene of interest correctly annealed were then transformed (**Table 13 - 3.1.8**) by an expression strain of *E. coli*, BL21(DE3). 5 mL expression tests was performed to identify over expression and optimal expression temperatures.

The expression test for *Bcl*RED (Figure 54) identified over expression at 16°C, 30°C and 37°C when cells were induced with 1 mM IPTG. Significant presence of insoluble protein was observed at 37°C and soluble protein expression was significantly greater at 30°C compared to other temperatures. Therefore 30°C was used for scaling up protein expression for *Bcl*RED.

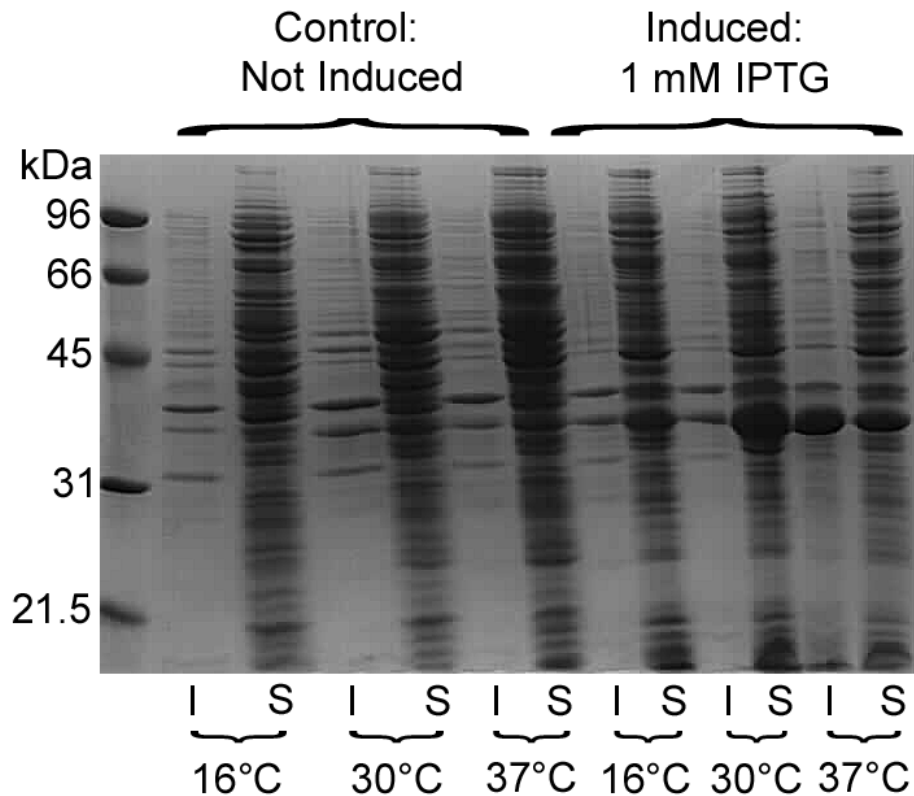


Figure 54 – The expression test for *Bcl*RED, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of *Bcl*RED is 36.5 kDa, including the 6*His*-tag and HRV3C precision site.

The expression test for *Nh*RED (Figure 55) identified leaky expression in the control samples and mostly insoluble over-expression at 16°C and 30°C when cells were induced with 1 mM IPTG. Small amounts of soluble *Nh*RED were present at 16°C in the induced sample. Therefore 16°C was used for scaling up protein expression for *Nh*RED.

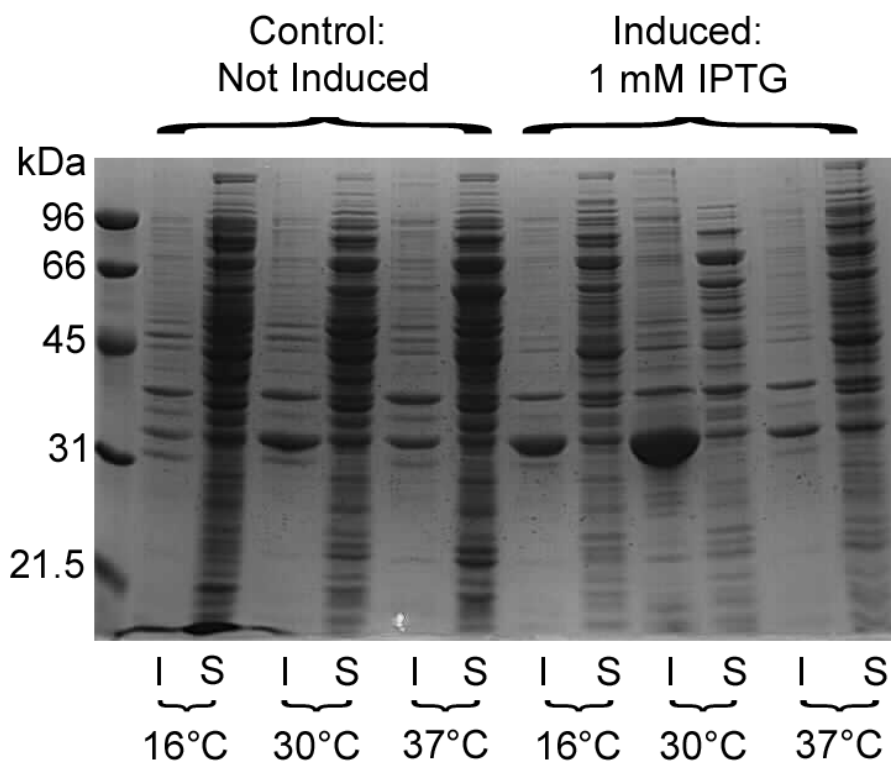


Figure 55 – The expression test for *NhIREd*, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of *NhIREd* is 33 kDa, including the 6His-tag and HRV3C precision site.

With the identified optimal expression temperatures, protein expression was scaled up. After harvesting cells and cell lysis, protein was purified (3.3) by Ni-affinity and SEC (Figure 56 and Figure 57 for *BclIREd* and *NhIREd* respectively), with samples analysed by SDS-PAGE (Figure 58).

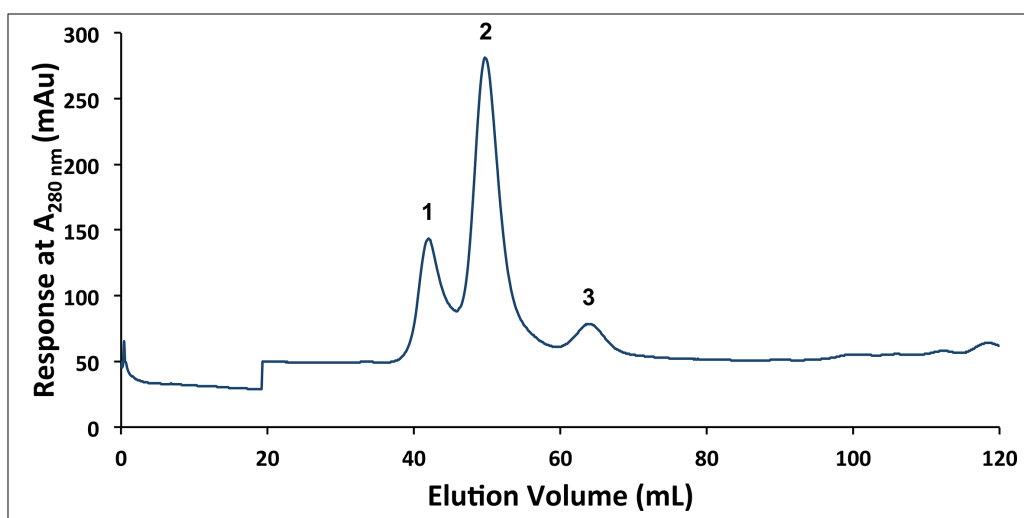


Figure 56 – Chromatogram of the purification of *BclIREd* by SEC. Peak 2 elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa.

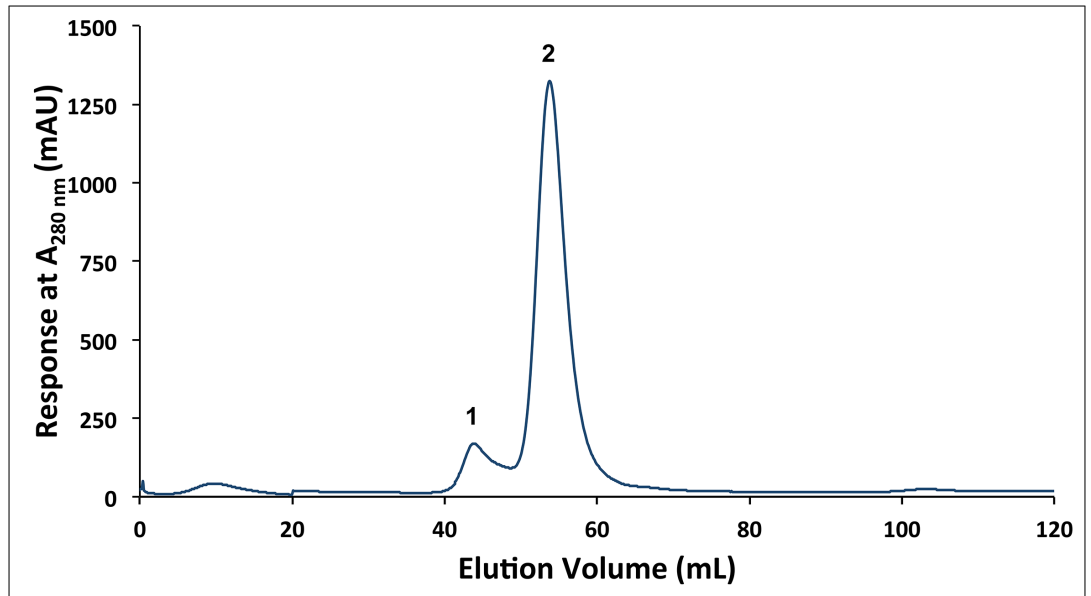


Figure 57 – Chromatogram of the purification of *NhlRED* by SEC. Peak 2 elutes between 50-60 mL, correlates to a species with an approximate molecular weight of 60 kDa.

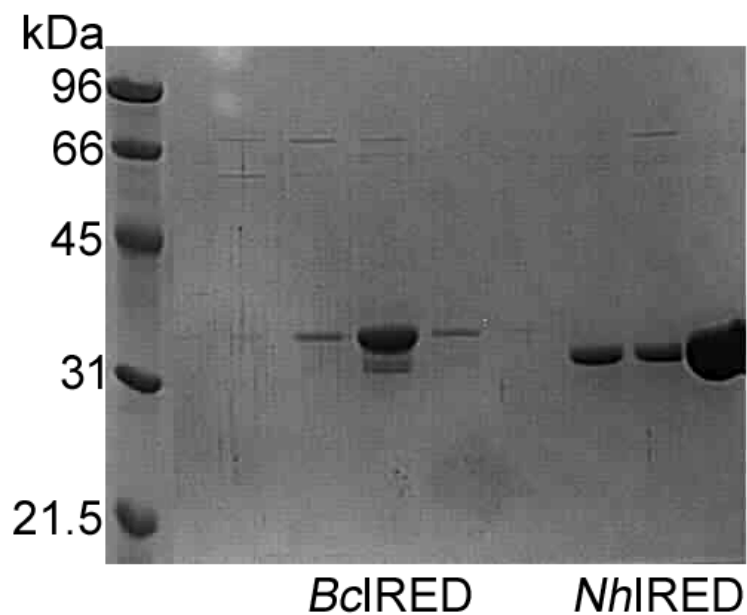


Figure 58 – SDS-PAGE analysis of SEC purified *BclRED* and *NhlRED*. Bands correspond to main peak elutions, peak 2, for both *BclRED* and *NhlRED* from Figure 56 and Figure 57 respectively. Both proteins were purified by Ni-affinity prior to SEC.

Both chromatograms for *BclRED* (Figure 56) and *NhlRED* (Figure 57) illustrate species eluting between 50-60 mL, which would correlate to species of approximately 60 kDa. The SDS-PAGE analysis (Figure 58) shows both *BclRED* and *NhlRED* with monomeric weights of approximately 32 kDa.

The kinetic assay (3.4.1) was performed using purified *Bc*IRED (Figure 59) and *Nh*IRED (Figure 60) with the substrate **1** (2MPN). The kinetic parameters were calculated for both IREDs (Table 25).

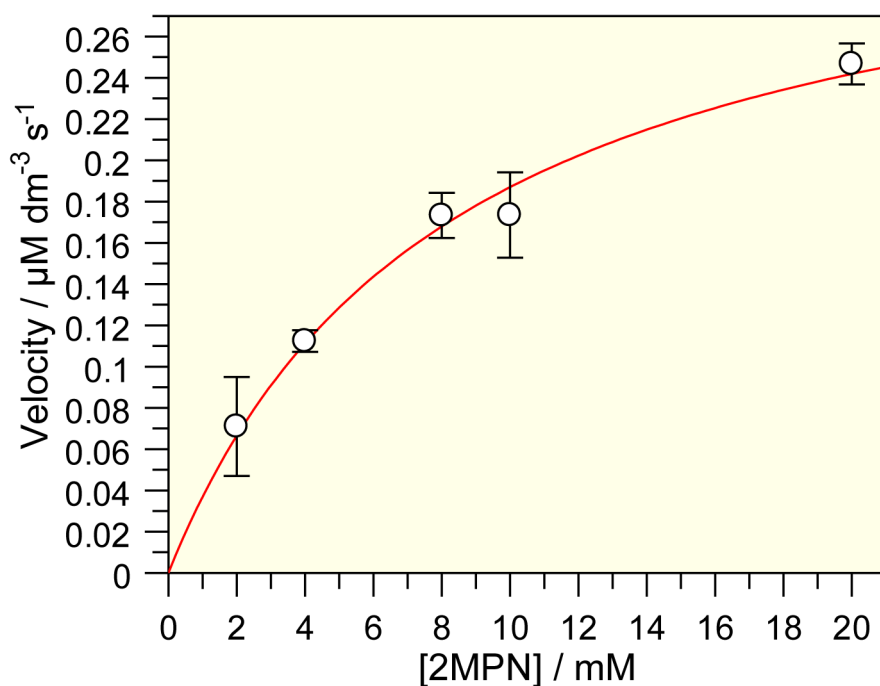


Figure 59 – The Michaelis-Menten fitted data for the kinetic assay with *Bc*IRED and the imine substrate **1** (2MPN).

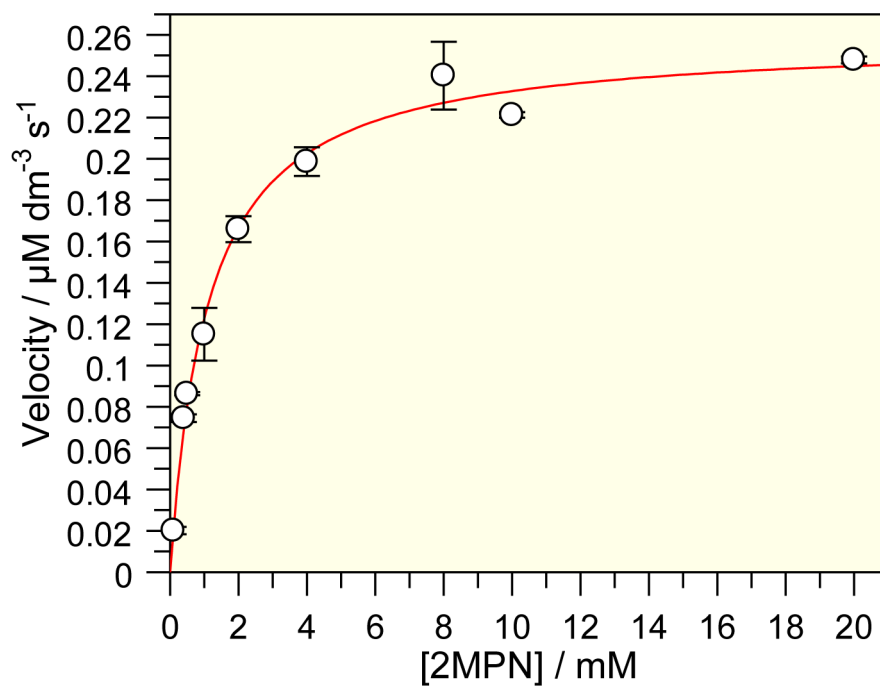





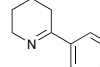
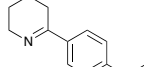

Figure 60 - The Michaelis-Menten fitted data for the kinetic assay with *Nh*IRED and the imine substrate **1** (2MPN).

Table 25 – The kinetic parameters for *Bc*IRED and *Nh*IRED for the imine substrate 1 (2MPN).

Enzyme	Kinetic Parameters			
	V_{\max} ($\mu\text{M dm}^{-3} \text{s}^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\text{s}^{-1} \text{M}^{-1}$)
<i>Bc</i> IRED	0.34	8.32	0.032	3.84
<i>Nh</i> IRED	0.26	1.10	0.013	11.82

Biotransformations for *Bc*IRED and *Nh*IRED were performed (Table 26), as described in Table 19 - 3.4.2, on substrates 1-6 with conversions, enantiomeric excess and absolute configuration determined (Table 20 - 3.4.2).

Table 26 – Biotransformations performed with *Bc*IRED and *Nh*IRED with imine substrates 1-6.

Enzyme	Substrate					
	1	2	3	4	5	6
						
<i>Bc</i> IRED	98%, >99% (S)-	90%, >99% (S)-	57%, 35% (S)-	57%, 35% (R)-	8%, 18% (S)-	94%, 81% (S)-
<i>Nh</i> IRED	80%, >99% (S)-	86%, >99% (S)-	90%, 82% (R)-	100%, 85% (S)-	53%, 98% (S)-	97%, 68% (S)-
percentage conversion, enantiomeric excess and absolute configuration						

The biotransformations indicate that both *Bc*IRED and *Nh*IRED are indeed (S)-selective IREDs towards smaller substrates such as 2MPN, **1**, and **2**. However, *Nh*IRED does not exhibit the strict facial preference that *Bc*IRED displays with substrates **3** and **4**. Interestingly with substrate **5**, both enzymes attack the same *re*-face, albeit with less specificity in the case of *Bc*IRED. The results seen with substrate **5** are most likely dictated by mesomeric electronic effects of the *para*-methoxy group, which would suggest attack of the enamine form rather than the imine form, hence the significantly lower conversions are achieved.

Purified *Bcl*RED was subjected to crystal trials with commercially available screens in 96-well sitting drop format (3.5.1) at a protein concentration of 50 mg mL⁻¹ in buffer B (3.3). Large crystals (Figure 61) appeared overnight at 18°C in conditions containing 0.2 M MgCl₂, 0.1 M HEPES pH 7.5 and 25% (w/v) PEG 3350 (INDEX G12).

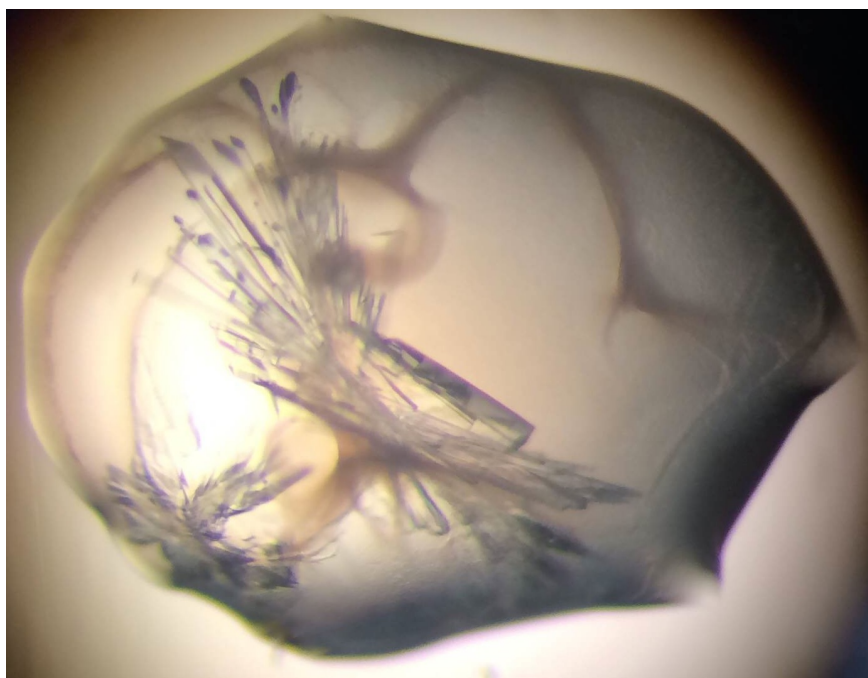


Figure 61 – Crystals of *Bcl*RED, at a protein concentration of 50 mg mL⁻¹, obtained in an 300 nL crystallisation drop in INDEX G12 from crystal trial screening.ik

These crystals were optimised in 24-well hanging drop format (3.5.3), in an attempt to reduce a clustered morphology. The best crystals (Figure 62) were obtained in conditions containing 0.25 M MgCl₂, 0.1 M HEPES pH 7.5 and 30% (w/v) PEG 3350 at a protein concentration of 225 mg mL⁻¹ after 3 days at 18°C.

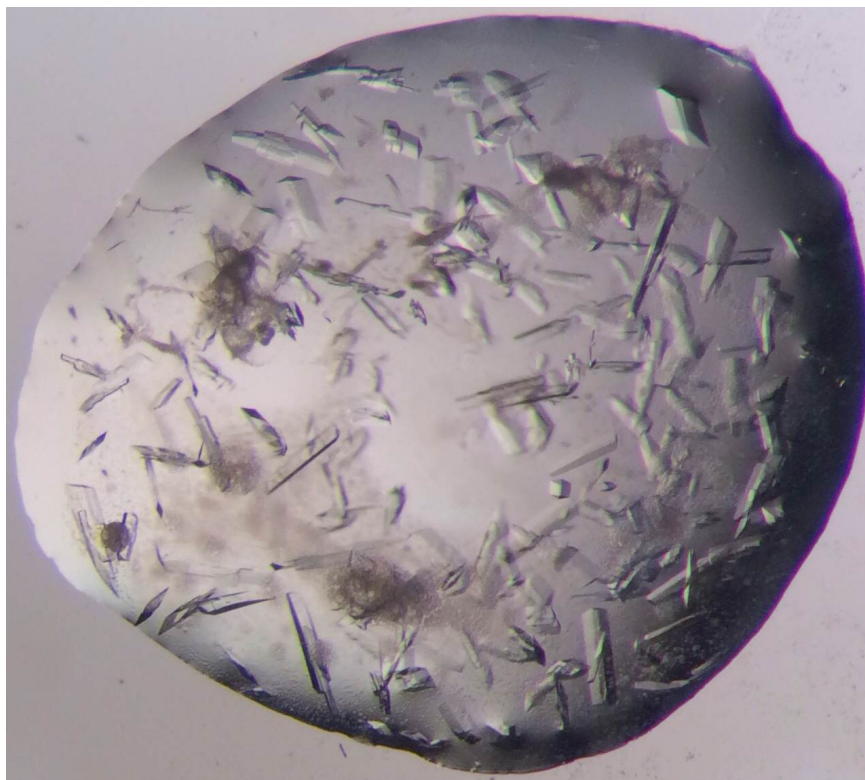


Figure 62 – Optimised, using hanging drop format, crystals of *BclRED* in conditions containing 0.25 M MgCl₂, 0.1 M HEPES pH 7.5 and 30% (w/v) PEG 3350 obtained in a 2 μ L crystallisation drop with a 1 mL mother liquor volume.

Co-crystallisation by incubating *BclRED* with 10 mM NADPH or by adding 20 mM NADPH to the optimised conditions did not yield any crystals. The NADPH complex was obtained by soaking (3.5.2.2) the optimised crystals (at 50 mg mL⁻¹ in buffer B (3.3)) in a cryoprotectant containing the mother liquor (0.25 M MgCl₂, 0.1 M HEPES pH 7.5 and 30% (w/v) PEG 3350), 20 mM NADPH and 20% (v/v) ethylene glycol for 1 h prior to flash cooling with liquid nitrogen.

Once the full dataset for the *BclRED*-NADPH complex had been collected to 1.8 Å resolution, the Xia2 auto processed data was used for structure solution and refinement. The crystal structure of *BclRED* was solved in the space group $P2_12_12_1$ with Molrep using *SkIRED*, PDB entry: 3ZHB, as the molecular replacement model. Iterative cycles of model building and refinement were performed. Water molecules were added last and the model was refined and checked. The residual density in the cleft formed between the dimer interface in the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was assigned as NADPH. The crystal structure is composed of two monomers in the asymmetric unit arranged as a single homodimer (Figure 63).

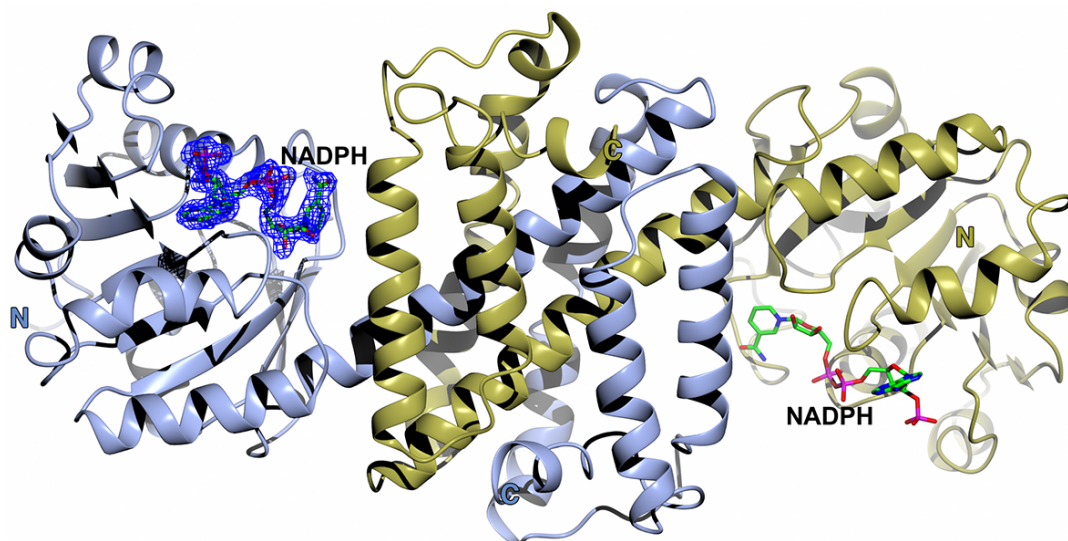


Figure 63 – The crystal structure of *BclRED*-NADPH complex with subunits A and B coloured in blue and gold respectively. The electron density for NADPH in subunit A is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining NADPH.

The final structure was refined to R_{factor} and R_{free} values of 17.4% and 19.7% respectively. The model was complete from residues 22 to 306 in both subunits. PISA analysis indicates a dimer interface of 4048 \AA^2 , that spontaneously forms with a Gibbs free energy (ΔG) of $-76.0 \text{ kcal mol}^{-1}$ and is composed of 25 hydrogen bonds and 12 salt bridges. DALI analysis indicates the closest structural homologue to *BclRED* is *SkIRED* (PDB entry 3ZHB), sharing 35% sequence homology and monomers superposing with an RMSD of 1.5 \AA over 283 residues out of 286. Similar to *SkIRED*, the N-terminal Rossmann domain is formed by residues 23-181, with a long interdomain helix (residues 181-211, α_8 in **Figure 64**) connecting to the C-terminal dimerising domain (residues 212-306).

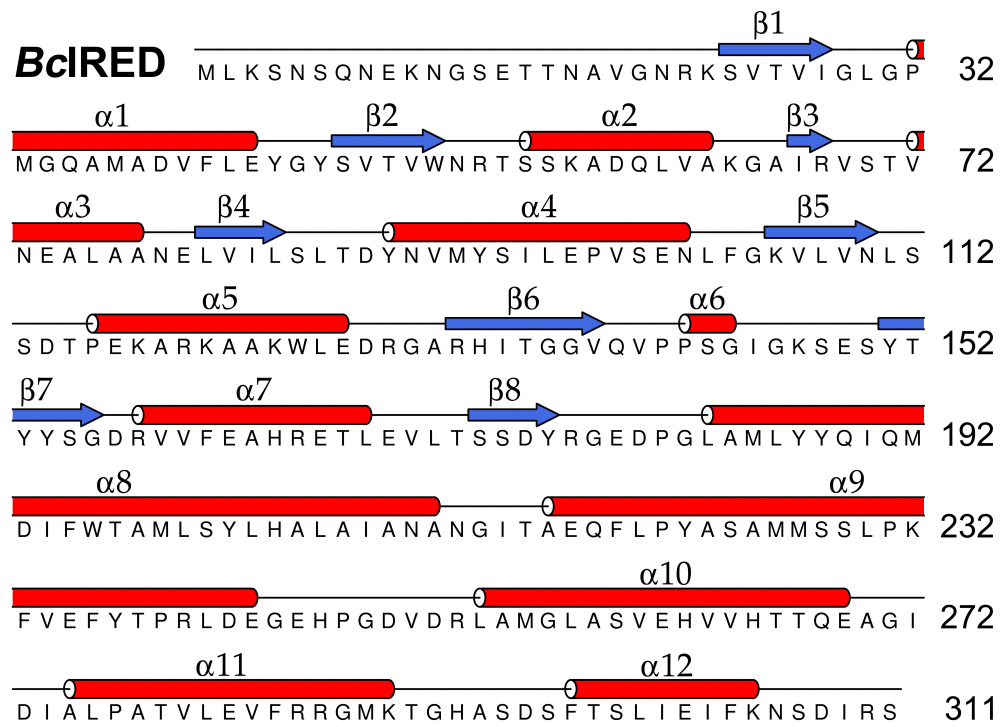


Figure 64 – The primary structure of *BclIRED* with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵²

Purified *Nh*IRED was also subjected to crystal trials with commercially available screens in 96-well sitting drop format (3.5.1) at a protein concentration of 60 mg mL⁻¹ in buffer B (3.3). Crystals (Figure 65) appeared after a month at 18°C in conditions containing 0.25 M MgCl₂, 0.1 M BisTris pH 5.5 and 18% (w/v) PEG 3350 (INDEX G10).



Figure 65 – Crystals of *Nh*IRED, at a protein concentration of 60 mg mL⁻¹, obtained from a 300 nL crystallisation drop in INDEX G10 from crystal trial screening.

*Nh*IRED was optimised using the commercially available Additive Screen HT (Hampton Research) in 96-well format. 10% of the water content of the mother liquor (0.25 M MgCl₂, 0.1 M BisTris pH 5.5 and 18% (w/v) PEG 3350) was replaced with solutions in the additive screen. Crystals appeared after a month at 18°C in conditions containing 0.25 M MgCl₂, 0.1 M BisTris pH 5.5, 18% (w/v) PEG 3350 and 1% (w/v) *n*-octyl- β -D-glucoside (OBDG) at a protein concentration of 60 mg mL⁻¹ in buffer B (3.3). These crystals were scaled up and again optimised in 24-well hanging drop format using a 1 mL mother liquor volume with a 2 μ L crystallisation

drop. The best crystals (**Figure 66**) appeared after 2 months at 18°C in the same conditions and protein concentration as mentioned above.



Figure 66 – Optimised, using hanging drop format, crystals of *NhIRED* in conditions containing 0.25 M MgCl₂, 0.1 M BisTris pH 5.5, 18% (w/v) PEG 3350 and 1% (w/v) OBDG obtained in a 2 µL crystallisation drop.

Co-crystallisation by incubating *NhIRED* with 10 mM NADPH or by adding 20 mM NADPH to the optimised conditions did not yield any crystals. Soaking the crystals in a cryoprotectant containing 20 mM NADPH also did not yield an NADPH complex.

Once the full dataset for the *NhIRED*-OBDG complex had been collected to 2.4 Å resolution, the Xia2 auto processed data was used for structure solution and refinement. The crystal structure of *NhIRED* was solved in the space group $P2_12_12$ with Molrep using *BclRED* as the molecular replacement model. Iterative cycles of model building and refinement were performed. Water molecules were added last and the model was refined and checked. The residual density in the cleft formed between the dimer interface in the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was assigned as OBDG. The crystal structure is composed of eight monomers in the asymmetric unit arranged as four homodimers. Each dimer contained one molecule of OBDG (**Figure 67**).

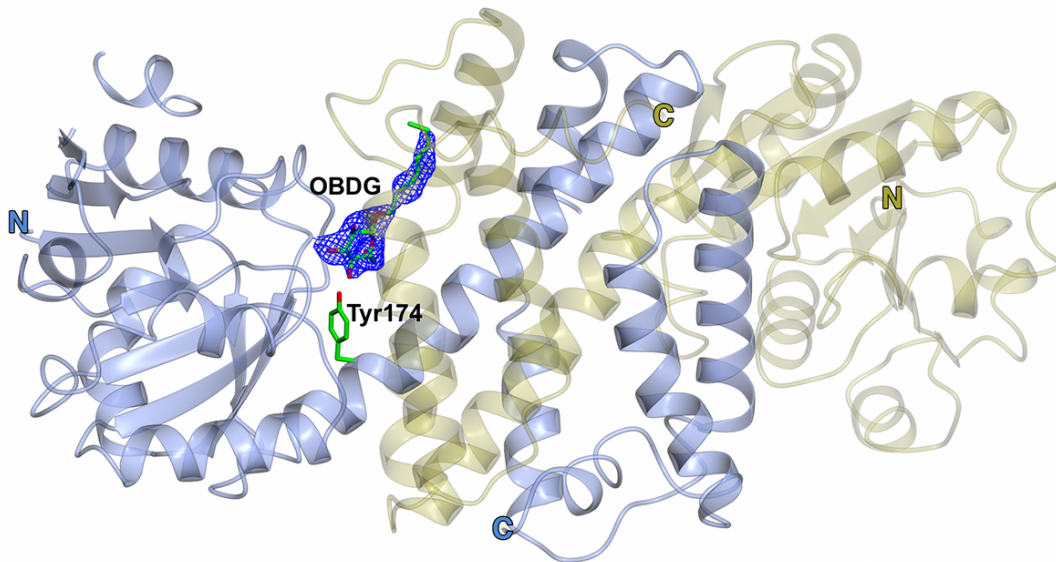


Figure 67 – The dimeric model of *NhlRED*-OBDG complex. The electron density for OBDG is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining OBDG.

The final structure was refined to R_{factor} and R_{free} values of 23.0% and 26.6% respectively. The model was mostly complete from residues 9 to 291 in both subunits. The N-terminal Rossman domain is formed by residues 9-169, with a long interdomain helix (residues 170-196, α_{10} in **Figure 68**) connecting to the C-terminal dimerising domain (residues 212-291).

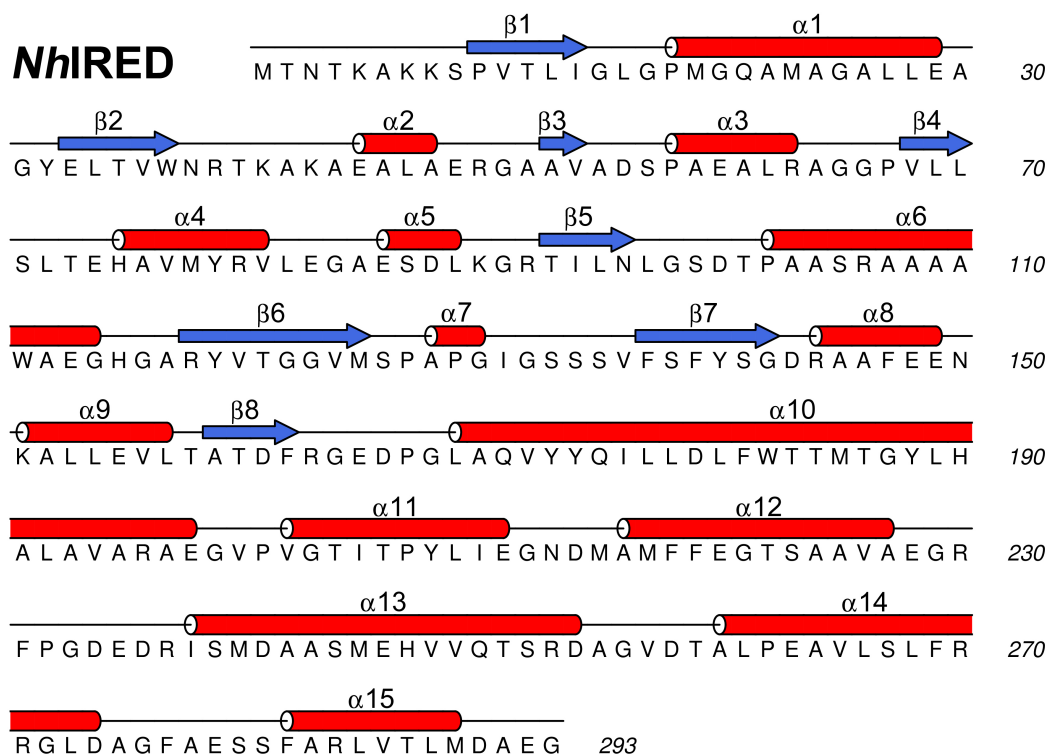


Figure 68 – The primary structure of *NhIRED* with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵²

DALI analysis of *NhIRED*-OBDG complex also indicates the closest structural homologue is *SkIRED* (PDB entry: 3ZHB), sharing 36% sequence homology with monomers superposing with an RMSD of 2.0 Å over 284 residues out of 288. However, the closest structural homologue is in fact *BcIRED*, sharing 50% sequence homology, with monomers superposing with an RMSD of 1.4 Å over 284 residues out of 289. PISA analysis indicates an average dimer interface 4082 Å² consisting of an average of 14 hydrogen bonds and 16 salt bridges with dimers spontaneously forming with an average ΔG of -78.2 kcal mol⁻¹.

As OBDG was present in the active site, an NADP⁺ oxidation assay was performed with *NhIRED* and varying concentrations of OBDG whilst monitoring the increase of absorbance at 340 nm (**Figure 69**).

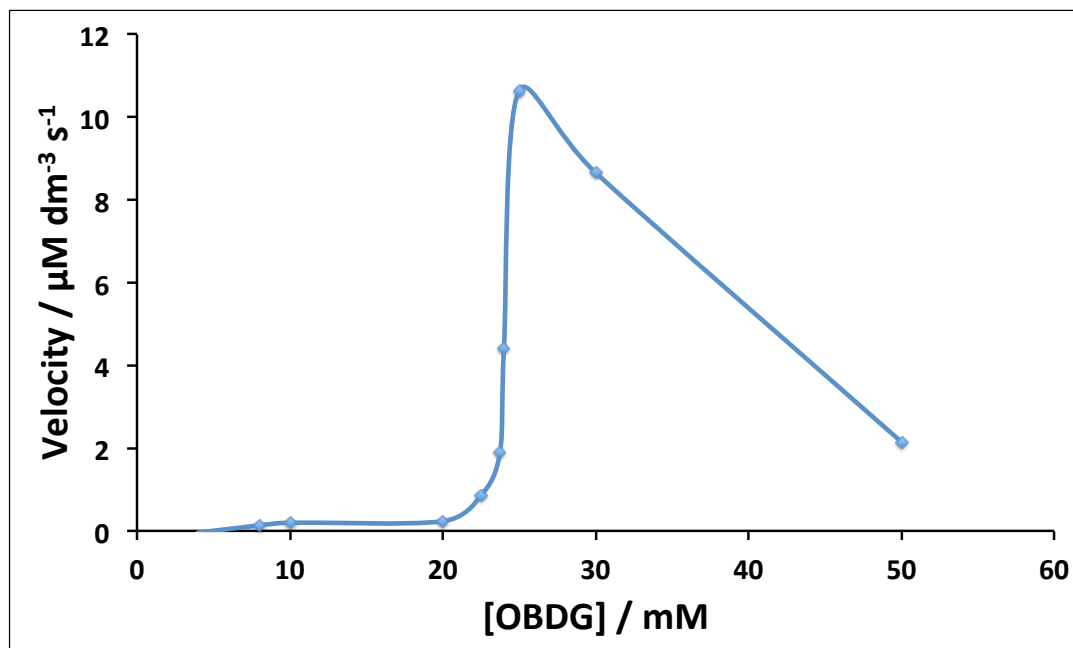


Figure 69 – The curve obtained for the oxidative kinetic assay for *Nh*IRED with increasing concentrations of OBDG, where the reaction rate has been plotted against increasing concentrations of OBDG.

Despite having performed the oxidative kinetic assay for *Nh*IRED and OBDG, it is unclear whether or not OBDG is a substrate (**Figure 69**) and which of the hydroxyls are being oxidised. The profile of the curve is not standard Michaelis-Menten kinetics, therefore the Michaelis-Menten model cannot be fitted to the data. OBDG is a non-ionic surfactant and has lipid characteristics; its critical micelle concentration is 25 mM. Interestingly, at 25 mM OBDG, the peak of activity is observed in the oxidation assay (**Figure 69**). However, the only conclusion that can be drawn from this assay is that a hydroxyl-group on OBDG has been oxidised.

The statistics for the crystal structures are summarised in **Table 27**.

Table 27 – Crystal data collection and refinement statistics for *Bc*IRED-NADPH complex and *Nh*IRED-OBDB complex. Numbers in brackets refer to the data in the highest resolution shell.

	<i>Bc</i>IRED-NADPH complex	<i>Nh</i>IRED-OBDB complex
Date	2nd Oct 2014	23 rd May 2014
Beamline	Diamond I03	Diamond I04-1
Wavelength (Å)	0.97950	0.92000
Resolution (Å)	107.25 – 1.81 (1.86-1.81)	67.57 – 2.39 (2.47-2.39)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2
Unit cell (Å)	a = 55.23; b = 62.78; c = 214.50; $\alpha = \beta = \gamma = 90.0$	a = 141.17; b = 153.92; c = 105.57; $\alpha = \beta = \gamma = 90.0$
Molecules in the asymmetric unit	2	8
Unique reflections	69092 (5080)	94683 (7705)
Completeness (%)	99.3 (99.8)	99.8 (100)
R_{merge}	0.03 (0.63)	0.07 (0.55)
$R_{\text{p.i.m.}}$	0.03 (0.41)	0.05 (0.34)
Multiplicity	6.5 (6.4)	6.8 (7.0)
$\langle I/\sigma(I) \rangle$	18.8 (2.7)	15.1 (3.6)
$CC_{1/2}$	0.99 (0.84)	0.99 (0.90)
Overall <i>B</i> factor from Wilson plot (Å ²)	24	39
$R_{\text{factor}} / R_{\text{free}}$ (%)	17.4/19.7	23.0/26.6
r.m.s.d 1-2 bonds (Å)	0.02	0.01
r.m.s.d 1-3 angles (°)	2.00	1.70
Average main chain <i>B</i> (Å ²)	28	51
Average side chain <i>B</i> (Å ²)	34	52
Average water <i>B</i> (Å ²)	40	42
Average ligand <i>B</i> (Å ²)	30	52

4.1.1. Discussion

Once *Bc*IRED and *Nh*IRED had been cloned (**Figure 53**), expressed and purified (**Figure 58**), initial characterisation indicated that both *Bc*IRED and *Nh*IRED

were similar to previously characterised IREDs; in that both enzymes purified as ~60 kDa dimers with a monomeric molecular weights of ~30 kDa. The kinetic, NADPH depletion, assay with substrate **1** and both enzymes confirmed exhibited IRED activity (**Table 25**).



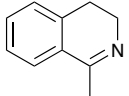
In comparison with SSiRED, both *Bc*IRED and *Nh*IRED are both superior catalysts for the reduction of **1**. However, *Nh*IRED exhibits a significantly higher affinity and catalytic efficiency for **1** compared to SSiRED and *Bc*IRED.

Table 28 – Comparison of kinetic parameters for SSiRED, *Bc*IRED and *Nh*IRED with substrate **1. [a] Kinetic constants for SSiRED were taken from Leipold and co-workers (2013).²⁰**

Enzyme	Kinetic Parameters		
	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (s ⁻¹ M ⁻¹)
SSiRED ^[a]	11.82	0.024	2
<i>Bc</i> IRED	8.32	0.032	3
<i>Nh</i> IRED	1.10	0.013	11

*Bc*IRED is an (*S*)-selective IRED for substrates **1-6** (**Table 26**). Despite *Bc*IRED displaying (*R*)-selectivity for substrate **4**, this is due to an inversion in Cahn-Ingold-Prelog priority rules. *Bc*IRED attacks the same equivalent face of imines substrates. Whereas, *Nh*IRED does not exhibit strict facial preference like *Bc*IRED; *Nh*IRED switches enantioselectivity upon larger substrates (**3-5**). Therefore, predicting enantioselectivity for a given substrate will be more difficult for *Nh*IRED compared to *Bc*IRED. Conversion and enantioselectivity is comparable with SSiRED, *Bc*IRED and *Nh*IRED for substrates **1**, **2** and **6**, however substrates **3-5** were not determined for SSiRED (**Table 29**), as such it is difficult to draw further comparisons between the three IREDs.

Table 29 – Comparison of biotransformations for SSiRED, BcIRED and NhIRED with substrates 1, 2 and 6. Biotransformations with substrates 3, 4 and 5 were not determined for SSiRED. [a] Biotransformation data for SSiRED were taken from Leipold and co-workers (2013).²⁰

Enzyme	Substrate		
	1	2	6
			
SSiRED ^[a]	57%, 95% (S)-	>98%, 98% (S)-	98%, 98% (S)-
BcIRED	98%, >99% (S)-	90%, >99% (S)-	94%, 81% (S)-
NhIRED	80%, >99% (S)-	86%, >99% (S)-	97%, 68% (S)-
percentage conversion, enantiomeric excess and absolute configuration			

Using the dimeric models of SkIRED (PDB entry: 3ZHB), BcIRED and NhIRED, with SkIRED set as the reference; BcIRED superposes onto SkIRED over 547 residues with a root-mean-square deviation (RMSD) of 2.22 Å, whereas NhIRED superposes onto SkIRED over 509 residues with an RMSD of 2.51 Å. The overall fold and secondary structures of the IREDs are conserved. Active sites are formed at the interface between the N-terminal domain of one subunit and the C-terminal domain of the adjacent subunit. The nicotinamide moiety of NADPH protrudes into this interface, which is lined with hydrophobic residues.

4.2. Structure of the IRED from *Amycolatopsis orientalis*

Previously published *apo*- and NADPH-complexes of IREDs have not been insightful regarding the elucidation of the mechanism of IREDs. Therefore, substrate or product complexes would be more useful in identifying binding sites. The imine reductase from *Amycolatopsis orientalis* (*AoIRED*) was interesting in that, from sequence alignment (**Figure 70**), it was predicted to have an asparagine, which is not a typical proton donor, in place of Tyr169-SSIRED and Asp187-SRIRED, which is analogous to Asp187 in *SkIRED*. Whereas, Tyr169-SSIRED and Asp187-SRIRED were previously hypothesised to be potential proton donors.

```

SSIRED      --MSKQSVTVIGLGPMQAMVNTFLDNGHEVTVWNRTASKAEALVARGAVLAPTVEDALS
SRIRED      MGDNRTPVTVIGLGLMGQALAAAFLEAGHTTVWNRSGAKAEQLVSGAVQAATPADAVA
AoIRED      MTDQNLFPVTVAGLGPMGSALAAALLDRGHDTVWNRSPGKAAPLVAKGARQADDIVDAVS
           ..  ***  ***  **  *..  ::*:  **  .*****:  .**  ***:  **  *  **::

SSIRED      ANELIVLSLTDYDAVYAILEPVTGSLSGKVIANLSSDTPDKAREAAKWAAKHGAKHLTGG
SRIRED      ASELVVVCLSTYDNMHDVIGSLGESLRGKVINLTSGSSDQGRETAAWAEKQGVVEYLDGA
AoIRED      ASRLLVVCLADYDALYSALGPAREALRGRVVNLNSGTPKEANEALRWAERHGTGYLDGA
           *..*:  *..*:  **  ::  :  :  *  *..*:  .**.*  :  ..:..*:  **  ::*.  :  *  .

SSIRED      VQVPPPLIGKPESSTYYSGPKDVFDHEDTLKVLTN-ADYRGEDAGLAAMYQAQMTIFW
SRIRED      IMITPPGIGTETAVLFYAGTQSVFEKYEPALKLLGGGTTYLGTDHGMPALYDVSLGLLMW
AoIRED      IMVPPAMVGHPGSVFLYSGSAEVFEEYKETLAGLGD-PVHLGTEAGLAVLYNTALLSMY
           :  :  *  :  *  :  *..*  .**::  ::  *  *  :  *  :  *  :  .**  :  :  ::

SSIRED      TTMLSYYQTLALGQANGVSAKELLPYATMMTSMMPHFLELYAQ--HVDSADYPGDVDRLA
SRIRED      GTLNSFLHGVAVVETAGVGAQQFLPWAHMWLEAIKMFADYAAQIDAGDGKFPANDATLE
AoIRED      SSMNGFLHAAALVGSAGVPAAEFTKLAVDWF LPAVIGQIIKAEAPTIDEGVYPGDAGSLE
           ::  :  :  *  :  *..*  :  :  *  *  :  *  :  *  :  .**::  *

SSIRED      MGAASVDHVLHTHQDAGVSTVLPAAVAEIEFKAGMEKGAENFSLSLIEVLKPPAV
SRIRED      THLAALKHLVHESEALGIDAELPKYSEALMERVISQGHAKNSYAAVLKAFRKPSE
AoIRED      MNVTTLKHIIIGTSQEQVDTEIPVRNKELLDRAVAAGFGESSYYSVIELWR----
           :  :  .**::  :  *..:  :  *  :  :  :  *..:  *  :  :  :  :

```

Figure 70 – Protein sequence alignment of SSIRED, SRIRED and *AoIRED*, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interested are highlighted in red and yellow. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹ *AoIRED* shares 78.3% and 65.0% sequence similarity with SSIRED and SRIRED respectively.

The synthetic gene for *AoIRED* was codon optimised for expression in *E. coli* and synthesised by Biomatik (Cambridge, Ontario, Canada). The gene was sub-cloned into pET-28a (+) vector and supplied by Aleku and co-workers.

An expression strain of *E. coli*, BL21(DE3), was transformed (**3.1.8**) by the plasmid. A 5 mL expression test (**Figure 71**) was performed to identify the optimal growth temperature. Optimal over expression was identified at 16°C, when the cells were induced with 1 mM IPTG. Despite over expression being observed at

30°C and 37°C, the ratio between insoluble protein to soluble protein was significantly greater at the higher temperatures compared to 16°C.

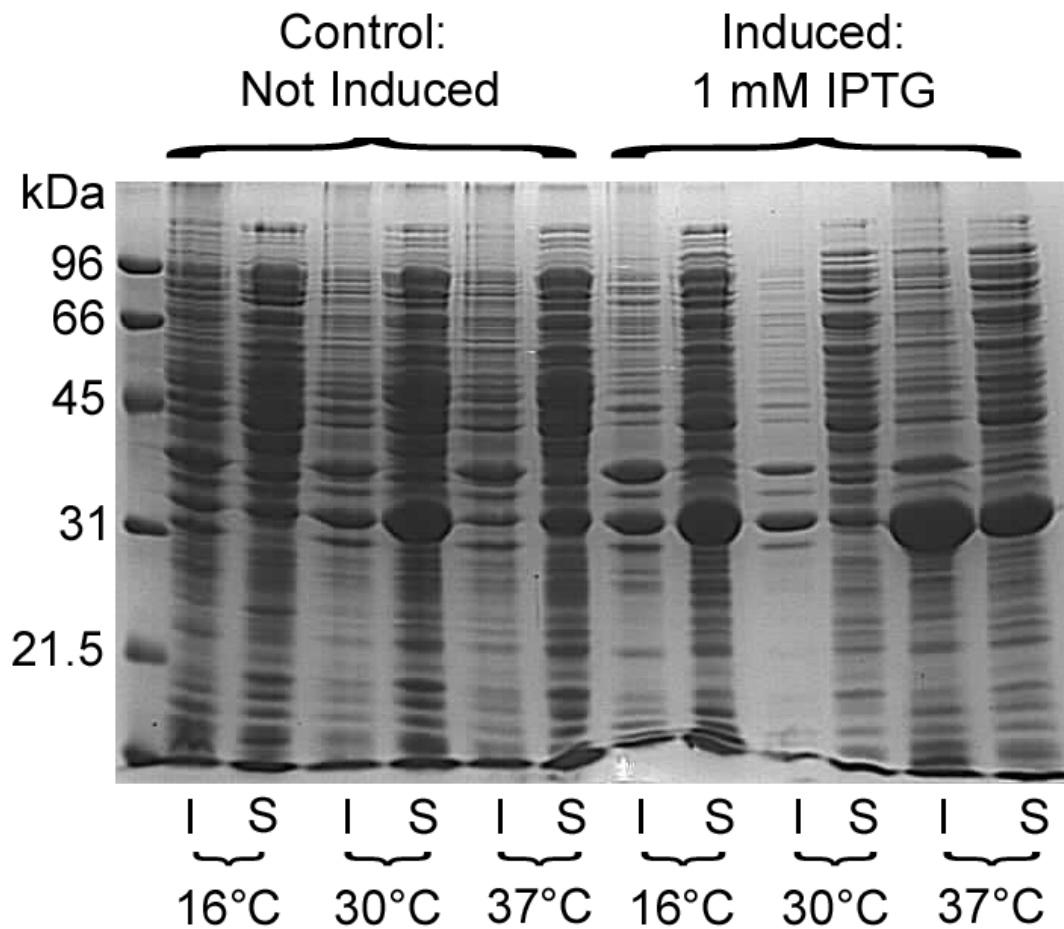


Figure 71 – The expression test for *AoIRED*, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of *AoIRED* is 33 kDa, including the 6His-tag and thrombin cleavage site.

Using the optimal growth conditions, gene expression was scaled up. Once the cells had been harvested and lysed, the protein was purified by Ni-affinity and SEC (Figure 72) with protein purity assessed by SDS-PAGE (Figure 73).

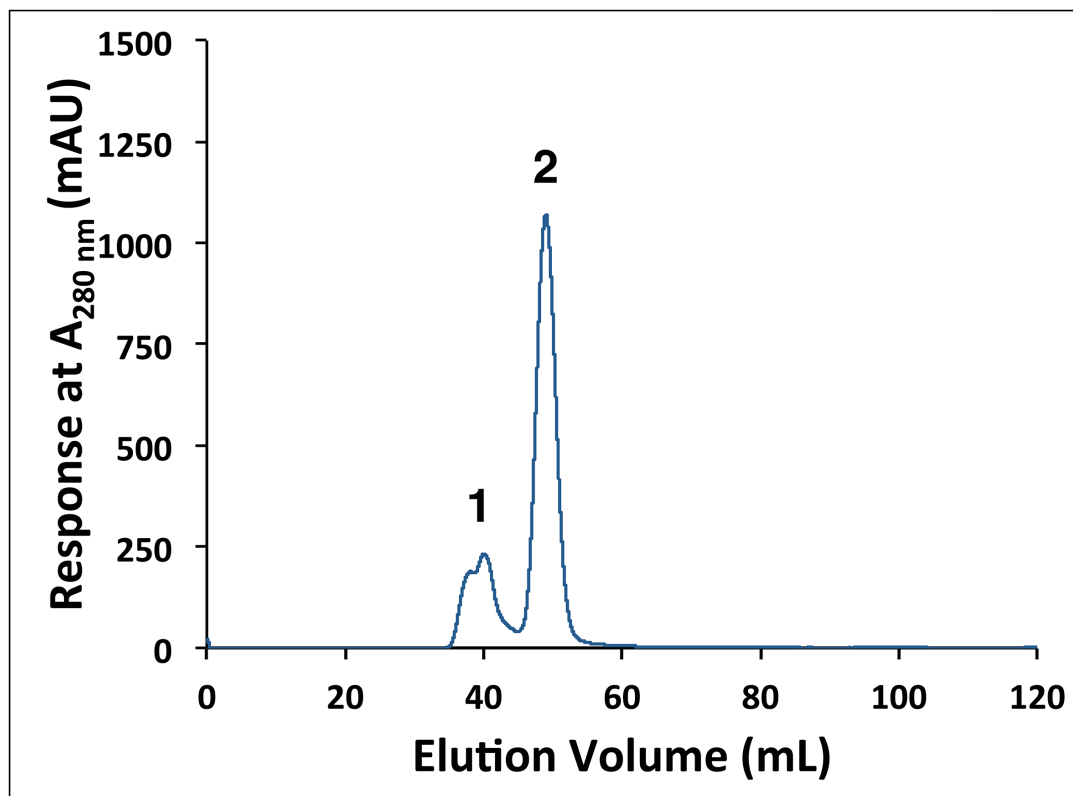


Figure 72 – Chromatogram of the purification of AoIRED by SEC. Peak 2 elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa.

The elution profile from the chromatogram obtained from SEC indicates a species eluting between 50-55 mL, with a molecular weight suggested by calibration standards of ~60 kDa. The SDS-PAGE analysis indicates a monomeric weight of ~30 kDa. The purity of the protein after SEC was deemed sufficiently pure to proceed with further characterisation.

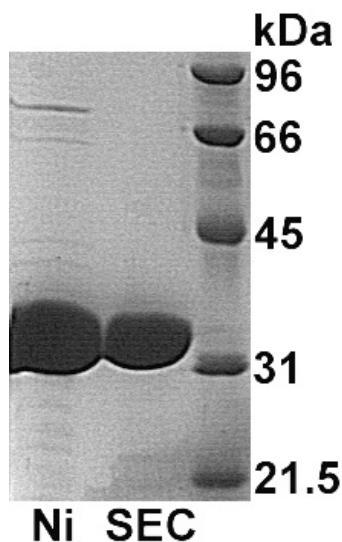
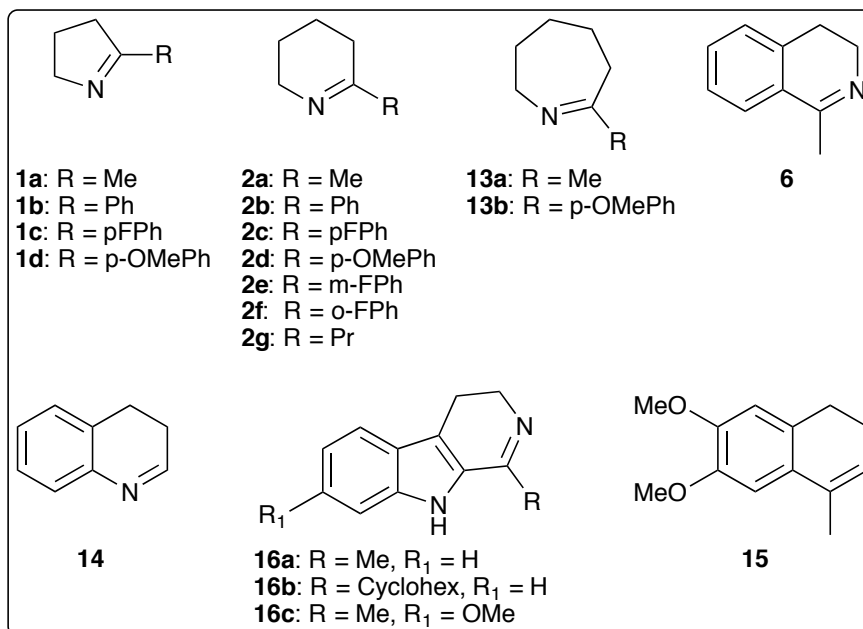


Figure 73 – SDS-PAGE analysis of nickel and SEC purified AoIRED. Ni is the sample obtained from the pooled purest fractions from nickel affinity chromatography. SEC corresponds to a sample of peak 2 from Figure 72.

Aleku and co-workers (personal communication) had performed biotransformations, using AoIRED with a broad range of substrates (**Figure 74**). AoIRED is an incredibly versatile IRED for obtaining very good yields for bulky substituted pyrrolines (substrates **1b-1d**) and piperidienes (substrates **2b-2f**), as well as bulky imines such as azepines (substrates **13a** and **13b**) and β -carbolines (substrate **16b**). The enantioselectivity of AoIRED is generally good across most substrates, however there are substrates where complete conversion is obtained with very little enantiomeric excess observed (substrate **2f**) and the opposite case is also seen with substrate **16a**. Interestingly, despite structural similarities between substrates **2b** and *para*-substituted variants **2c** and **2d**, different enantiomers are obtained where CIP-inversion is not the case. One reaction of interest was with imine **6**, 1-methyl-3,4-dihydroisoquinoline (MDQ) which formed the amine product 1-methyl-1,2,3,4-tetrahydroisoquinoline (MTQ). However, differences in enantioselectivity were observed when using whole cells with AoIRED over expressed, or with purified protein stored at 4°C for several days. Biotransformations with **6** were performed with whole cells with AoIRED over expressed, cell lysate, freshly purified AoIRED, and AoIRED stored for different periods of time.



Substrate	Conversion (%)	<i>e.e</i> (%)	Absolute Configuration
1a	96	n.d.	(<i>S</i>)-
1b	100	95	(<i>S</i>)-
1c	98	>99	(<i>S</i>)-
1d	85	>99	(<i>R</i>)-
2a	>99	98	(<i>S</i>)-
2b	>99	40	(<i>S</i>)-
2c	>99	60	(<i>R</i>)-
2d	77	90	(<i>R</i>)-
2e	>99	71	(<i>S</i>)-
2f	>99	4	(<i>R</i>)-
2g	>99	>99	(<i>R</i>)-
13a	95	>99	(<i>S</i>)-
13b	76	96	n.d.
6	100	81	(<i>S</i>)-
14	100	n/a	n/a
15	50	79	(<i>S</i>)-
16a	5	>99	(<i>R</i>)-
16b	66	71	(<i>R</i>)-
16c	15	79	(<i>S</i>)-

Figure 74 – Prochiral imines reduced by AoIRED. Performed by Aleku and co-workers (personal communication). n.d. = not determined. Substrate 1a had lack of baseline to accurately determine *e.e.* substrate 14 is achiral.

Biotransformations performed with whole cells over-expressing AoIRED and substrate **6** result in (*S*)-MTQ with an enantiomeric excess of 85%. The enantiomeric excess decreases slightly when crude cell extract is used with the

same reaction with substrate **6**. The enantiomeric excess is further slightly decreased when freshly nickel purified *Ao*IRED is used to perform the biotransformation with substrate **6**. When the biotransformation with substrate **6** is performed with the same freshly purified protein that has been left at 4°C for 4 h, the enantiomeric excess is significantly reduced to near racemic quantities. However, when the biotransformation with substrate **6** is performed with the same purified protein that has been left at 4°C for 8h, it results in (*R*)-MTQ with an enantiomeric excess of ~50%. Interestingly, when the biotransformation with substrate **6** is performed with purified protein that has been left at 4°C for greater than 24 h, it results in (*R*)-MTQ with enantiomeric excess of >99%. The results are summarised in **Figure 75**.

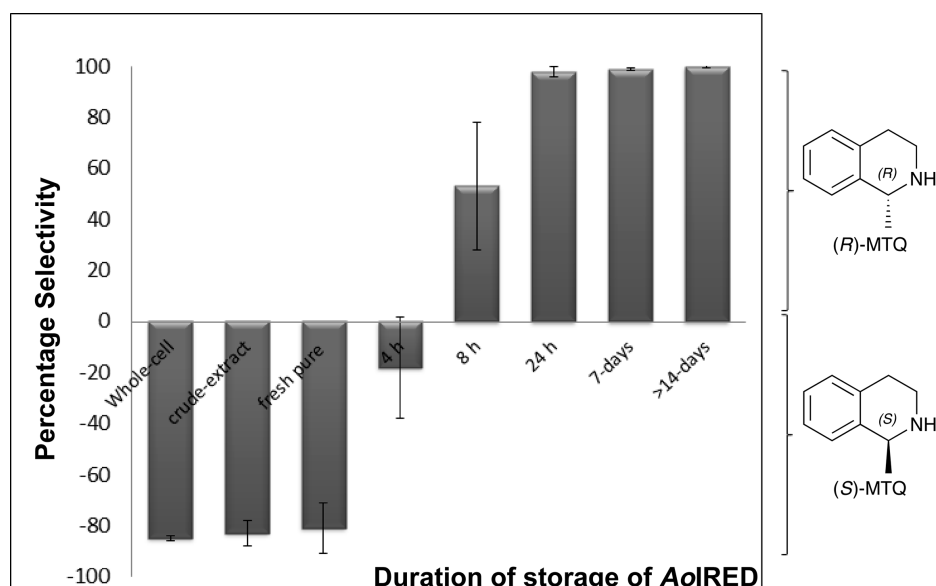


Figure 75 – Enantioselective switch of *Ao*IRED towards the imine substrate **6** (MDQ). Whole cells, over expressing *Ao*IRED, cell lysate and freshly purified *Ao*IRED yielded (*S*)-MTQ, whereas purified *Ao*IRED that had been stored at different periods yielded (*R*)-MTQ. Enantiomeric excess is represented by percentage selectivity. Negative selectivity represents (*S*)-selectivity, whereas positive selectivity represents (*R*)-selectivity. Figure adapted from Aleku and co-workers (2016).¹⁰⁷

A native gel (**3.2.1**) was performed to attempt to track conformational changes upon purification (**Figure 76**). Once *Ao*IRED had been purified by Ni-affinity, the purified protein was pooled and stored at 4°C. Samples were taken every 30 min for 5 h and to prevent the protein from ageing, samples were mixed

with the loading dye and flash cooled with liquid nitrogen. Once all samples had been collected, samples were rapidly thawed prior to loading onto the gel.

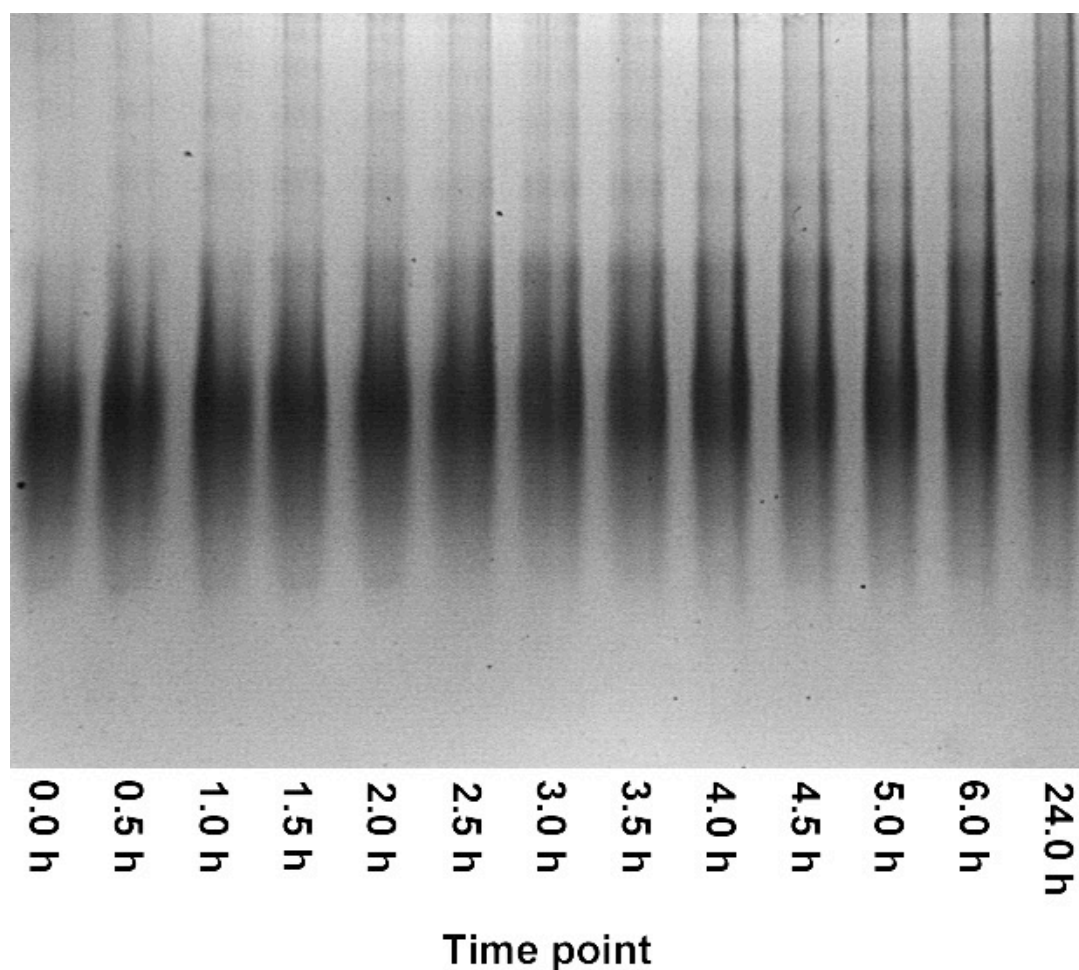


Figure 76 – Native gel analysis of purified AoIRED upon different lengths of storage at 4°C. This figure has also been published in the supporting information of Aleku and co-workers (2016).¹⁰⁷

The gradual smearing observed with the time point native gel (**Figure 76**) could be associated with the inversion of selectivity towards substrate **6**. One explanation to the stereo-inversion could be due to the aid of chaperone proteins, such as GroEL, within *E. coli* whole cells and cell lysate. Upon purification, the chaperones are removed and the effects of the chaperones aid rapidly diminish over time. However, this is only speculative without further investigation.

Kinetic parameters were determined for AoIRED (**Table 30**) using purified AoIRED, performed by Aleku and co-workers (personal communication).

Table 30 – Kinetic parameters for selected AoIRED substrates.

Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1} mM^{-1}$)
1a	0.87 ± 0.43	0.15 ± 0.08	0.17 ± 0.18
2a	0.65 ± 0.11	0.35 ± 0.01	0.53 ± 0.11
2b	0.22 ± 0.01	0.12 ± 0.00	0.38 ± 0.08
2c	0.51 ± 0.16	0.14 ± 0.01	0.27 ± 0.07
6	0.72 ± 0.05	0.74 ± 0.02	1.03 ± 0.04
14	0.36 ± 0.04	0.94 ± 0.01	2.58 ± 0.30

AoIRED was further characterised by structural investigations in an attempt to identify the active site. Purified native AoIRED was subjected to crystal trials using commercially available screens in 96-well sitting drop format (3.5.1) at a protein concentration of 20 mg mL^{-1} in buffer B (3.3). Crystals appeared (Figure 77) after 2 days at 18°C in conditions containing 0.2 M calcium acetate, 0.1 M Tris pH 9.0, 8% (w/v) PEG 550 MME and 8% (w/v) PEG 20K (CSS II B6 with 0.1 M Tris pH 9.0).

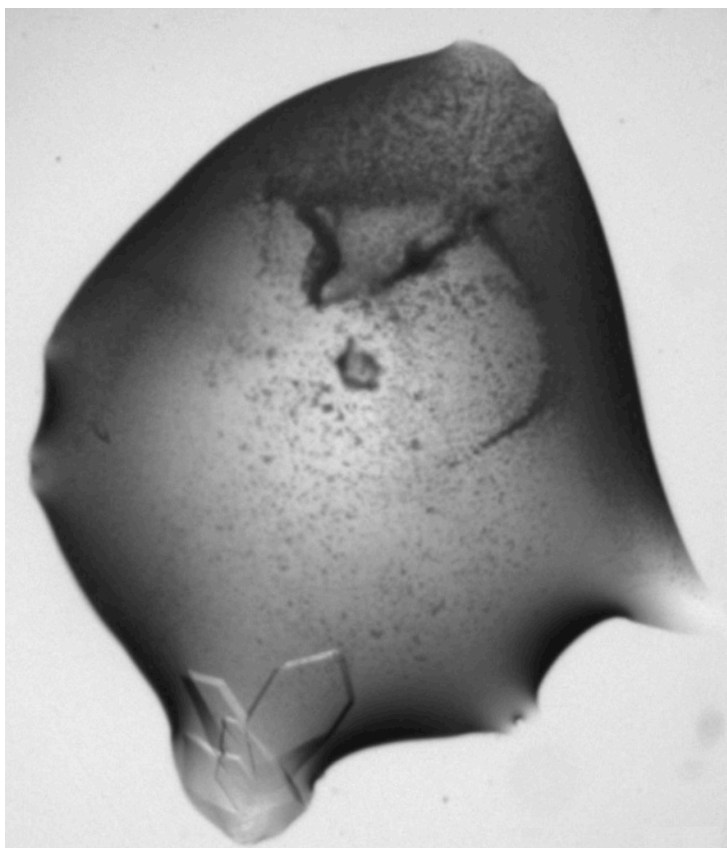


Figure 77 – Crystals of *apo*-AoIRED, at a protein concentration of 20 mg mL^{-1} , obtained in a 300 nL crystallisation drop in CSS II B6 with 0.1 M Tris pH 9.0 from crystal trial screening.

These crystals were subsequently optimised in 24-well hanging drop format using a 1 mL mother liquor volume with a 2 μ L crystallisation drop (3.5.3), in an attempt to reproduce larger crystals with a reduced clustered morphology. The best crystals (**Figure 78**) were obtained in the same conditions as the initial hit from crystal trial screening. These crystals appeared within 24 h at 18°C, however the best diffracting crystals came from crystals matured after one week.

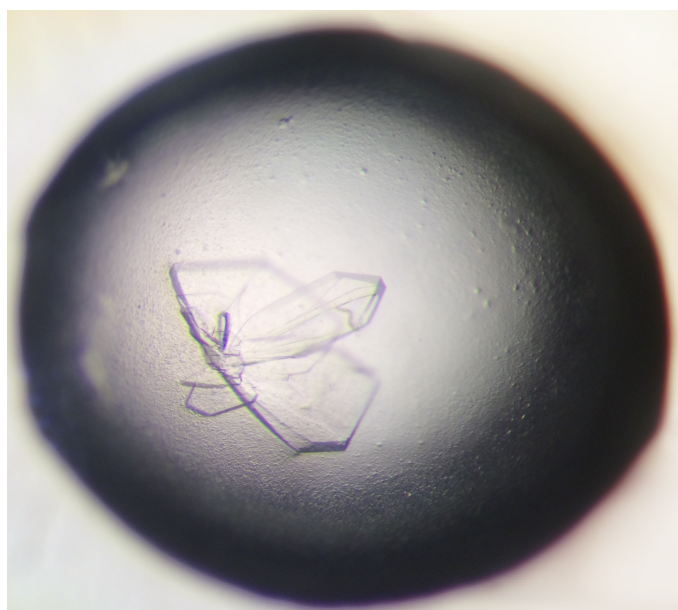


Figure 78 – Optimised, using hanging drop format, crystals of *apo*-AoIRED in conditions containing 0.2 M calcium acetate, 0.1 M Tris pH 9.0, 8% (w/v) PEG 550 MME and 8% (w/v) PEG 20K.

Apo-AoIRED crystal structure was obtained by fishing the *apo*-crystals without cryoprotectant, which were then flash cooled with liquid nitrogen. Co-crystallisation by adding NADPH to the conditions that yielded the *apo*-crystals did not yield any further crystals. The NADPH complex was obtained by soaking the *apo*-crystals in a cryoprotectant in conditions containing the mother liquor (0.2 M calcium acetate, 0.1 M Tris pH 9.0, 8% (w/v) PEG 550 MME and 8% (w/v) PEG 20K) with additional 20% (v/v) ethylene glycol and 10 mM NADPH for 5 min prior to flash cooling with liquid nitrogen.

Full datasets had been collected for *apo*-AoIRED and AoIRED-NADPH complex to 1.5 Å and 2.0 Å respectively. The Xia2 auto processed data for AoIRED-NADPH complex was acceptable, however the dataset for *apo*-AoIRED required reprocessing in XDS. The datasets were solved by molecular replacement using Molrep in the space groups $C2\ 2\ 2_1$ and $P2_1\ 2_1\ 2_1$, for *apo*-AoIRED and AoIRED-

NADPH complex respectively, using a monomer *BcIRED*-NADPH complex as the molecular replacement model. The solution required searching for the N- and C-terminal domains separately. Iterative cycles of model building and refinement were performed. Water molecules were added last and the model was again refined and checked. For the *AoIRED*-NADPH complex, residual density at the cleft formed between the N- and C-terminal domains in the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was assigned as NADPH. The *apo*- crystal structure is composed of one monomer in the asymmetric unit, however a homodimer is formed when symmetry related units are viewed. The crystal structure of the NADPH complex is composed of two monomers in the asymmetric unit arranged as a single homodimer. The secondary and tertiary structures of *AoIRED* are similar to previous IRED structures, now termed the IRED fold. The final structures were refined to R_{factor}/R_{free} values of 18.1%/21.4% and 17.3%/22.5% for *apo*-*AoIRED* and *AoIRED*-NADPH complex respectively. The *apo*-model was complete from residues 5 to 290 (Figure 79).

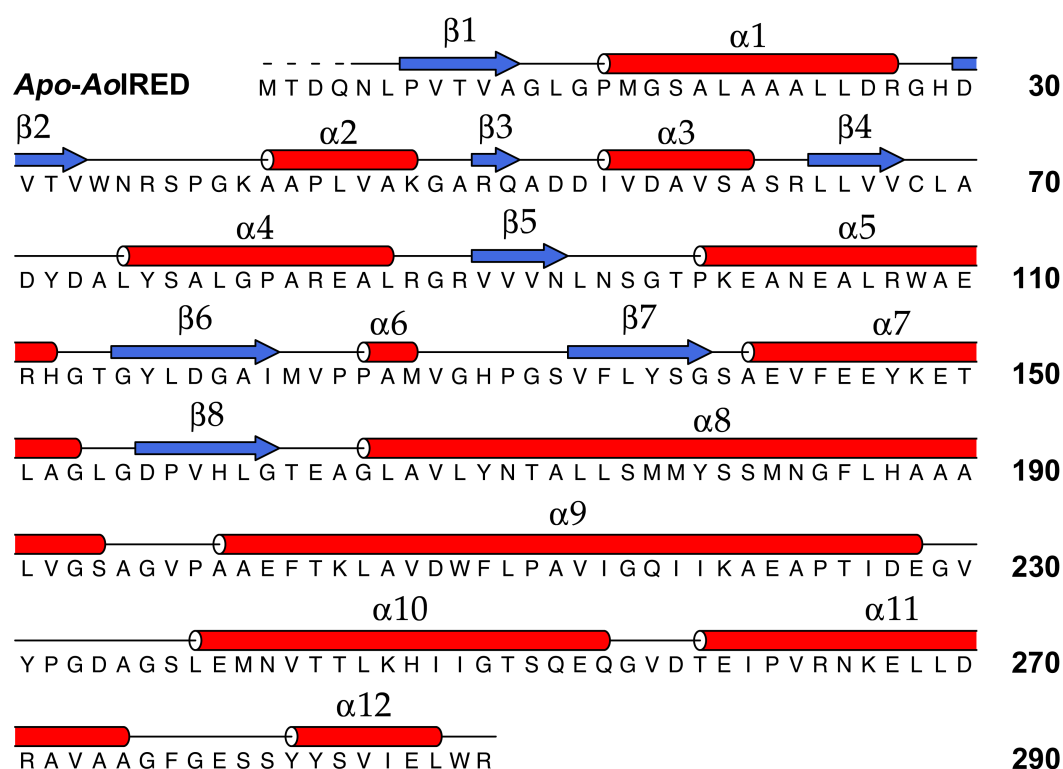


Figure 79 – The primary structure of *apo*-*AoIRED* with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.

The model for the NADPH complex had a chain break from residues 227 to 234, the model was otherwise complete between residues 2 to 227 and 234-290 (Figure 80).

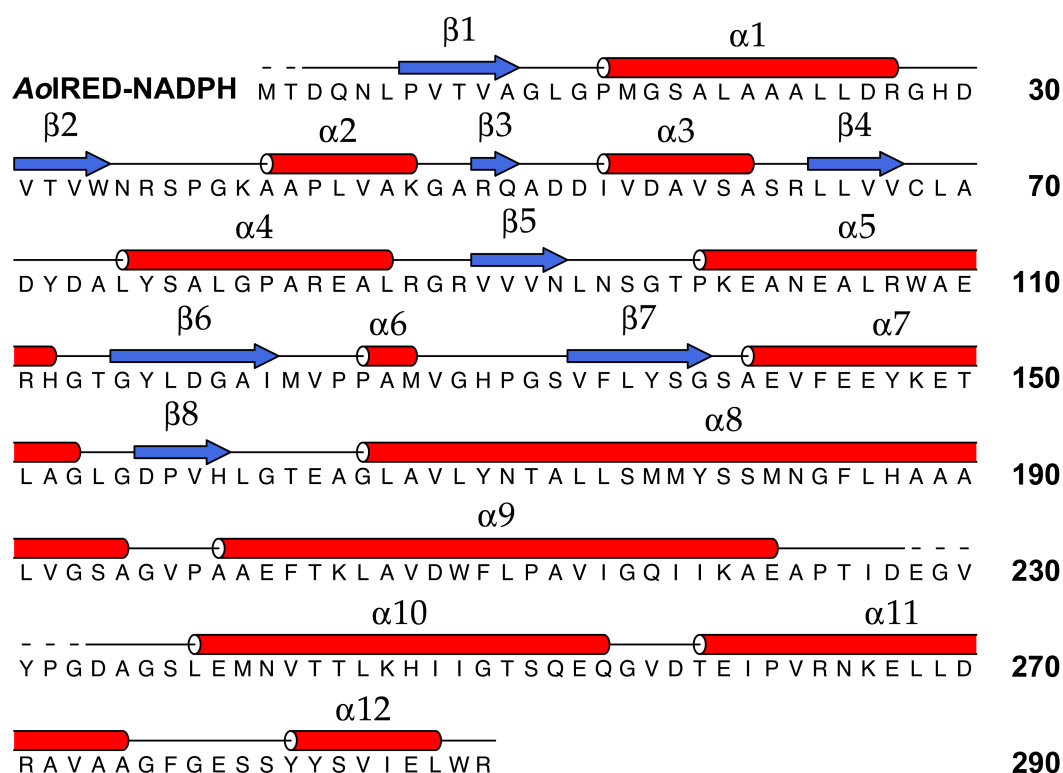


Figure 80 – The primary structure of AoIRED-NADPH complex with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.

DALI analysis indicates that the closest structural homologue of *apo*-AoIRED is the *BcIRED*-NADPH complex that has a sequence homology of 37%, with monomers superposing with an RMSD of 2.4 Å over 283 residues out of 289 residues. However besides the other IRED structures, the next closest structural homologue of *apo*-AoIRED is a dehydrogenase-like protein, Glyoxylate reductase 1 from *Arabidopsis thaliana* (PDB entry: 3DOJ) with some sequence homology to hydroxyisobutyrate dehydrogenases. Glyoxylate reductase 1 shares 22% sequence homology to *apo*-AoIRED and superposes with an RMSD of 2.6 Å over 210 residues out of 292.

Interestingly, DALI analysis of the AoIRED-NADPH complex indicates the closest structural homologue is different to that of *apo*-AoIRED. The closest structural homologue of the AoIRED-NADPH complex is a 6-phosphogluconate

dehydrogenase with hydroxyisobutyrate dehydrogenase activity from *Geobacter metallireducens* (PDB entry: 3PEF) that shares sequence homology of 24% with the AoIRED-NADPH complex and superposes with an RMSD of 4.6 Å over 213 residues out of 287.

The dimeric model of *apo*-AoIRED was generated by pairing the symmetry related units. The dimeric model of *apo*-AoIRED superposes onto crystal structure of AoIRED-NADPH complex with an RMSD of 1.23 Å over 543 residues. PISA analysis of the dimeric model of *apo*-AoIRED indicated a dimer interface of 3883 Å², consisting of 14 hydrogen bonds and 3 salt bridges that spontaneously dimerised with a ΔG value of -78.2 kcal mol⁻¹. Whereas, PISA analysis of AoIRED-NADPH complex indicates a larger dimer interface of 4165 Å², that consists of 15 hydrogen bonds and 13 salt bridges that spontaneously dimerises with a ΔG value of -85.3 kcal mol⁻¹. This indicates that formation of the NADPH-complex is more favourable and more stable with a greater number of hydrogen bonds and salt bridges formed, compared to the *apo*-form.

Interestingly the distance observed between the C4N of NADPH to the OD1 of Asn171 (**Figure 81**) is the same 8 Å as observed with the hydroxyl of Asp187-SkIRED to the C4N-NADPH. Upon closer inspection, there is a water molecule within hydrogen bonding distance of residue Asn171.

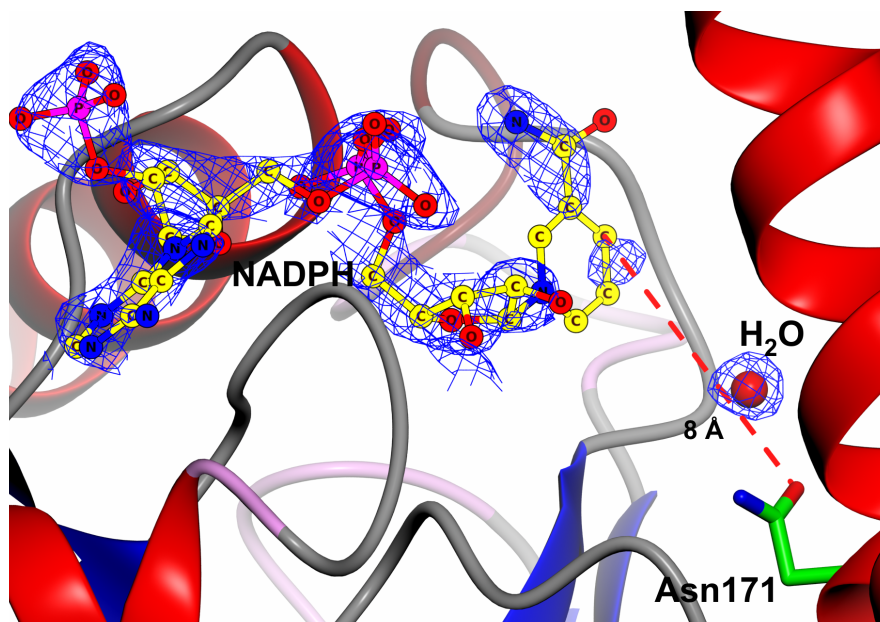


Figure 81 – The active site of AoIRED, an Asn-IRED, in complex with NADPH with a water molecule within hydrogen bonding distance of Asn171. The distance observed between C4N-NADPH and OD1-Asn171 is 8Å. The electron density for NADPH is the $mF_o - DF_c$ map contoured to 3σ obtained from refinement prior to modelling and refining NADPH. The carbon atoms of NADPH have been highlighted in yellow and in circles for illustrative purposes.

Aleku and co-workers (personal communication) performed biotransformations with purified AoIRED that had been stored at 4°C for 24 h with the substrates listed in **Figure 82**. Despite structural similarities with substrates **4** and **5**, AoIRED exhibits different enantioselectivities for the two substrates. Differences observed with enantioselectivity between freshly purified AoIRED and ‘aged’ AoIRED by storage at 4°C with substrate **6** were interesting, as seen with previously with **Figure 75**. Also, AoIRED could accept bulky substrates such as the dibenzazepine **12**, whereas other IREDs such as SSIREd and SRIREd could not accept such large substrates. As such, structural investigations with substrates **4**, **5**, **6** and **12** would hopefully reveal further insights into the mechanism.

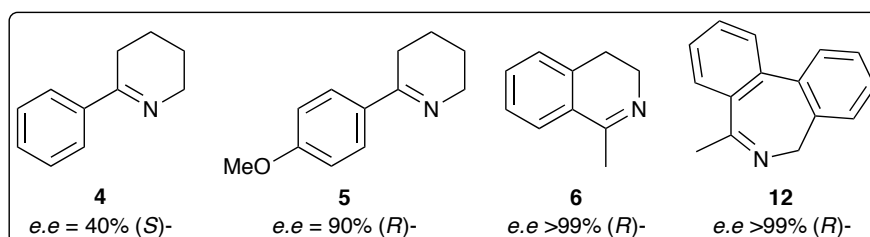


Figure 82 – Substrates of interest with AoIRED for structural investigations.

Apo-AoIRED crystals were reproduced in 24-well hanging drop format. Once the crystals had fully matured after one week, different crystallisation drops with matured crystals were soaked (3.5.2.2) with racemic amine product compounds of **6** for 18 h, **4** and **5** for three days, and **12** for five days. These crystals were fished without further cryoprotectants and rapidly flash cooled with liquid nitrogen.

Full datasets were collected for each soak experiment to 1.5 Å, 1.4 Å 1.5 Å and 1.7 Å resolution for the amine products complexes of **4**, **5**, **6** and **12** respectively. In all four cases, the auto processed data by Xia2 were adequate for structure solution. The crystal structures for all four product complexes were solved by molecular replacement using Molrep in the space group $C2\ 2\ 2_1$ using the previously obtained *apo-AoIRED* crystal structure as the molecular replacement model with similar cell dimensions to each other and the *apo*-structure. Iterative cycles of model building and refinement were performed. Water molecules were added last and the models were refined and checked. In each case, the residual density between the dimer interface in the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was modelled as the ligand of interest.

The enantiomers of the amine products that were bound in each respective crystal structures were the same enantiomers as obtained through biotransformations. From the biotransformation data, a specific face must have been presented to NADPH for the specific enantiomer to be formed; the faces presented in **Figure 83** in the highlighted boxes are the correct facial presentation. NADPH at site **S2** has been superposed from the previously obtained NADPH complex. The faces presented for **4** and **12** are the same, interestingly they are bound in the same place at the front of the active site, at site **S1**. However, the opposite face for **5** is presented to NADPH and it was found to be bound near the rear of the active site at sites **S3**. The product for **5** was also found at site **4**, which is clashing, with where NADPH would be usually found at site **S2**. As such, due to the binding sites being far from the nicotinamide ring of NADPH; the amine product complexes of **4**, **5** and **12** could simply be crystallographic artefacts.

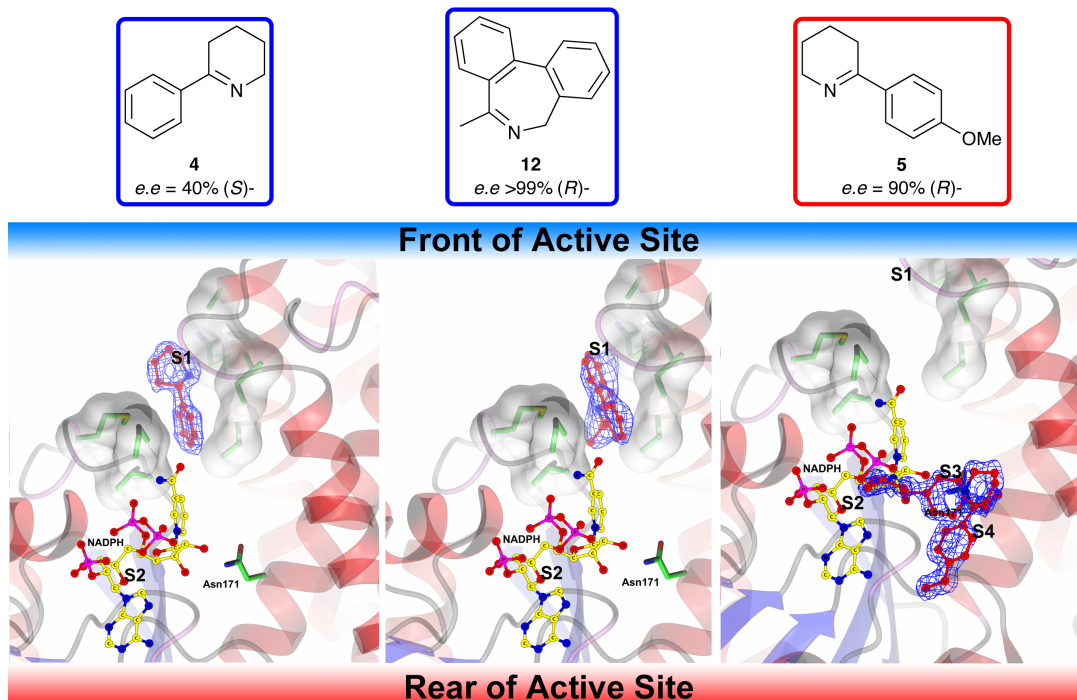


Figure 83 – Binding sites obtained for products of 4, 5 and 12, obtained by soaking *apo*-crystals. The electron density for products of 4, 5 and 12 is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining products of 4, 5 and 12. NADPH is represented in yellow circles and ligands 4, 5 and 12 are represented in red circles for illustrative purposes. Binding sites of products 4 and 12 are bound towards the front of the active site at site S1, the residues that surround both products are hydrophobic. One of the two potential binding sites, site S4, of product 5 clashes with that of the NADPH binding site, site S2. The binding site for product 5 at site S3 is situated towards the rear of the active site and similar to products bound at site S1, are all too far from the C4N of NADPH to interpret the mechanism.

The product of 6, MTQ, was bound in a different location compared with 4, 5 and 12. (R)-MTQ was bound between the helical bundle of the adjacent dimerising subunits C-terminal domain.

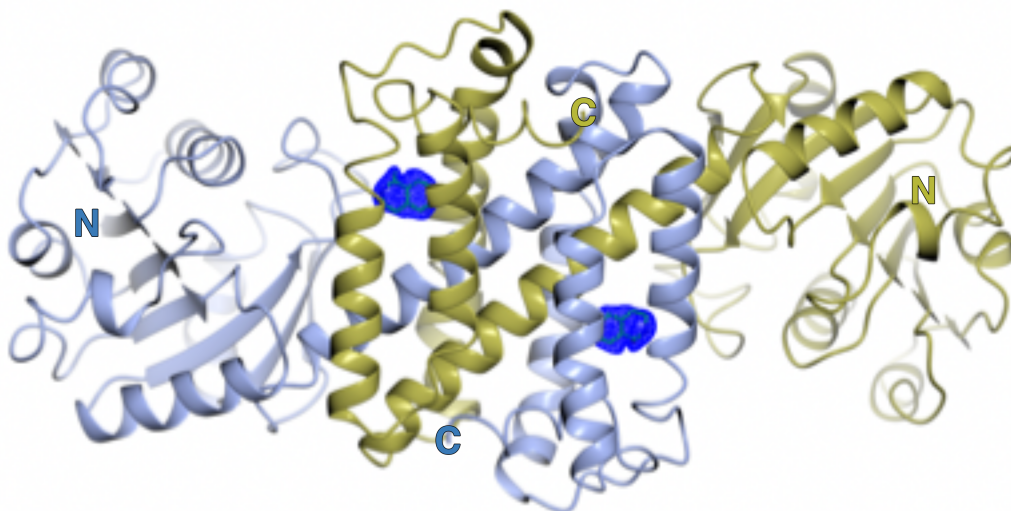


Figure 84 – The dimeric model of the AoIRED (*R*)-MTQ complex. Subunits A and B are in blue and gold respectively. The electron density for (*R*)-MTQ is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining (*R*)-MTQ.

The binding site of (*R*)-MTQ is very hydrophobic (**Figure 85**), consisting of residues Leu175, Met178, Tyr179 and Met182 from one subunit, and residues Asn241, Thr244, Tyr282 and Tyr283 from the adjacent dimer subunit. A water molecule is seen to be hydrogen bonding to the hydroxyl of Tyr179 (3.7 Å), amino group of Asn241 (3.6 Å), the hydroxyl of Thr244 (2.8 Å) and the amine of (*R*)-MTQ (3.4 Å). Besides the network of hydrogen bonding to the water molecule, there is no notable direct hydrogen bonding or electrostatic interactions between the residues in the binding site. The distance between the imine bond of (*R*)-MTQ and NADPH from the NADPH complex is too far for catalysis. However it is possible that IREDs in general can close up to form tighter complexes, as illustrated by previous structures of *apo*-AoIRED and the AoIRED-NADPH complex, which have dimer interfaces of 3883 Å² and 4165 Å² respectively, that suggests there is mobility between the domains and may result in a more closed complex.

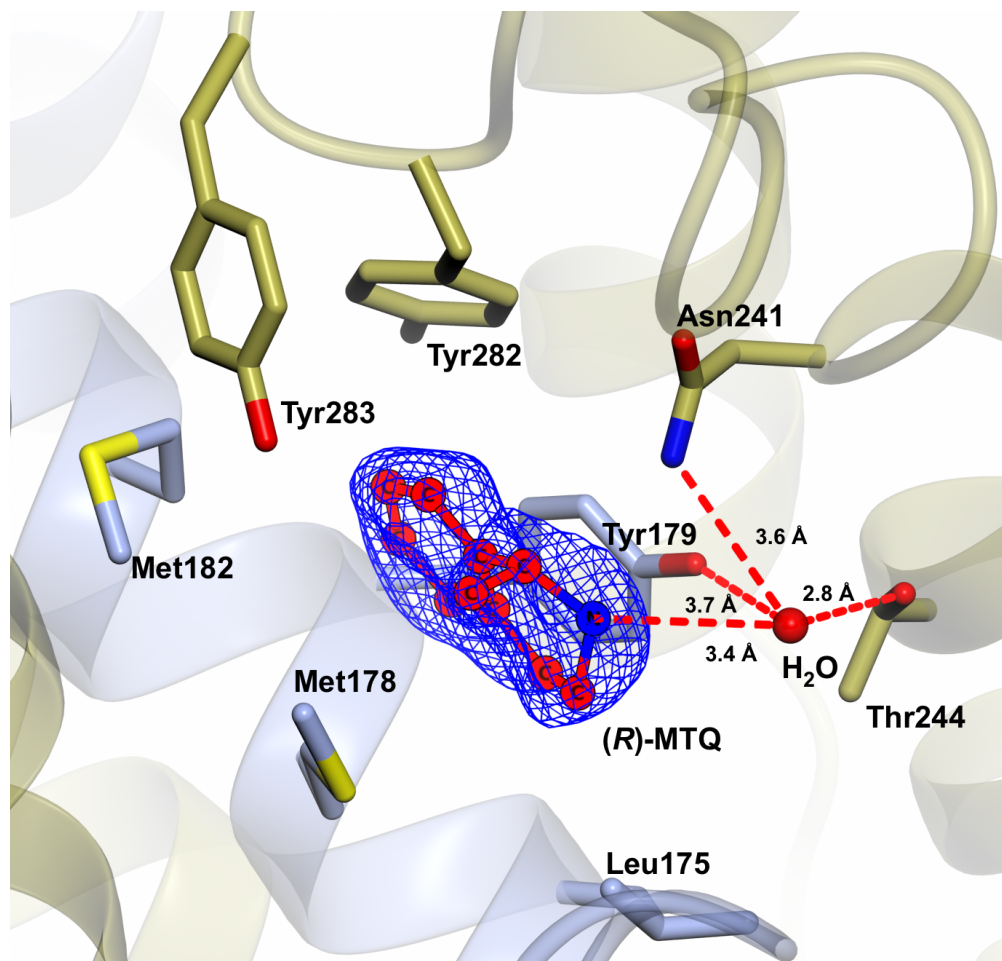


Figure 85 – The hydrophobic binding pocket for (*R*)-MTQ. Subunits A and B are in blue and gold respectively. (*R*)-MTQ is shown in red circles for illustrative purposes. The electron density for (*R*)-MTQ is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining (*R*)-MTQ.

*Ao*IRED is capable of existing in different conformations (Figure 86). Upon comparing the *apo*-structure and NADPH complex, it is clear that a large conformational change has taken place between in the region between residues 207 and 247. This shows that *Ao*IRED is very flexible and dynamic. Therefore, in this case, obtaining a ternary complex by soaking was not possible, as well as the large conformational change, a large change in cell dimension and space group is likely to occur, which would result in cracked or dissolved crystals impairing diffraction quality.

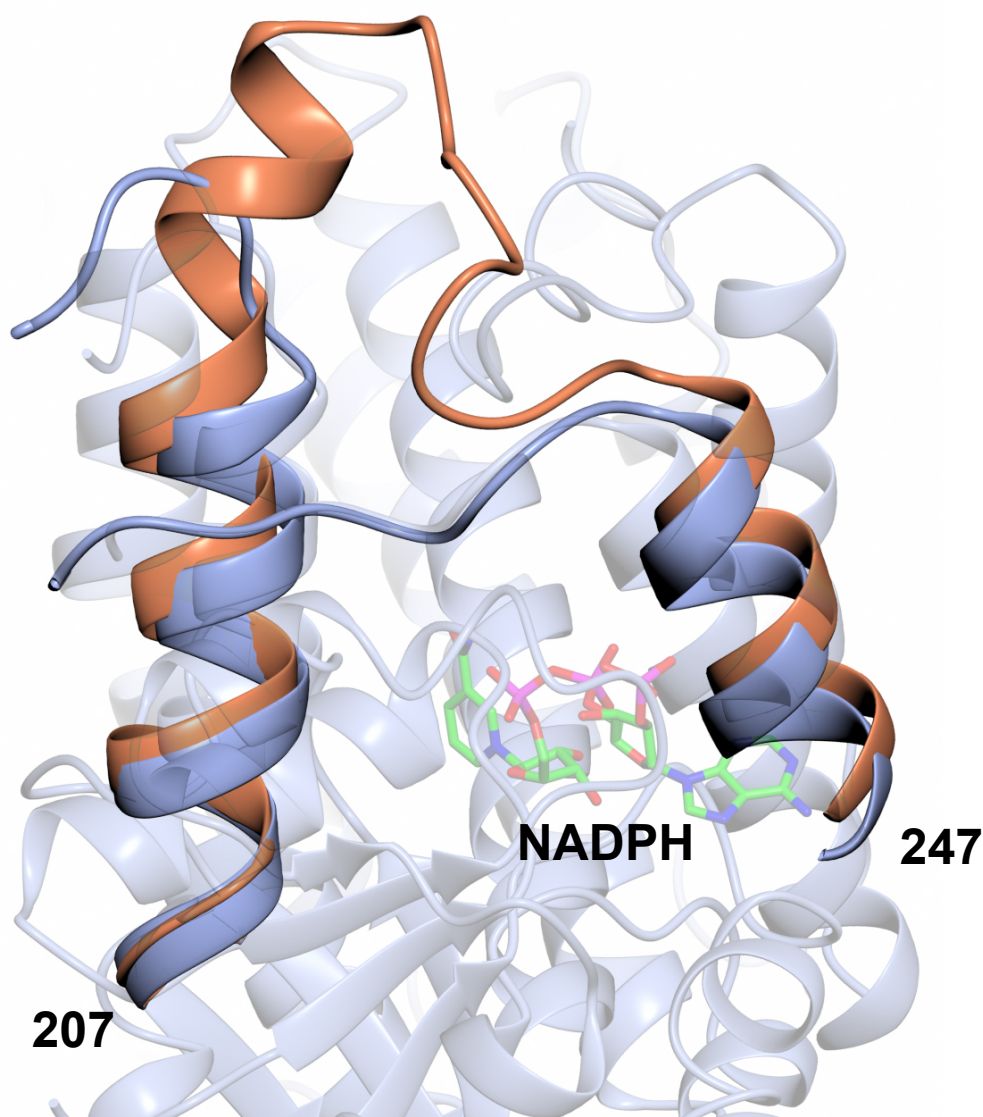


Figure 86 – Comparison between *apo*-AoIRED (coral) and the AoIRED-NADPH complex (ice blue). The region between residues 207 to 247 is highlighted.

To obtain a ternary complex, purified native AoIRED was rescreened with commercially available crystallisation screens with the addition of 1 mM NADPH and 1 mM *rac*-MTQ added as an additive into the mother liquor (3.5.2.3), at a protein concentration of 25 mg mL⁻¹ in buffer B (3.3). New crystal hits were observed, different to the hits observed with *apo*- screening, with moderately sized crystals (Figure 87) appearing after two weeks at 18°C in conditions containing 0.2 M lithium sulfate, 0.1M BisTris pH6.5, 25% (w/v) PEG 3350, 1 mM NADPH and 1 mM *rac*-MTQ.

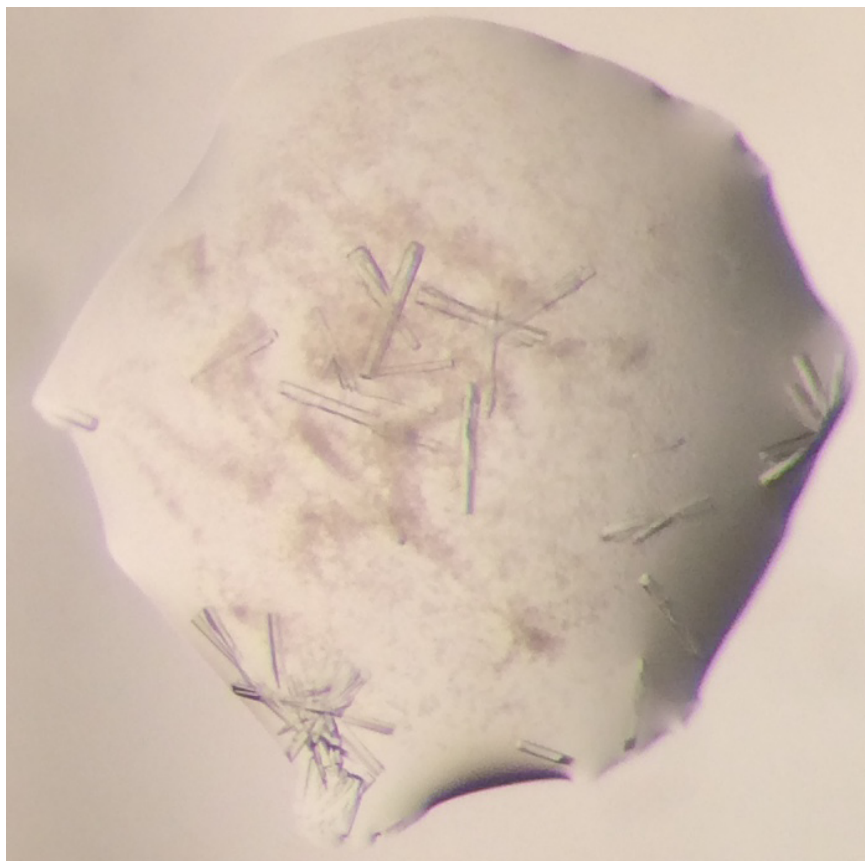


Figure 87 – Crystals of AoIRED in complex with NADPH and MTQ, at a protein concentration of 25 mg mL^{-1} in buffer B (3.3), obtained in a 300 nL crystallisation drop in INDEX (with 1 mM NADPH and 1 mM MTQ) G3 from crystal trial screening.

These crystals were optimised in 48-well sitting drop format (3.5.3) utilising a mother liquor volume of $150 \mu\text{L}$ and a $2 \mu\text{L}$ crystallisation drop, in an attempt to reproduce larger crystals. The best crystals (**Figure 88**) appeared after two weeks at 18°C were obtained in conditions 0.25 M lithium sulphate, 0.1 M BisTris pH 6.5, 25% (w/v) PEG 3350, 1 mM NADPH and 1 mM MTQ at a protein concentration of 25 mg mL^{-1} in buffer B (3.3).

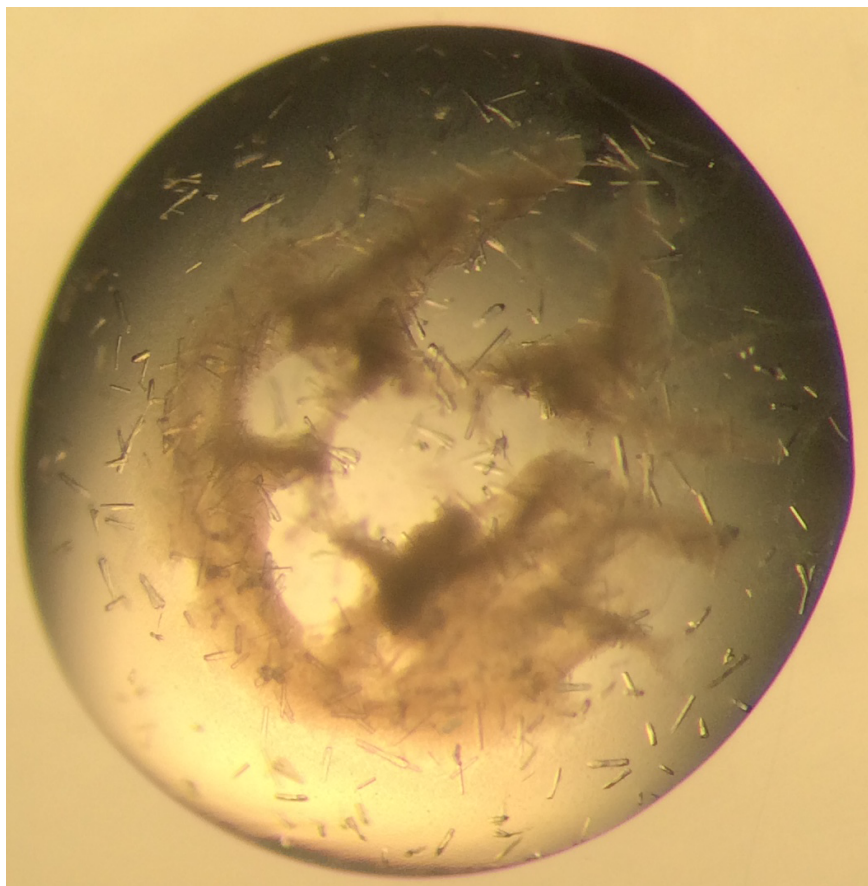


Figure 88 – Optimised crystals of AoIRED in complex with NADPH and MTQ using sitting drop format with a 2 μ L crystallisation drop, in conditions containing 0.25 M lithium sulphate, 0.1 M BisTris pH 6.5, 25% (w/v) PEG 3350, 1 mM NADPH and 1 mM MTQ.

Initially a full dataset was collected using 100% beam transmission, however upon solving and refining the model, it was clear that AoIRED was in a ‘closed’ state and there was only partial NADPH density. The ligands could have been excited and liberated by the high-energy X-ray beams. Therefore, subsequent crystals were soaked with compound NADPH and *rac*-MTQ for 5 min prior to flash cooling with liquid nitrogen without further cryoprotectants. Along with the additional soak, a full dataset was collected using 10% beam transmission (3.6.1) combined with helical data collection (3.6.2).

A full dataset was collected to 2.1 Å resolution with the Xia2 auto processed data suitable for further structure solution and refinement. This crystal structure of AoIRED was solved by molecular replacement using Molrep in the space group $P2_1 2_1 2_1$ with the previous AoIRED-NADPH complex as the molecular replacement model. Iterative cycles of model building and refinement were performed. Water molecules were added last and the model was refined and checked. The large residual density within the omit ($mF_o - DF_c$) map contoured to

$\pm 3\sigma$ was assigned as NADPH missing the nicotinamide ring and (*R*)-MTQ. This crystal structure is a ternary complex of AoIRED, NADPH and (*R*)-MTQ.

Upon comparing the length of the dimer with all previously obtained crystal structures, between residue 57 of subunit A and residue 57 of subunit B (**Figure 89**), the ternary complex and the NADPH complex share a similar length of approximately 90 Å, whereas, *apo*-AoIRED and the (*R*)-MTQ complex share a similar distance of approximately 92 Å. This is further highlighted by the superimposition of structures. The NADPH-complex superposes onto the ternary complex with an RMSD of 0.3 Å over 564 residues, whereas the *apo*- and xenobiotic ligand complexes superpose with an RMSD of 1.35 Å and 1.41 Å over 540 and 535 residues respectively. However, these differences could be attributed to different crystal packing.

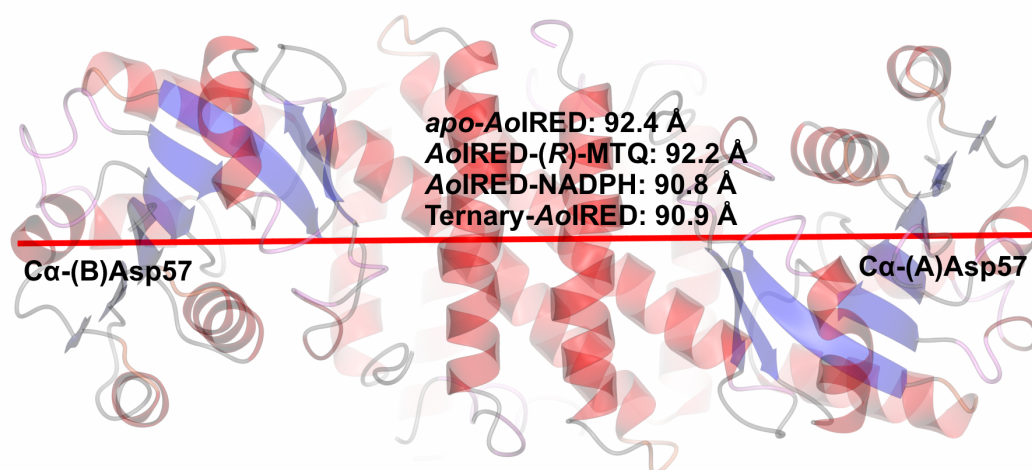


Figure 89 – Comparison of dimer length between previously obtained crystal structures of AoIRED. The NADPH-complex and the ternary complex are noticeably more closed compared to the *apo*- and binary xenobiotic ligand complexes.

Therefore, the ternary complex with partial NADPH density and (*R*)-MTQ, was termed AoIRED-‘closed’-(*R*)-MTQ complex, due to the partial NADPH density and AoIRED has adopted a significantly more closed conformation. The previous binary xenobiotic ligand complexes has been termed AoIRED-‘open’-ligand complexes, as it was obtained using *apo*-AoIRED, which is seen to be in a more open conformation.

The crystal structure was refined to R_{factor} and R_{free} values of 17.1% and 22.4% respectively. Besides the chain breaks between residues 228 to 230 and 277

to 289, the model is complete from residues 3 to 290 (**Figure 90**). DALI analysis resulted in the same results as the analysis with the *Ao*IRED-NADPH complex. PISA analysis indicates a dimerising interface of 4137 Å², consisting of 16 hydrogen bonds and 7 salt bridges that spontaneously dimerises with a ΔG value of -87.4 kcal mol⁻¹. The area covered in the dimerising interface of the ternary complex is more similar to that of the NADPH-complex compared to *apo-Ao*IRED.

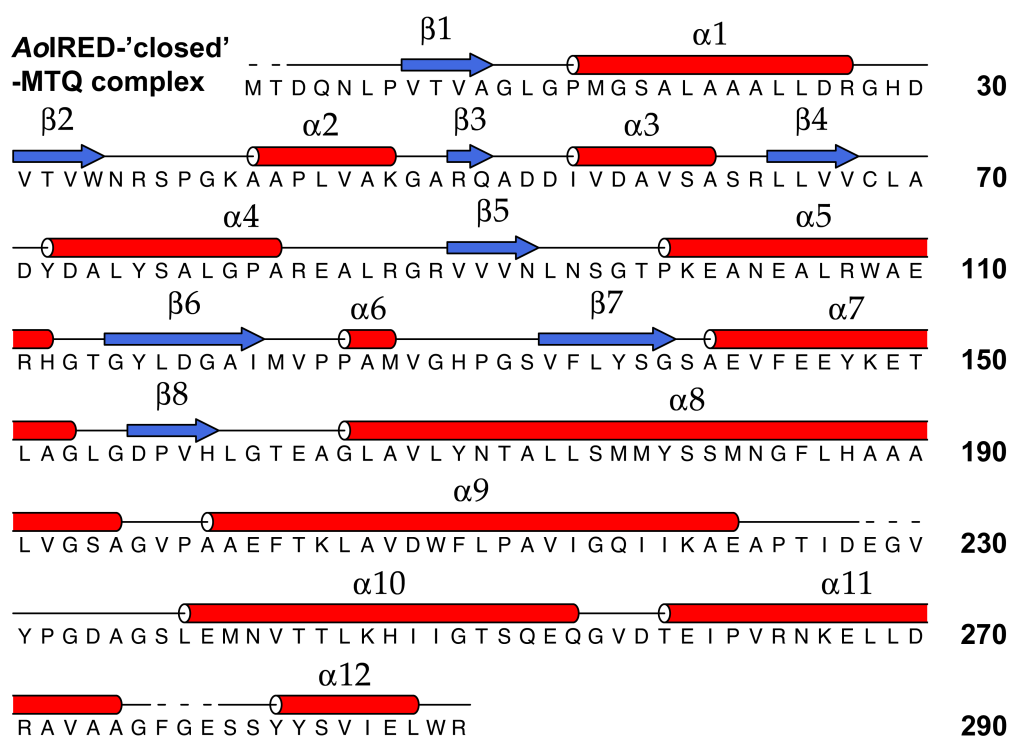


Figure 90 – The primary structure of *Ao*IRED-'closed'-(R)-MTQ complex with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.

The active site of the ternary *Ao*IRED-'closed'-(R)-MTQ complex (**Figure 91**) consists of similar residues as the previous *Ao*IRED-'open'-(R)-MTQ complex, however due to the structures more compact conformation, more residues are now involved. The active site contains residues Asn171, Leu175, Met178 Tyr179, Val214, Ile218, Met240, Asn241 and Thr244. Interestingly in this ternary complex, there is no longer a network of hydrogen bonding to a water molecule, however there is now end-on π -interaction between Met178 and the aromatic ring of Tyr177.

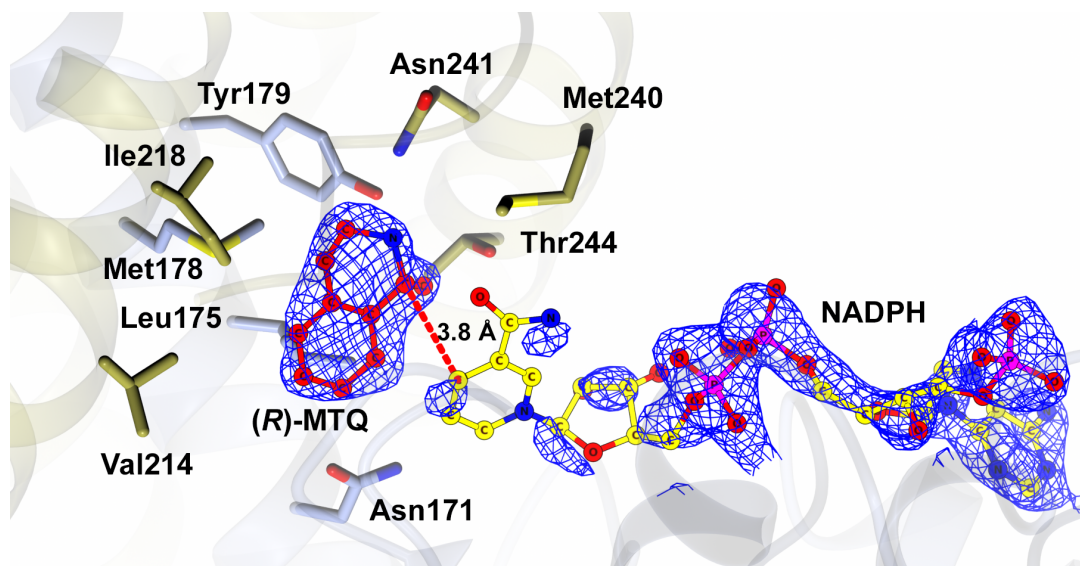


Figure 91 – The AoIRED-‘closed’-(*R*)-MTQ complex. Subunits A and B are in blue and gold respectively. NADPH and (*R*)-MTQ is in yellow and red circles respectively for illustrative purposes only. The electron density for the ligands is the omit ($mF_o - DF_c$) map contoured to 3σ obtained from refinement prior to modelling and refining NADPH and (*R*)-MTQ.

The crystal data statistics for all complexes of AoIRED are summarised in **Table 31**.

Table 31 – Crystal data collection and refinement statistics for all AoIRED data collection. Numbers in brackets refer to the data in the highest resolution shell.

	AoIRED 'open' apo-	AoIRED 'closed' NADPH complex	AoIRED 'open' (R)-MTQ complex	AoIRED 'open' complex with (S)-4 amine	AoIRED 'open' complex with (R)-5 amine	AoIRED 'open' complex with (R)-12 amine	AoIRED 'closed' (R)-MTQ complex
Date Obtained	2nd Oct 2014	25th May 2014	19th Oct 2014	9th Feb 2015	9th Feb 2015	9th Feb 2015	1st Feb 2016
Beamline	Diamond i03	Diamond i04	Diamond i03	Diamond i04	Diamond i04	Diamond i04	Diamond i02
Wavelength (Å)	0.97950	0.97935	0.97626	0.97949	0.97949	0.97949	0.97949
Resolution (Å)	57.75-1.55 (1.59-1.55)	60.32-2.06 (2.11-2.06)	60.10-1.46 (1.50-1.46)	59.87-1.50 (1.53-1.50)	60.00-1.40 (1.42-1.40)	60.63-1.73 (1.76-1.73)	73.78-2.14 (2.20-2.14)
Space Group	C2 2 2 ₁	P2 ₁ 2 ₁ 2 ₁	C2 2 2 ₁	C2 2 2 ₁	C2 2 2 ₁	C2 2 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit Cell (Å)	a = 75.58; b = 89.62; c = 91.48; α = γ = β = 90.00	a = 52.79; b = 66.06; c = 147.88; α = γ = β = 90.00	a = 80.00; b = 91.06; c = 90.98; α = γ = β = 90.00	a = 79.83; b = 90.52; c = 90.98; α = γ = β = 90.00	a = 79.89; b = 90.87; c = 90.87; α = γ = β = 90.00	a = 83.22; b = 88.51; c = 88.89; α = γ = β = 90.00	a = 52.85; b = 65.67; c = 147.56; α = γ = β = 90.00
No. of molecules in the asymmetric unit	1	2	1	1	1	1	2
Unique reflections	297497 (44373)	211842 (32832)	320639 (52377)	335203 (52130)	414714 (64912)	230895 (34608)	230211 (29186)
Completeness (%)	99.0 (99.6)	99.9 (99.9)	98.7 (53.4)	99.5 (88.2)	99.9 (92.6)	100.0 (100.0)	99.8 (100.0)
R _{merge} (%)	3.5 (69.7)	2.1 (69.9)	3.1 (45.0)	2.1 (56.5)	2.3 (56.9)	4.0 (97.0)	0.05 (0.71)
R _{p, int.} (%)	2.2 (43.3)	1.3 (44.7)	2.0 (34.7)	1.3 (38.4)	1.5 (32.7)	2.6 (60.2)	0.02 (0.28)
Multiplicity	6.0 (6.8)	5.2 (6.4)	5.8 (3.9)	5.8 (5.0)	5.8 (4.4)	5.6 (6.8)	6.4 (7.8)
<I/σ(I)>	50.5 (3.1)	54.7 (2.9)	55.0 (2.5)	64.5 (2.6)	62.5 (2.1)	44.0 (2.0)	28.3 (4.0)
CC _{1/2}	1.00 (88.0)	1.00 (80.0)	99.8 (77.5)	99.7 (83.8)	99.5 (74.6)	98.7 (72.4)	0.99 (0.89)
Overall B factor from Wilson plot (Å ²)	25	36	23	20	17	30	21
R _{factor} /R _{free} (%)	18.3/21.9	16.4/22.7	17.8/20.3	22.1/25.7	17.3/18.7	19.7/23.7	17.1/22.4
r.m.s.d 1-2 bonds (Å)	0.022	0.015	0.025	0.020	0.025	0.019	0.015
r.m.s.d 1-3 angles (°)	2.08	1.78	2.00	2.00	2.56	2.06	1.83
Avg main chain B (Å ²)	23	33	20	18	20	28	27
Avg side chain B (Å ²)	28	42	24	22	26	32	30
Avg water B (Å ²)	30	39	29	29	29	31	33
Avg NADPH B (Å ²)	-	91	-	-	-	-	105
Avg Ligand B (Å ²)	-	-	35	32	57	49	51
Normalised B factor	0.92	0.92	0.87	0.90	1.18	0.93	1.29

4.2.1. Discussion

*Ao*IRED is an interesting IRED as it has Asn171 in place of Tyr169 in *SS*IRED and Asp187 in *Sk*IRED, as well as being a highly selective IRED with a large substrate scope and unpredictable selectivity (**Figure 74**). Due to this, it becomes inappropriate to assign *Ao*IRED as an (*S*)- or (*R*)-selective IRED. Furthermore, *Ao*IRED does not exhibit strict *re*- or *si*- preference towards the imine like *SS*IRED or *SR*IRED.

*Ao*IRED is also considerably more active compared to *SS*IRED, *Bc*IRED and *Nh*IRED considering imine **1** (2MPN) is not the best substrate for *Ao*IRED.

Table 32 – Comparison of kinetic parameters for IREDs that are (*S*)-selective for the substrate **1** (2MPN). [a] kinetic constants taken from Leipold and co-workers (2013).²⁰ [b] kinetic constants provided by Aleku and co-workers (personal communication).

Enzyme	Kinetic Parameters for imine 1 (2MPN)		
	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (s ⁻¹ mM ⁻¹)
<i>SS</i> IRED ^[a]	11.82	0.024	0.002
<i>Bc</i> IRED	8.32	0.032	0.003
<i>Nh</i> IRED	1.10	0.013	0.011
<i>Ao</i> IRED ^[b]	0.87	0.15	0.17

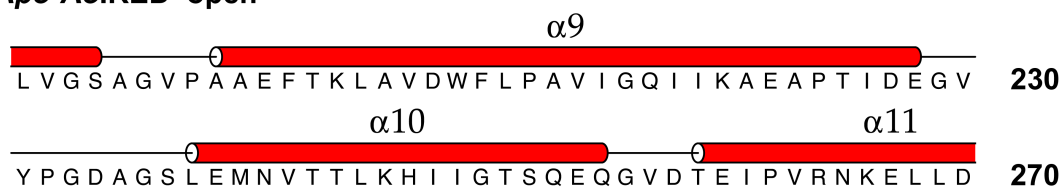
When comparing *Ao*IRED against *SR*IRED (**Table 33**), it appears that *SR*IRED has a preference for smaller substrates and is more active than *Ao*IRED. However, *Ao*IRED is still a useful IRED for the reduction of bulky imines, such as imines **13**, **15**, **16a**, **16b** and **16c**, which *SR*IRED is unable to accept (Hussain and co-workers, personal communication).

Table 33 – Comparison of kinetic parameters between *SR*IRED and *Ao*IRED for imine substrates **1**, **6** and **14**. [a] kinetics constants taken from Hussain and co-workers (2015).⁴³ [b] kinetic constants provided by Aleku and co-workers.

Substrate	Kinetic parameters for substrates 1 , 6 and 14 . [K_M (mM); k_{cat} (s ⁻¹); k_{cat}/K_M (s ⁻¹ mM ⁻¹)]	
	<i>SR</i> IRED ^[a]	<i>Ao</i> IRED ^[c]
1	1.88; 0.35; 0.698	0.87; 0.15; 0.17
6	0.16; 0.19; 0.22	0.72; 0.74; 1.03
14	0.32; 0.14; 0.45	0.36; 0.94; 2.58

AoIRED is the only IRED studied thus far with such distinct ‘open’ and ‘closed’ forms. 6 residues at the terminus of $\alpha 9$ (Figure 92) become unstructured to accommodate additional ligands. Tighter ligand complexes have been observed with previously investigated IREDs, however none of the IREDs undergo such drastic conformational changes.

Apo-AoIRED ‘open’



AoIRED-NADPH ‘closed’

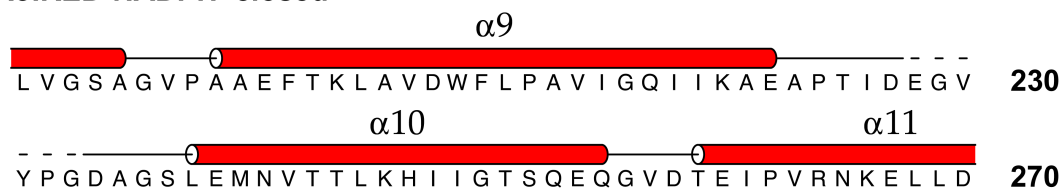


Figure 92 – Comparison between the dssp of ‘open’ and ‘closed’ forms of AoIRED between residues 191 to 279.

The ternary (partial NADPH with (*R*)-MTQ) ‘closed’ complex illustrates a tighter complex where the binding site for (*R*)-MTQ is the same as the site obtained in the ‘open’ binary complex of (*R*)-MTQ. As such, the ‘open’ binary complexes of the amine product of **4**, **5** and **12** may be crystallographic artefacts (Figure 93). Whereas, the ‘open’ binary complex with (*R*)-MTQ and the partial ternary complex have shed some light on the mechanism for IREDs.

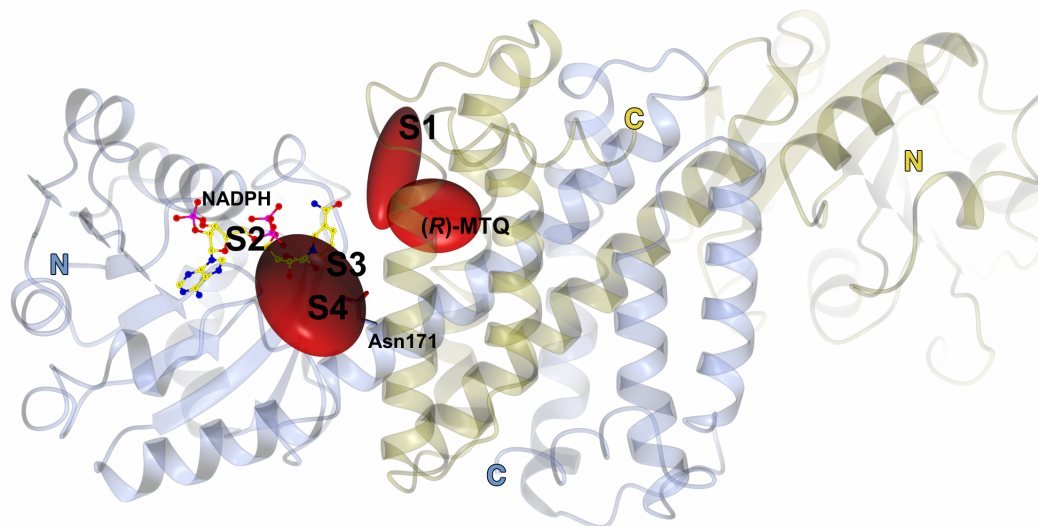


Figure 93 – The dimeric ‘open’ model of AoIRED with the binding sites for ligands 4, 5, 12, (R)-MTQ and NADPH. Subunits A and B are in blue and gold respectively. NADPH is shown in yellow circles for illustrative purposes only. Ligand binding sites for ligands 4, 5, 12 and (R)-MTQ are shown in red blobs for illustrative purposes only. Sites S1 is where ligands 4 and 12 were bound. Site S2 is where NADPH was bound. Sites S3 and S4 are where ligand 5 were bound. Site (R)-MTQ is where (R)-MTQ was bound.

Various residues in the binding site for **6**, as in **Figure 91**, were subjected to SDM experiments to probe the role of the residues within the active site. The mutation data, obtained by Aleku and co-workers (personal communication), is summarised in **Table 34**.

Table 34 – Biotransformations for variants of AoIRED. Substrate numbers refer to those in **Figure 74**. Data is presented as: conversion (%), *e.e* (%) and absolute configuration.

	Imine	Variants of AoIRED					
		WT	Y179A	Y179F	N241A	N171A	N171D
1a		96; 82 (S)-	23; 99 (S)-	35; 47 (S)-	43; 81 (S)-	31; >99 (S)-	2; >99 (S)-
1b		100; 95 (S)-	8; 99 (S)-	99; 92 (S)-	43; 81 (S)-	100; 4 (S)-	100; 95 (S)-
2b		99; 40 (S)-	32; 43 (R)-	93; 8 (R)-	99; 0	41; 68 (S)-	80; 80 (S)-
2d		70; 90 (R)-	3; 59 (R)-	25; 84 (R)-	56; 72 (R)-	2; 99 (R)-	11; 80 (R)-
15		50; 79 (S)-	15; >99 (R)-	5; 99 (R)-	99; 85 (R)-	2; 31 (S)-	2; 99 (S)-
16c		15; 62 (S)-	10; 99 (S)-	9; 99 (S)-	96; 60 (R)-	18; 71 (S)-	6; 99 (S)-

All variants of AoIRED were still active. For variants Y179A and Y179F, overall yield was compromised, but stereoselectivity was retained compared to WT-

AoIRED for smaller substrates such as **1a** and **1b**, however stereoselectivity was inverted with larger substrates such as **2b** and **15**. Variant N241A exhibited compromised yields for smaller substrates, however yield and enantioselectivity for **15** and **16c** were notably higher despite an inversion of stereoselectivity. Variants N171A and N171D retained the same enantiopreference as WT-AoIRED, however overall yields were significantly lower in majority of the cases.

The kinetic data for variants of AoIRED was obtained for many substrates, obtained by Aleku and co-workers (personal communication), however only the data for **1a** (2MPN) is summarised in **Table 35**.

Table 35 – Biotransformation and kinetic data for substrate 1a (2MPN) for variants of AoIRED.

AoIRED variant	<i>Biotransformation and kinetic data for substrate 1a</i>				
	Conversion (%)	<i>e.e</i> (%)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1} mM^{-1}$)
WT	96	82 (S)-	0.83	0.15	0.17
Y179A	23	99 (S)-	1.77	0.03	0.05
Y179F	35	47 (S)-	0.87	0.01	0.02
N241A	43	81 (S)-	1.27	0.22	0.17
N171A	31	>99 (S)-	7.90	0.12	0.02
N171D	2	>99 (S)-	2.22	0.17	0.01

It is clear that all variants of AoIRED, except from N241A, perform worse than WT-AoIRED. In majority of the cases, yields, substrate affinity, turnover number and catalytic efficiency is significantly lowered. However with N241A, turnover numbers have actually increased and despite a decrease in substrate affinity the overall catalytic efficiency is retained compared to WT-AoIRED. The data suggests that residues Tyr179 and Asn171 play a role in catalysis, whereas Asn241 plays a lesser role. Therefore variants at Asn241 would not severely impact activity but could be used to expand substrate specificity.

The crystal structure data along with the mutation and kinetic data have revealed information on protein dynamics which could be attributed to the inversion of selectivity towards substrate **6** as well as reveal insights to the mechanism of AoIRED. Despite some clarification, to fully understand the mechanism, further investigations are required.

4.3. Structure and mechanism of the reductive aminase (RedAm) from *Aspergillus oryzae*

Reductive amination is the condensation between a carbonyl and amine to form an amine *via* an imine intermediate. Towards the end of 2015, considerable effort was put into performing reductive amination with IREDs (Figure 94).^{101, 153} As a pre-formed imine was no longer required, substrate constraints are lifted as reductive amination is a synthetically versatile method for obtaining chiral amines. The majority of imine substrates are restricted to being aryl or cyclic, otherwise the imine is destabilised by lack of aryl groups or further conjugation and tends to be labile in solution. With reductive amination, any desired ketone and amine building blocks can be used to form the imine intermediate, which is subsequently reduced.

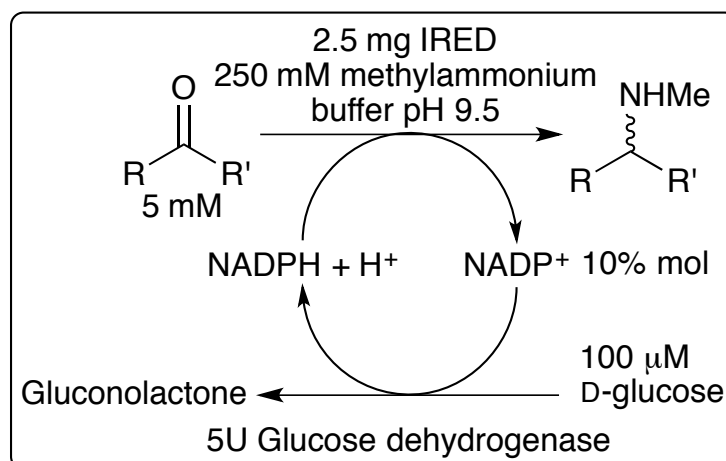


Figure 94 – Reductive amination performed with large quantities of IRED requiring up to 50 times excess amine donor source to drive the equilibrium towards imine formation.

IREDs that are capable of reductive amination are therefore highly desirable biocatalysts. However, the currently investigated IREDs used for reductive amination experiments are not thought to catalyse the formation of the imine. Typically a large excess of amine donor source, of up to 50 times and is often the buffer source itself, to drive the equilibrium towards imine formation. In these cases, large quantities of enzyme are required to achieve very modest yields.

The reductive aminase (RedAm) from *Aspergillus oryzae* (*AspRedAm*) was initially characterised as an IRED with good selectivity and substrate specificity by Aleku and co-workers (personal communication). With the increase in popularity with performing reductive amination with IREDs, *AspRedAm* was screened for RedAm activity and was found to produce significantly larger yields of amine product. A biotransformation of interest (**Figure 95**) was the reductive amination between 1-indanone, **9**, and propargylamine, **c**, with *AspRedAm* to yield the human monoamine oxidase B inhibitor, (*R*)-rasagiline. Not only was the yield a significant improvement over previously investigated IRED-based reductive amination, but *AspRedAm* still retained excellent enantioselectivity.

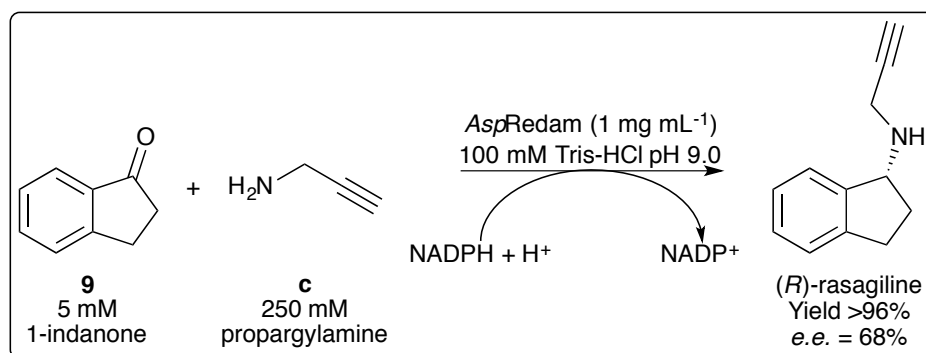


Figure 95 – Scheme for the production of (*R*)-rasagiline with *AspRedAm* from starting materials 1-indanone, **9**, and propargylamine, **c**. Performed by Aleku and co-workers (personal communication).

Further biotransformation and kinetic assay investigations into the reductive amination capabilities of *AspRedAm* revealed that it could perform the reaction with certain substrates at a carbonyl-to-amine ratio of 1:1 and catalysed the imine formation. Whereas, with previous reductive amination investigations where higher enzyme and amine substrate loadings were used, the imine formation was chemically driven. Therefore, IREDs capable of catalysing reductive amination at a carbonyl-to-amine ratio of 1:1 are designated as ‘RedAms’. Hence, obtaining crystal structures and ligand complexes of *AspRedAm* would reveal active site differences between IREDs and RedAms.

The synthetic gene for *AspRedAm* was codon optimised for expression in *E. coli* and synthesised by Biomatik (Cambridge, Ontario, Canada). The gene was sub-cloned into pET-28a (+) vector and supplied by Aleku and co-workers (personal communication).

It was observed that *AspRedAm* shares a lot of sequence homology with *AoIRED* (**Figure 96**). However from sequence alignment alone, it is not clear why *AspRedAm* is capable of catalysing reductive amination.

```

AoIRED      MTDQNLPVTVAGLGPMGSALAAALLDRGHDVTVWNRSPGKAAPLVAKGARQADDIVDAVS
AspRedAm    --MSKHIGIFGLGAMGTALAAKYLEHGKYKTSVWNRRTAKAIPLVEQGAKLASTISEGVN
           .  :  :  **  **:****  *:::..:****:  .**  ***  **:  *  *  :..*

AoIRED      ASRLLVVCLADYDALYSALGPAREALRGRVVVNLNSGTPKEANEALRWAERHGTGYLDGA
AspRedAm    ANDLIIICLLNNQVEDALRDALQTLPSKTI VNLTNNGTPNQARKLADFVTSHGARYIHGG
           *.  *:::**  :  ::  .**  *  ::*  ..:****..****:*.  :  .  **:  *:.

AoIRED      IMVPPAMVGHPGSVFLYSGSA-EVFEEYKETLAGLGDPVHLGTEAGLAVLYN TALLSMY
AspRedAm    IMAVPTMIGSPHAVLLYSGESLELFQSIESHLSLLGMSKYLGTDAGSASLH LALLSGMY
           **.  *:::*  *  :*:****.:  *:*  .  .  *  :  **  :****:*  *  *::  ****  **

AoIRED      SSMNGFLHAAALVGSAGVPAA---EFTKLAVDWFLPAVIGQIIKAEAPTIDEGVYPGDAG
AspRedAm    GLFSGFLHAAVALIKSGQDTSTTATGLLPLLPWLSA-M-TGYLSSIAKQIDDGDYATQGS
           .  .  .****.  **.  *  .  :  :  :  *  .  *  :  :  :  *  **:*  *  ..

AoIRED      SLEMNVTTLKHHIIGTSQEQVDTEIPVRNKELLDRAVAAGFGESSYYSVIELWR-----
AspRedAm    NLGMQLAGVENIIRAGEEQRVSSQMLLPKALIEQAVGEGHGGEDLSALIEYFKVGNVD
           .  *  *:::  ::.*  :::**  *:::  :  *  *:::**.  *  *  ..  :***  :

```

Figure 96 – Sequence alignment of *AoIRED* and *AspRedAm*, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹ *AspRedAm* shares 64.2% sequence similarity with *AoIRED*.

Gene expression protocols were optimised by Aleku and co-workers (personal communication). Once expression was scaled up, *AspRedAm* was purified (**3.3**) by nickel-affinity and SEC (**Figure 97**). *AspRedAm* noticeably precipitates at high imidazole concentrations and during concentration for SEC.

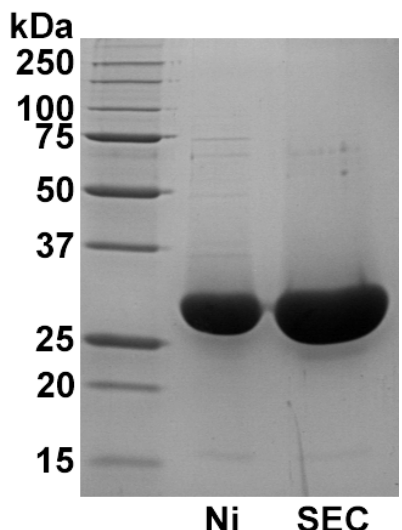


Figure 97 – SDS-PAGE analysis from purification of *AspRedAm* by nickel-affinity chromatography and SEC. Ni is the sample obtained from purification by nickel-affinity. SEC is the sample obtained from purification by SEC. The expected molecular weight of *AspRedAm* is 33 kDa, including the 6His-tag and thrombin cleavage site.

The chromatogram obtained from the SEC (**Figure 98**) displayed a similar elution volume as the previously purified IREDs, suggesting a dimeric state in solution, therefore a dimeric weight of ~60 kDa. However the SDS-PAGE analysis suggests a monomeric weight of ~30 kDa which is in agreement with previously purified IREDs. *AspRedAm* also precipitates when the SEC purified samples are left at 5°C overnight. Therefore, a lot of yield is lost to the protein precipitating during and after purification.

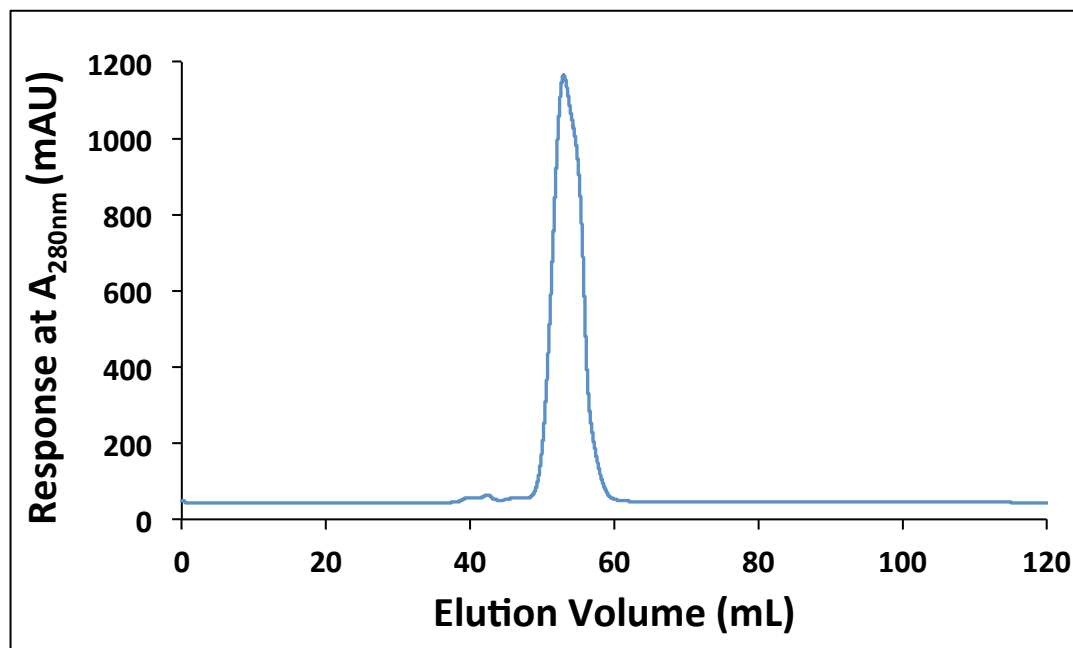



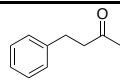
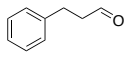
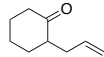
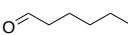
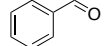

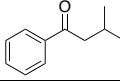
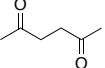
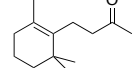
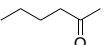
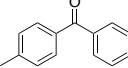

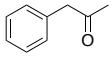
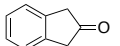

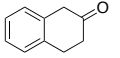
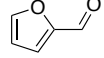
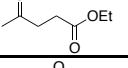


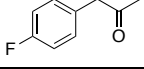
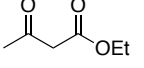
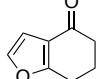

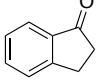
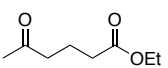
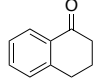
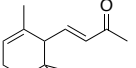
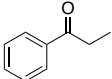
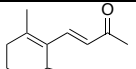
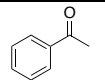


Figure 98 – Chromatogram of the purification of *AspRedAm* by SEC. Main peak elutes between 50-55 mL, correlates to a species of an approximate molecular weight of 60 kDa.

All reductive amination biotransformations with *AspRedAm* were performed by Aleku and co-workers (personal communication). Relative specific activity was obtained for a range of carbonyl acceptors and either propargylamine or methylamine as the amine donor at a 1:4 ketone-to-amine ratio (**Table 37**). The best carbonyl acceptors appear to be compounds with substituents that exert minimal positive and negative mesomeric effects, and instead with substituents that exert positive inductive electronic effects, as seen with entries 1-4 (**Table 37**).


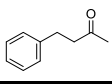

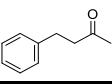
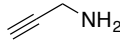
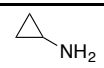

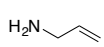
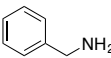
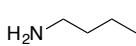

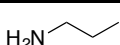
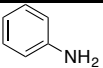




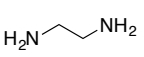
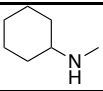
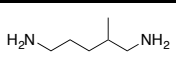
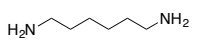
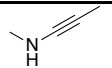
Table 36 – Relative activity of carbonyl acceptors in AspRedAm catalysed reductive aminations. Specific activities (i) cyclohexanone/propargylamine = 6.68 U/mg (100%); (ii) cyclohexanone/methylamine = 2.23 U/mg (100%). Reaction conditions: 15 mM carbonyl acceptor, 60 mM amine nucleophile (1:4 ketone-to-amine ratio), 0.3 mM NADPH and 100 mM borate buffer pH 9.0.

Entry	Carbonyl acceptor	Relative specific activity [%]		Entry	Carbonyl acceptor	Relative specific activity [%]	
			CH ₃ NH ₂				CH ₃ NH ₂
1		100	100	17		10	8
2		75	85	18		10	6
3		7	93	19		14	12
4		44	40	20		11	13
5		33	21	21		9	8
6		30	26	22		7	8
7		30	25	23		4	4
8		35	33	24		4	3
9		33	28	25		5	5
10		8	75	26		4	3
11		28	26	27		4	1
12		7	60	28		4	4
13		24	20	29		3	1
14		7	45	30		2	1
15		19	18	31		4	3
16		16	15	32		2	1

Once a good and poor carbonyl acceptor had been established based on relative specific activity from **Table 37**, the relative specific activities for a wide range of amines were obtained (**Table 38**) with cyclohexanone and 4-phenylbutanone as the carbonyl acceptors. A good amine donor source appears to

be a primary amine with substituents that exert positive inductive electronic effects, as seen with entries 1-5 in **Table 37**, which is similar to the carbonyl preference.

Table 37 – Relative activity of amines in *AspRedAm* catalysed reductive amination. Specific activities: (i) cyclohexanone/propargylamine = 6.68 U/mg (100%); (ii) 4-phenyl-2-butanone/propargylamine = 0.54 U/mg (100%). Reaction conditions 15mM carbonyl acceptor: 60mM amine nucleophile (1:4 ketone-to-amine ratio), 0.3mM NADPH and 100 mM borate buffer pH 9.

Entry	Amines	Relative Specific Activity [%]		Entry	Amines	Relative Specific Activity [%]	
							
1		100	100	11	NH ₃	16	0
2		125	40	12		10	0
3		115	65	13		13	0
4		59	40	14		8	3
5		65	33	15		12	0
6		62	18	16		5	0
7		55	12	17		6	2
8		14	5	18		3	1
9		7	4	19		2	1
10		10	0				

Structural data was required to identify differences between RedAms, such as *AspRedAm*, and previously investigated IREDs. Therefore, purified native *AspRedAm* was subjected to crystallisation trials with commercially available screens (**3.5.1**) at a protein concentration of 20 mg mL⁻¹ in buffer B (**3.3**). After several rounds of optimisation in 48-well sitting drop format (**3.5.3**) utilising a 2 μ L crystallisation drop with a 150 μ L mother liquor, the crystals in were obtained in conditions containing 50 mM Tris pH 8.5, 23% (w/v) PEG 3350, 5% (v/v) butan-1,4-diol and 6% (w/v) xylose. However despite these optimisations and further

optimisations with cryoprotectants, these crystals did not result in any reflections with diffraction analysis.

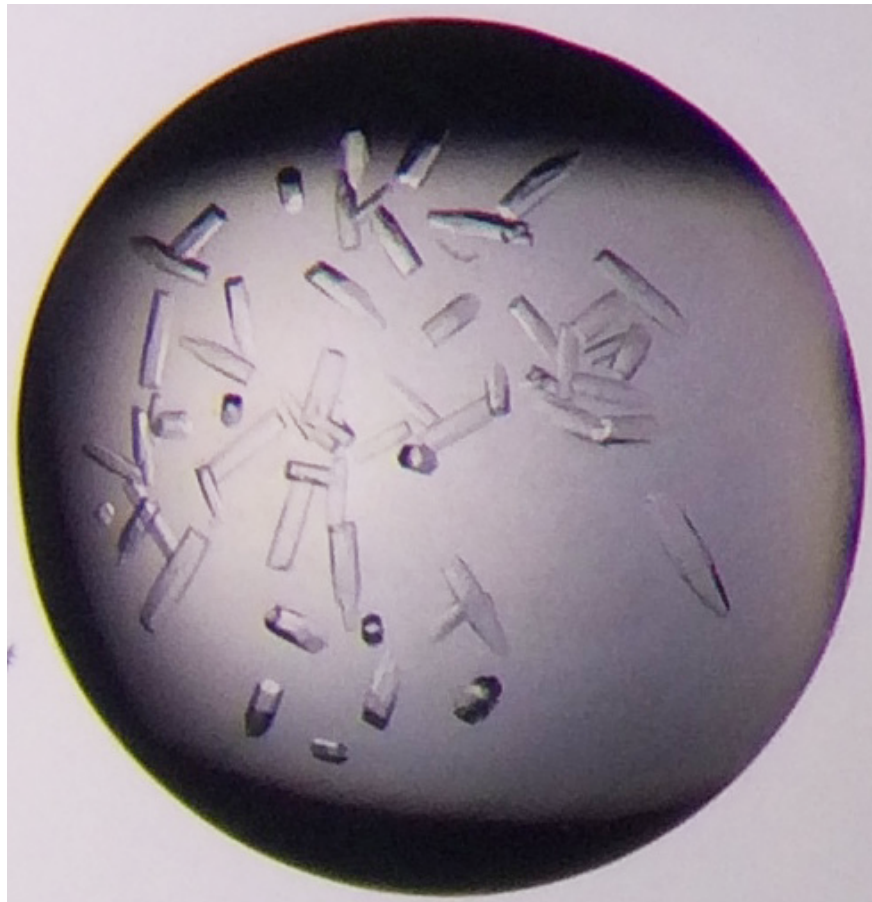


Figure 99 – Optimised crystals of *AspRedAm*, using a 2 μL crystallisation drop in sitting drop format, obtained in conditions containing 50 mM Tris pH 8.5, 23% (w/v) PEG 3350, 5% (v/v) butan-1,4-diol and 6% (w/v) xylose.

AspRedAm was screened again with commercially available crystallisation screens utilising a 300 nL crystallisation drop and 54 μL mother liquor with a protein concentration of 17 mg mL^{-1} in buffer B (3.3) with the addition of 1 mM NADPH and 5 mM methylamine added into the mother liquor as an additive (3.5.2.3). New crystal hits appeared within 48 h at 18°C in different conditions to the hits obtained with native *AspRedAm* screening. The crystals of the most regular morphology were obtained in CSS II, in conditions containing 15% (w/v) PEG 4K, 0.2 M calcium acetate, 0.1 M MES pH 6.0, 1 mM NADPH and 5 mM methylamine (Figure 100). These crystals diffracted beyond 3 Å in-house and therefore did not need further optimisation. The crystals were fished and flash cooled in liquid nitrogen without further cryoprotectants.

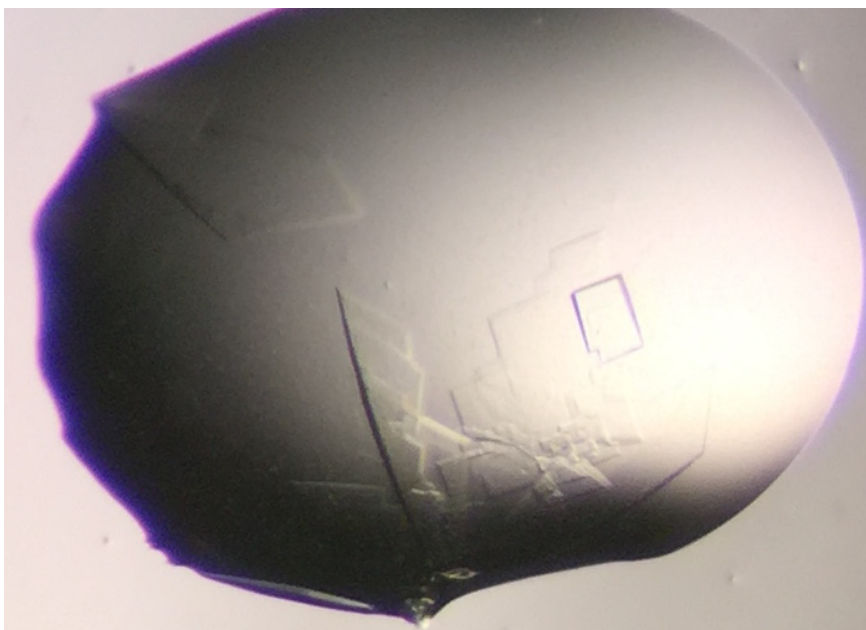


Figure 100 – Crystals of *AspRedAm*, at a protein concentration of 17 mg mL^{-1} in buffer B (3.3), obtained a 300 nL crystallisation drop in CSS II + 0.1 M MES pH 6.0 + 0.5 mM NADPH A6 from crystal trial screening.

A full dataset was collected for *AspRedAm*-NADPH complex to 1.8 \AA resolution using 10% beam transmission (3.6.1) and helical data collection (3.6.2). The Xia2 auto processed data was used for structure solution and refinement. The crystal structure of *AspRedAm* was solved in the space group $P2_1$ with Molrep using *apo-AoIRED* as the molecular replacement model. Iterative cycles of model building and refinement were performed. Water molecules were added last with the model refined and checked. Finally the residual density between the dimer interface at the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was assigned as NADPH.

The crystal structure is composed of two monomers in the asymmetric unit arranged as a single homodimer as with other IREDs (Figure 101).

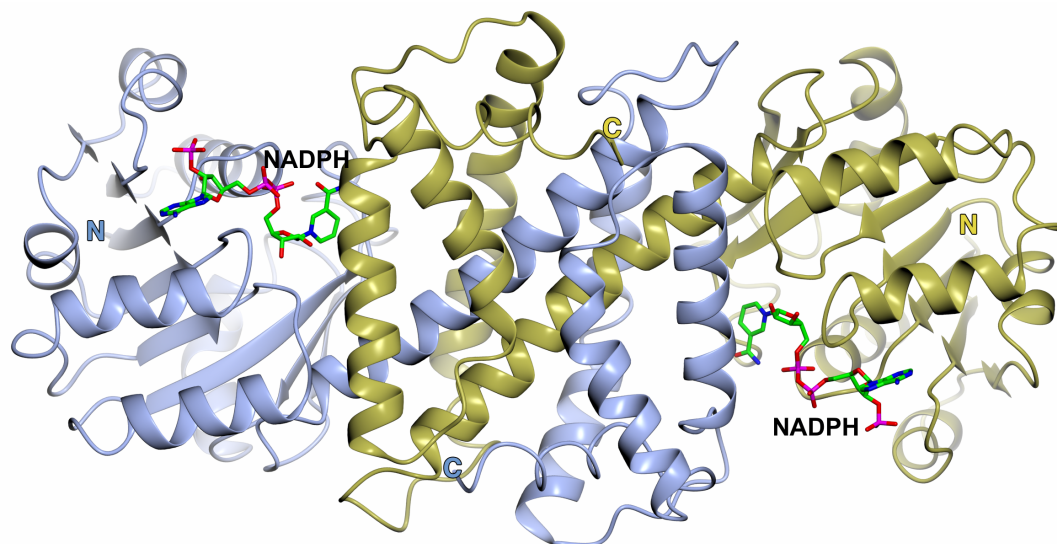


Figure 101 – The crystal structure of *AspRedAm*-NADPH complex. Subunits A and B are in blue and gold respectively.

The final structure was refined to R_{factor} and R_{free} values of 18.5% and 22.5% respectively. The model was complete from residues 3 to 291 in both subunits (**Figure 102**). The structure adopts the canonical IRED fold, similar to previously obtained IRED structures. DALI analysis indicates that the *AspRedAm*-NADPH complex is most structurally similar to the IRED from *Streptomyces aurantiacus*¹⁰¹ (PDB entry: 4OQZ), which has a sequence homology of 28% with monomers superposing with an RMSD of 1.4 Å over 282 out of 283 residues. PISA analysis indicates the dimer interface binds with a ΔG value of $-78.8 \text{ kcal mol}^{-1}$ and has an interfacial area of 3908 \AA^2 that consist of 23 hydrogen bonds and two salt bridges.

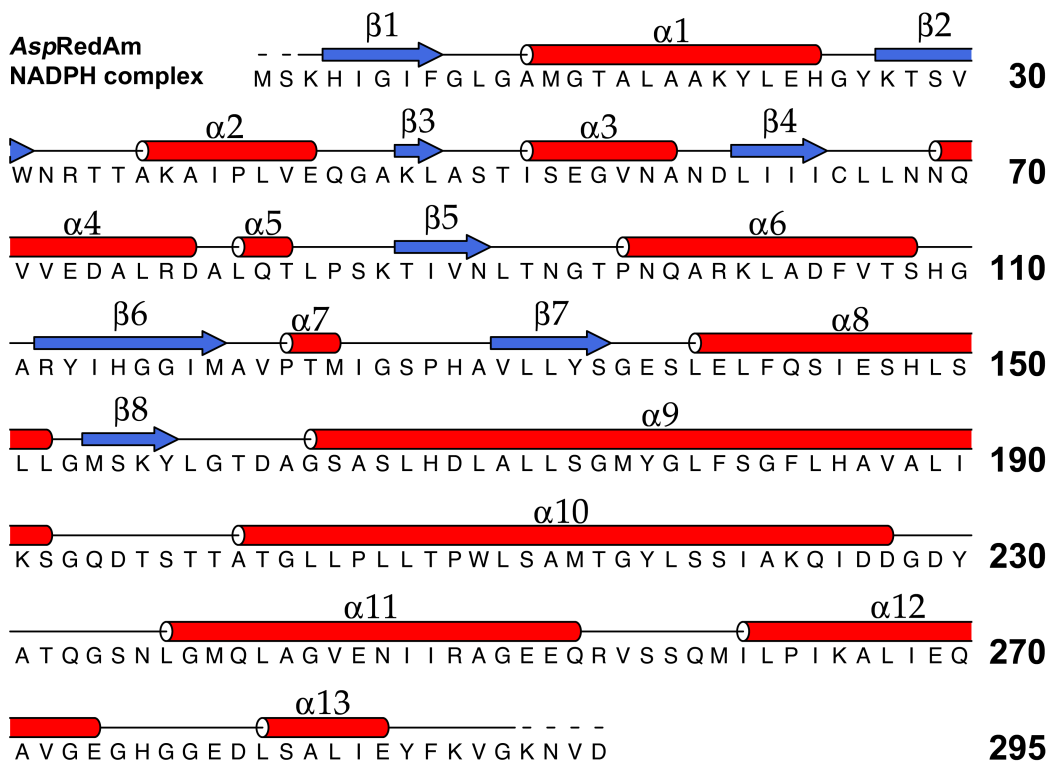


Figure 102 – The primary structure of AspRedAm-NADPH complex with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.

The active site (**Figure 103**) reveals that AspRedAm has an Asp169 at the same position of Asn171-AoIRED. The binding of NADPH is mostly hydrophobic consisting of N-terminal domain residues Met13, Arg33, Thr34, Cys65, Leu67, Asn93 and C-terminal domain residue Asp169. Arg33 and Met13 forms π -stacking interactions, with distances of approximately 3.5 Å, with the adenine ring and the nicotinamide ring of NADPH respectively. The distance between the hydroxyl of Asp169 to the C4N-NADPH has been observed to be 7.6 Å. Interestingly there is a water molecule hydrogen bonding, with a distance of 2.9 Å, to Asp169, similar to that as seen in AoIRED. However, this NADPH complex does not reveal any insight to differences between an IRED and a RedAm. Therefore, a ternary complex was pursued in an attempt to identify structural differences.

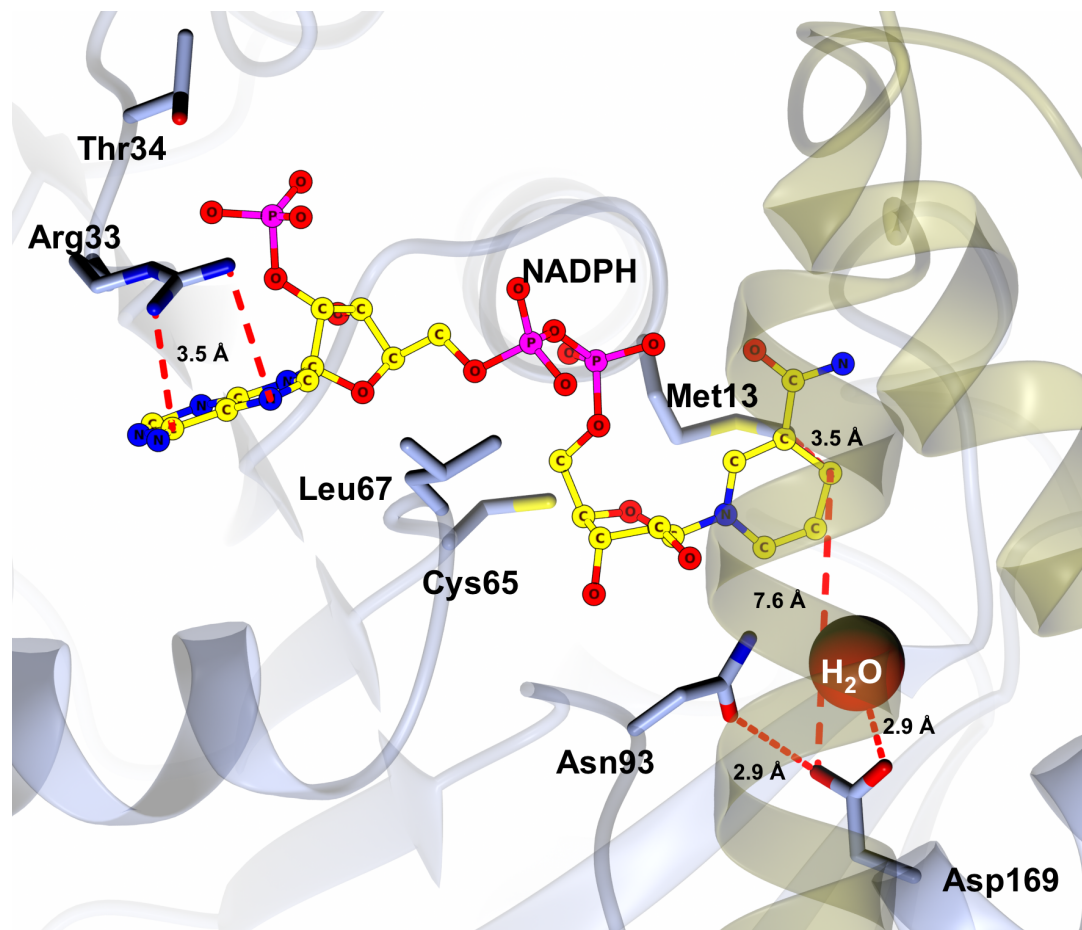


Figure 103 – The binding site of NADPH in *AspRedAm*. Subunits A and B are in blue and gold respectively. NADPH is in yellow circles for illustrative purposes only. Key interactions include Arg33 contributing π -stacking with the adenine ring of NADPH and Met 13 contributing π -stacking to the nicotinamide ring of NADPH.

One of the biotransformations of interest with *AspRedAm* was the production of the human monoamine oxidase B inhibitor (*R*)-rasagiline from cheap starting materials, 1-indanone and propargylamine (Figure 104). Substrates from the biotransformation were chosen for crystallographic studies.

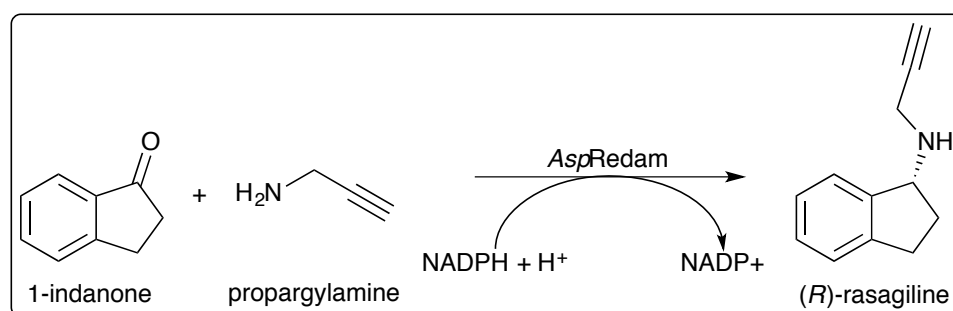


Figure 104 – Scheme of the reaction of interest: the production of (*R*)-rasagiline from 1-indanone and propargylamine with *AspRedAm*.

The racemic product, *rac*-rasagiline, was synthesised by Anibal Cuetos-Fernandez (personal communication, in our group). In an attempt to obtain a ternary complex, *AspRedAm* was rescreened with commercially available screens utilising a 300 nL crystallisation drop and a 54 μ L mother liquor with 0.5 mM NADPH and 0.5 mM *rac*-rasagiline, so that the ligands would not turn over in solution with the active enzyme. Crystals appeared (**Figure 105**) overnight at 18°C in conditions containing 0.2 M lithium sulphate, 0.1 M BisTris pH 6.5, 25% (w/v) PEG 3350, 0.5 mM NADPH and 0.5 mM *rac*-rasagiline at a protein concentration of 17 mg mL⁻¹ in buffer B (**3.3**).

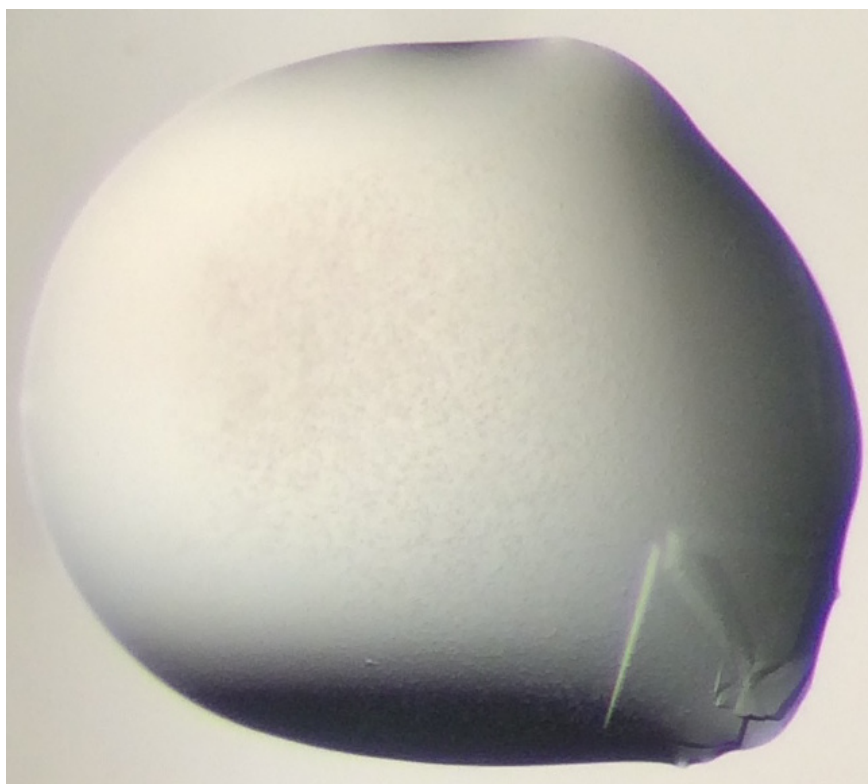


Figure 105 – Crystals of *AspRedAm*, at a protein concentration of 17 mg mL⁻¹ in buffer B (**3.3**), obtained in a 300 nL crystallisation drop in INDEX + 0.5 mM NADPH + 0.5 mM *rac*-rasagiline G3.

These crystals were optimised in an attempt to reduce a clustered morphology. The crystals were optimised (**Figure 106**) with NADPH and NADP⁺, as well as higher concentrations of *rac*-rasagiline. However, despite obtaining large crystals with a non-clustered morphology, the optimised crystals did not diffract beyond 3.5 Å resolution at the synchrotron.



Figure 106 – Optimised crystals of *AspRedAm* in conditions containing: (left) 0.2 M lithium sulphate, 0.1 M BisTris pH 6.5, 27% (w/v) PEG 3350, 10 mM NADPH and 10 mM *rac*-rasagiline, and (right) 0.2 M lithium sulphate, 0.1 M BisTris pH 6.5, 27% (w/v) PEG 3350, 5 mM NADP⁺ and 5 mM *rac*-rasagiline

AspRedAm was rescreened with commercially available screens utilising a 300 nL crystallisation drop and 54 μ L mother liquor with 0.5 mM NADP⁺, 0.5 mM 1-indanone and 0.5 mM propargylamine in an attempt to obtain binding sites for both substrates, as the oxidised cofactor should not reduce the substrates in solution. By obtaining the crystal structure with binding sites for the amine and carbonyl substrate, potential key hydrogen bonding residues that may be catalytically relevant can be identified, which would allow for a mechanism to be hypothesised. The crystal (**Figure 107**) was obtained in conditions containing 0.2 M magnesium chloride, 0.1 M BisTris pH 6.5, 25% (w/v) PEG 3350, 0.5 mM NADP⁺, 0.5 mM 1-indanone and 0.5 mM propargylamine at a protein concentration of 18 mg mL⁻¹ in buffer B (**3.3**). This crystal did not require further optimisation and was fished and flash cooled with liquid with liquid nitrogen without the addition of cryoprotectants.

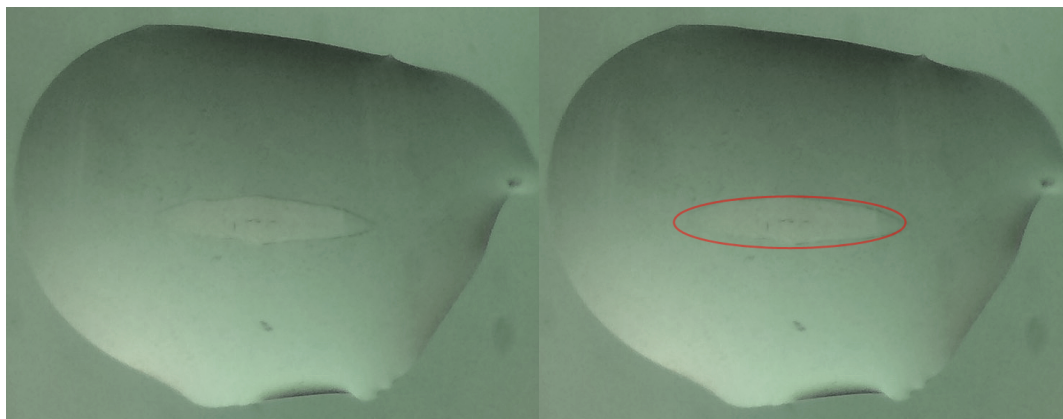


Figure 107 – Crystal (Left), highlighted in red (Right), of *AspRedAm*, at a protein concentration of 18 mg mL^{-1} , obtained in INDEX + 0.5 mM NADP^+ + $0.5 \text{ mM 1-indanone}$ + $0.5 \text{ mM propargylamine G11}$.

A full dataset was collected for *AspRedAm* ternary complex to 2.4 \AA resolution using 10% beam transmission (**3.6.1**) and helical data collection (**3.6.2**). The dataset was reprocessed using XDS¹⁴¹ using the xds GUI. The crystal structure of *AspRedAm* was solved in the space group *P1* with Molrep using *AspRedAm*-NADPH complex as the molecular replacement model. Iterative cycles of model building and refinement were performed. Water molecules were added last and the model was refined and checked. Finally the residual density between the dimer interface in the omit ($mF_o - DF_c$) map contoured to $\pm 3\sigma$ was assigned as NADP(H) and the amine product (*R*)-rasagiline. The crystal structure is composed of eight monomers in the asymmetric unit arranged as a tetramer of homodimers that forms a ring (**Figure 108**).

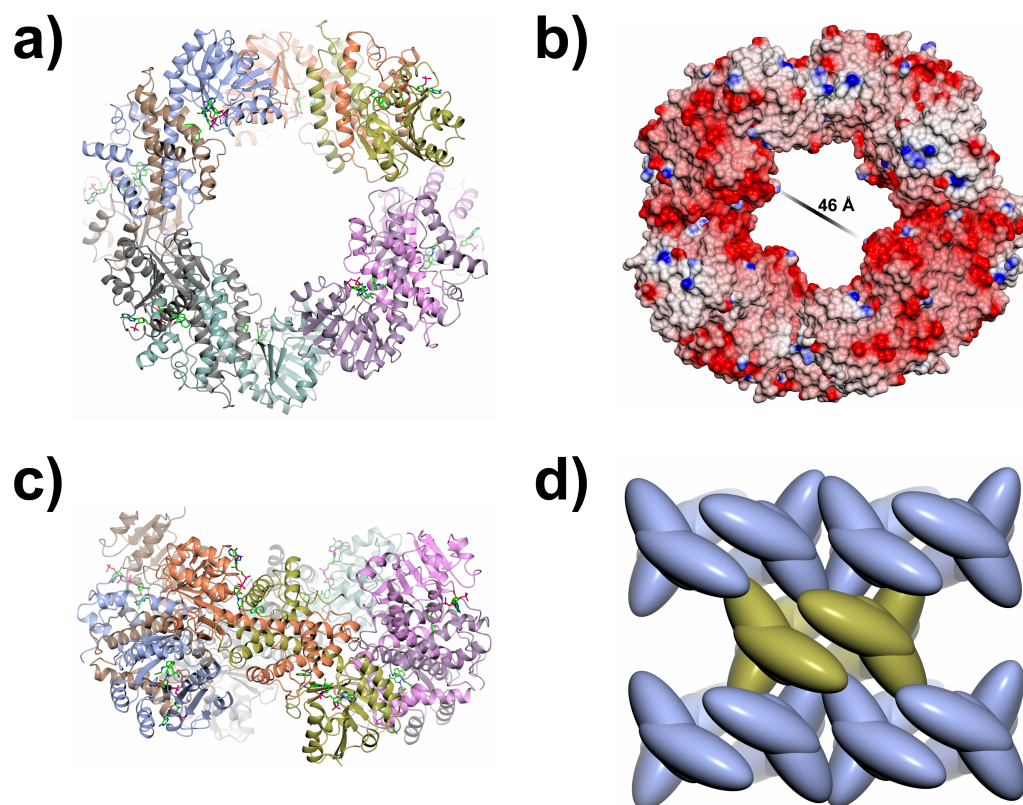


Figure 108 – The crystal structure of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. a) Top down secondary structure view. b) Top down surface view, the central cavity measures a distance of 46 Å. c) Side on secondary structure view. d) Side on simplified bloboid crystal contacts view where the gold models represent the asymmetric unit and the ice blue represents the nearby crystal packing.

The final structure was refined to R_{factor} and R_{free} values of 20.4% and 22.9%. The model was complete from residues 2 to 290 in all eight subunits (**Figure 109**).

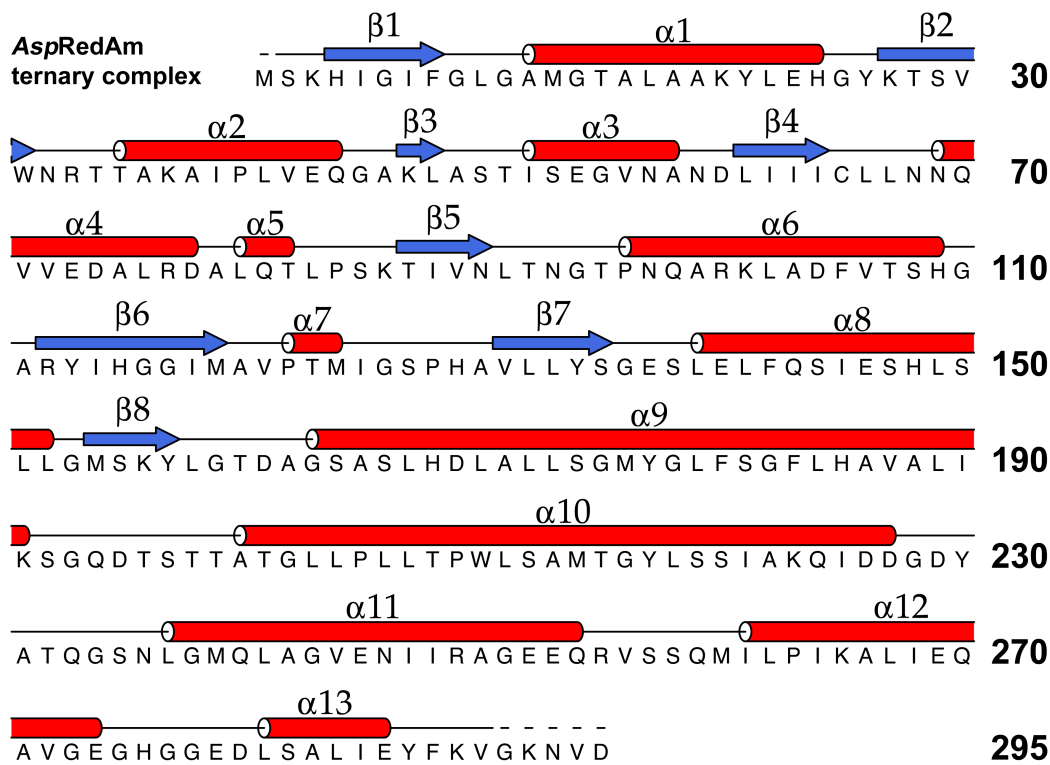


Figure 109 – The primary structure of *AspRedAm*-ternary complex with secondary structure assignment, created using DSSP¹⁵⁰⁻¹⁵¹ and represented using ALINE.¹⁵² Dashed lines represents areas of missing density that could not be modelled.

The dimeric structure adopts the IRED-fold (**Figure 110**), similar to previously obtained IRED structures.

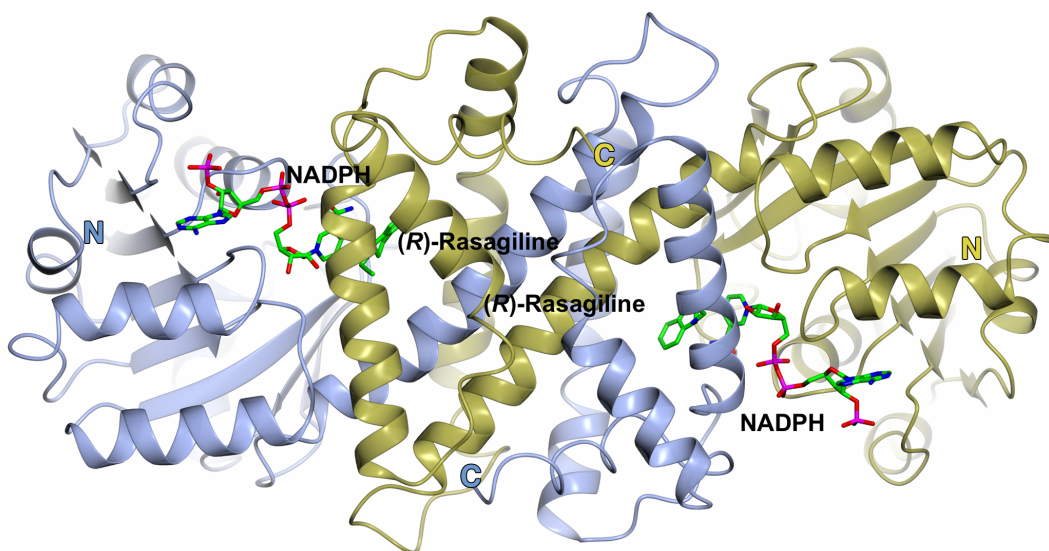


Figure 110 – The dimeric model of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. Subunits A and B are in blue and gold respectively.

Upon closer inspection of the active site, rather than having obtained a complex with NADP(H), 1-indanone and propargylamine, it appears that the reaction had turned over and formed the product with (*R*)-rasagiline trapped in the active site (**Figure 111**). This was not unexpected, as chemicals in the crystallisation condition could have aided the reaction in turning over whilst the enzyme is still active during crystallisation, along with potential excitation from X-ray sources upon data collection. At this resolution, 2.4 Å, ring puckering at the nicotinamide ring cannot be differentiated, therefore the cofactor was assigned as NADP(H) as its oxidation state could not be determined.

(*R*)-rasagiline is surrounded by residues (**Figure 111 - Left**) Leu173, Met176, Tyr177, Phe180, Trp210, Tyr217, Ser235 and Met239 creating a hydrophobic pocket, which is ideal for binding substrates and products with significant hydrophobic character, like (*R*)-rasagiline. Notable interactions include, facial π -stacking between Met176 and Tyr177, end on π -interaction between Met176 and the aromatic ring of (*R*)-rasagiline, and π -stacking between (*R*)-rasagiline and the nicotinamide ring of NADPH which aligns the substrate in an ideal position to be attacked by hydride of NADPH. The product was bound ideally between NADP(H) and Tyr177 (**Figure 111 - Right**), with distances that are not too dissimilar to what would be expected of hydride delivery from NADPH, 3.8 Å, and subsequent proton abstraction from Tyr177, 3.9 Å.

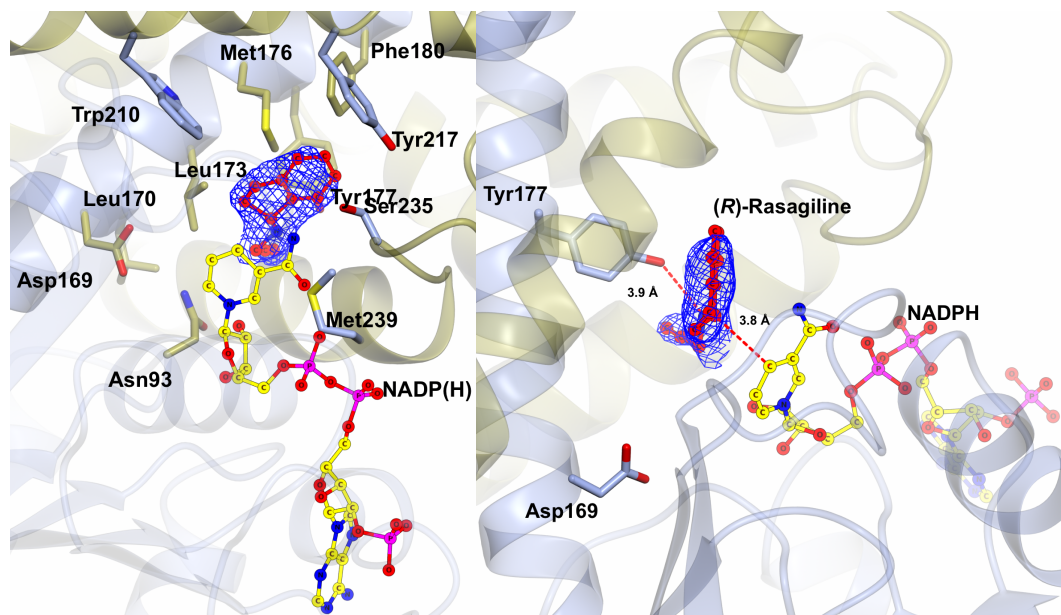


Figure 111 – Active site of the ternary complex of *AspRedAm* with NADP(H) and (*R*)-rasagiline. Subunits A and B are in blue and gold respectively. The ligands NADP(H) and (*R*)-rasagiline are in yellow and red circles respectively for illustrative purposes only. The electron density for (*R*)-rasagiline is the omit map ($mF_o - DF_c$) contoured to $\pm 3\sigma$ obtained from refinement prior to modelling and refining (*R*)-rasagiline. Left – View of active site with surrounding residues. Right – View of active site with distances within bonding distances shown.

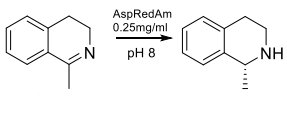

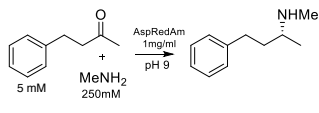
The statistics for the crystal structures are summarised in (Table 38).

Table 38 – Crystal data collection and refinement statistics for all AspRedAm data collection. Numbers in brackets refer to the data in the highest resolution shell.

	AspIREd 'closed' NADPH complex	AspIREd ternary complex with NADPH and (<i>R</i>)-rasagiline
Date Obtained	19th September 2015	2nd October 2015
Beamline	Diamond i02	Diamond i03
Wavelength (Å)	0.97950	0.97625
Resolution (Å)	89.91-1.95 (2.00-1.95)	75.48-2.35 (2.39-2.35)
Space Group	<i>P</i> 1 2 ₁ 1	<i>P</i> 1
Unit Cell (Å)	a = 39.27; b = 92.44; c = 84.51; α = γ = 90.00°; β = 100.3°	a = 79.82; b = 80.25; c = 123.89; α = 108.75; γ = 108.77; β = 90.05
No. of molecules in the asymmetric unit	2	8
Unique reflections	53033 (3964)	111053 (5488)
Completeness (%)	99.6 (99.7)	97.4 (97.1)
<i>R</i> _{merge} (%)	0.04 (0.77)	0.20 (0.55)
<i>R</i> _{p.i.m.} (%)	0.03 (0.66)	0.20 (0.55)
Multiplicity	2.0 (2.1)	1.8 (1.8)
<1/σ(<i>I</i>)>	15.9 (1.7)	2.9 (1.2)
CC _{1/2}	0.99 (0.50)	0.98 (0.60)
Overall <i>B</i> factor from Wilson plot (Å ²)	29	31
<i>R</i> _{cryst} / <i>R</i> _{free} (%)	18.5/22.5	20.4/22.9
r.m.s.d 1-2 bonds (Å)	0.02	0.01
r.m.s.d 1-3 angles (°)	1.94	1.35
Average main chain <i>B</i> (Å ²)	37	38
Average side chain <i>B</i> (Å ²)	43	41
Average water <i>B</i> (Å ²)	41	37
Average NADPH <i>B</i> (Å ²)	63	32
Average Ligand <i>B</i> (Å ²)	-	67
Normalised <i>B</i> factor	1.28	1.23

Residues in the hydrophobic binding site (**Figure 111**) for (*R*)-rasagiline were subjected to SDM studies. The mutations D169A, D169N and Y177A were produced by Aleku and co-workers. Once the mutant plasmids were prepared for expression in *E. coli*; expression was scaled up, and the mutants were purified with the activity assays (**Table 39**) performed by Aleku and co-workers .

Table 39 – Kinetic data of AspRedAm variants.

Enzyme variants									
	K_M [mM]	k_{cat} [S^{-1}]	k_{cat} / K_M [$S^{-1} mM^{-1}$]	K_M (ketone) [mM]	k_{cat} [S^{-1}]	k_{cat} / K_M [$S^{-1} mM^{-1}$]	K_M (ketone) [mM]	k_{cat} [S^{-1}]	k_{cat} / K_M [$S^{-1} mM^{-1}$]
WT	0.352	3.243	9.213	1.901	1.470	0.773	0.912	0.112	0.123
D169A	1.101	0.016	0.015	2.700	0.008	0.003	-	n.d	n.d
D169N	0.320	0.009	0.028	2.080	0.007	0.003	-	n.d	n.d
Y177A	0.689	0.063	0.091	2.212	0.050	0.023	-	n.d	n.d

The mutant Y177A indicates that Tyr177 is required RedAm activity however, Asp169 is significantly more important for RedAm, with k_{cat} values being approximately 7 times higher than those from the mutations at Asp169. The activity assays show that all mutants retain some IRED activity albeit with significantly reduced catalytic efficiency. Interestingly for all mutations, RedAm activity is significantly compromised more than IRED activity. The overall outcome of the mutations is that the mutants turn a RedAm into an IRED. Therefore Tyr177 and Asp169 are catalytically relevant for RedAm activity.

4.3.1. Discussion

The data obtained from the reductive amination biotransformations in **Table 37** and **Table 37** were plotted (**Figure 112**) on a relative specific activity graph. The graph suggests that substrates that fall within the green areas, high activity, should be able to utilise an amine loading between 1 and 4. This gradually decreases to a suggested amine loading between 4 and 10. Substrates of moderate activity, suggest utilising an amine loading between 10 and 20. Substrates of low activity suggest utilising an amine loading between 20 and 50. Where as substrates of minimal activity, where amination is potentially not possible, suggests using an amine loading of 50 or more.

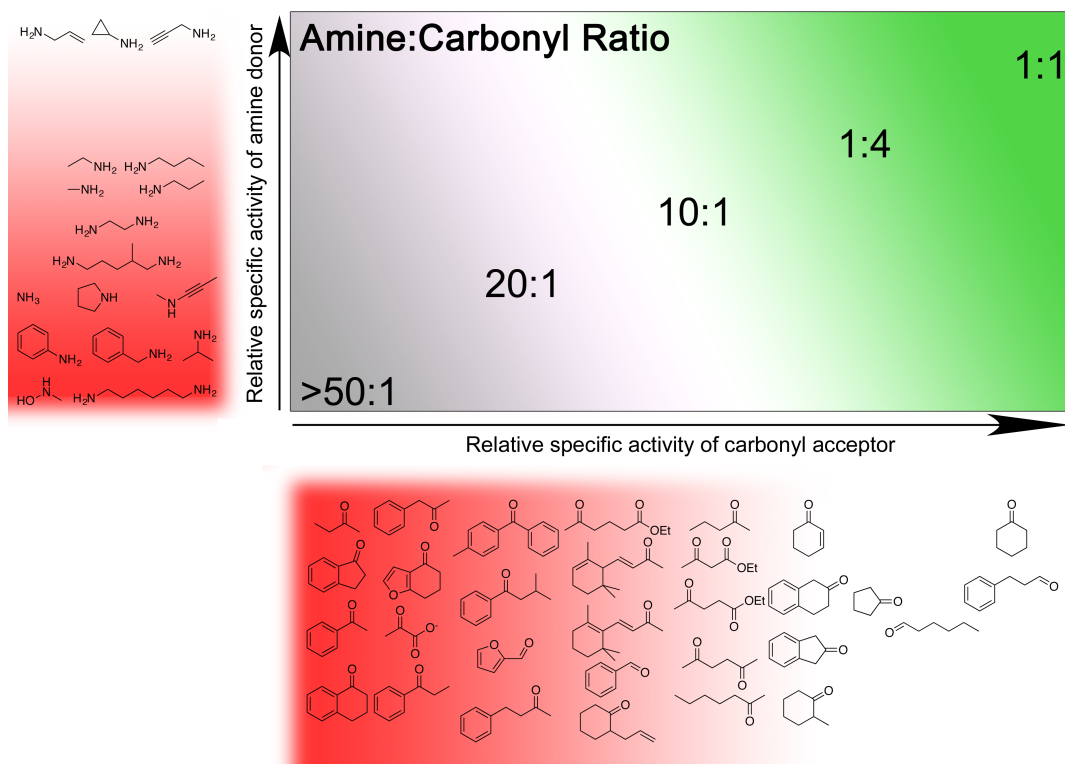


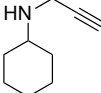


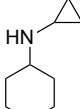

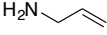
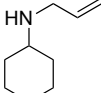





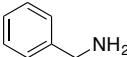
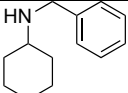
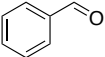
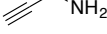
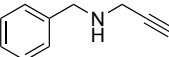
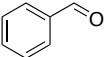

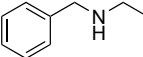
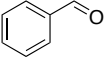
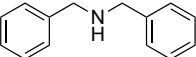
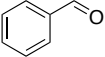
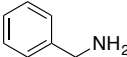
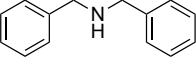
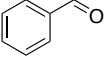

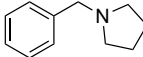


Figure 112 – Chart displaying relative activity of amine/carbonyl pairs in reductive amination reactions with suggested amine-to-carbonyl ratios for biotransformations. Compounds presented in the plot area are representative examples of products obtained in biotransformations.

For example, a reaction with propargylamine and cyclohexanone has a suggested carbonyl-to-amine ratio of 1:1-4. Whereas, a reaction with propargylamine with 1-indanone has a suggested carbonyl-to-amine ratio of 1:≥50. The suggested amine loading from **Figure 112** was experimentally tested with further biotransformations (**Table 40**) performed Aleku and co-workers (personal communication)

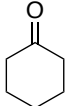
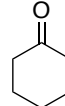
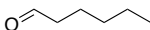
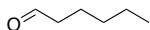
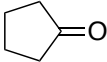
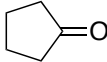
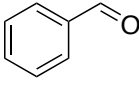
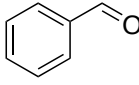
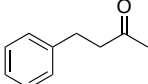
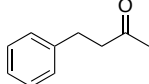
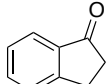
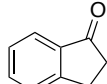
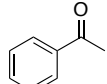
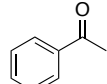
Table 40 – AspRedAm catalysed reductive amination of carbonyl compounds at low molar ratios.

Entry	Ketone	Amine	Ketone : Amine ratio	Product	Conversion (%)
1			1:1		94
2			1:1		90
3			1:1		73
4		—NH_2	1:4		>99
5		NH_3	1:4		>99
6			1:1		84
7			1:1		92
8			1:4		>99
9		NH_3	1:4		50
10			1:4		>99
11			1:10		>99

Interestingly, despite benzaldehyde being considered a carbonyl of moderate activity and pyrrolidine being considered an amine of low activity, the reaction was still performed at a carbonyl-to-amine ratio of 1:10 with >99% conversion (entry 11 – **Table 40**). The biotransformation between benzaldehyde and ammonia (entry 9 – **Table 40**) only forms the double amination product, which suggests the initial formation of benzylamine, which then attacks benzaldehyde to yield dibenzylamine. Entry 9 (**Table 40**) only results in 50% yield, whereas, dibenzylamine can be obtained in >99% yield with using benzylamine as the starting amine as in entry 10 (**Table 40**).

The ternary complex (**Figure 111**) revealed a binding site for (*R*)-rasagiline, which is analogous to the binding site for (*R*)-MTQ as in *Ao*IRED (**Figure 91**). The residues surrounding (*R*)-rasagiline were probed by SDM studies, which revealed that Asp169 and Tyr177 were necessary for RedAm activity, however they were not necessary for IRED activity (**Table 39**). Interestingly, *Ao*IRED can also perform reductive amination (Aleku and co-workers) albeit with significantly reduced activity (**Table 41**), this may be due to *Ao*IRED having the residue Tyr179 that is analogous to Tyr177 in *Asp*RedAm. Wild type *Ao*IRED's reductive amination activity could be comparable to the D169N *Asp*RedAm mutant.

Table 41 – Comparison of reductive amination activities between *Asp*RedAm and *Ao*IRED.

Entry	Carbonyl acceptor	Specific activity with propargylamine (U/mg)		Entry	Carbonyl acceptor	Specific activity with methylamine (U/mg)	
		<i>Asp</i> RedAm	<i>Ao</i> IRED			<i>Asp</i> RedAm	<i>Ao</i> IRED
1		6.7	1.2 × 10 ⁻¹ (55.8)	8		2.2	1.6 × 10 ⁻² (137.5)
2		4.7 × 10 ⁻¹	8.4 × 10 ⁻² (55.9)	9		2.1	2.3 × 10 ⁻¹ (9.1)
3		2.9	7.5 × 10 ⁻² (38.6)	10		8.9 × 10 ⁻¹	1.7 × 10 ⁻² (52.4)
4		9.3 × 10 ⁻¹	4.1 × 10 ⁻² (22.7)	11		2.7 × 10 ⁻¹	1.6 × 10 ⁻² (16.9)
5		6.8 × 10 ⁻¹	1.6 × 10 ⁻² (42.5)	12		1.7 × 10 ⁻¹	7.0 × 10 ⁻³ (24.3)
6		1.7 × 10 ⁻¹	1.2 × 10 ⁻² (14.2)	13		2.2 × 10 ⁻²	1.3 × 10 ⁻³ (16.9)
7		1.3 × 10 ⁻¹	7.0 × 10 ⁻³ (18.6)	14		1.7 × 10 ⁻²	4.0 × 10 ⁻⁴ (42.5)

The native, *apo*-*Asp*IRED crystal structure could not be obtained. This is likely due to the dynamic nature of the IRED-fold, as seen with low-resolution native *Asp*RedAm crystals (**Figure 99**) and the dynamic structure of the IRED from *Amycolatopsis orientalis*, *Ao*IRED (**Figure 86**). As such, substrates and ligands are required to stabilise the IRED-fold in a 'closed' conformation, also seen with the

IREd from *Nocardiopsis halophila*, NhIREd in complex with OBDG (**Figure 67**) where the native structure could also not be obtained.

Scheller and co-workers, who created the IREd database^{102, 105} analysed origins of the current IREds under investigation, as well as potential new IREds (**Figure 113**). *AspRedAm*, as a fungal IREd/RedAm was found to be in a separate sequence space compared to bacterial IREds. *AspRedAm* and other potential fungal IREds were grouped together in a relatively isolated island loosely related to bacterial IREds and each with the homologous Asp169 in *AspRedAm*. The potential exists for these fungal IREds to also be RedAms.

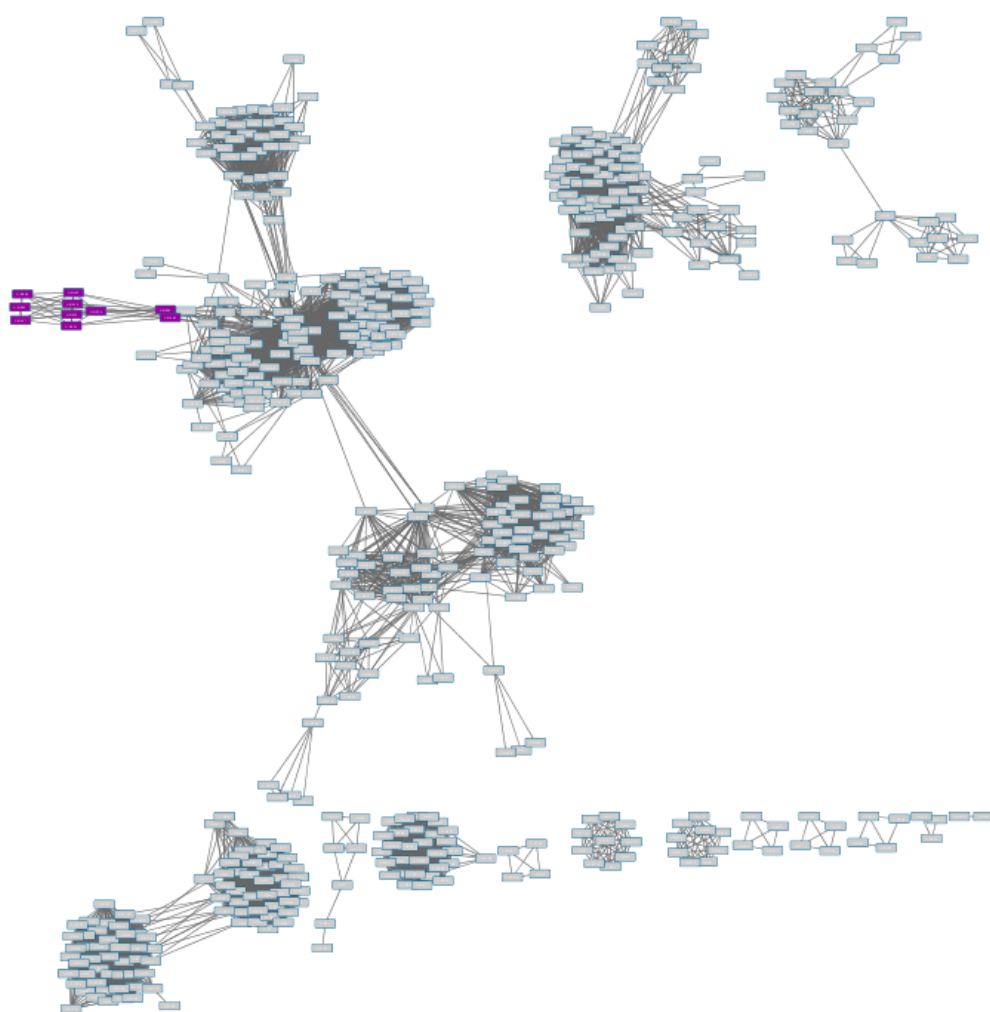


Figure 113 – IREd superfamily analysis. Grey sequences represent potential IREds of bacterial origin. Purple sequences represent potential IREds of fungal origin. Sequence analysis performed by Scheller and co-workers (personal communication).

The remaining fungal IREds have been characterised by Mahima Sharma (personal communication, in our group), and were found to be capable of catalysing

reductive aminations. These fungal RedAms have been named *AdRedAm*, *AnRedAm*, and *AtRedAm* for *Ajellomyces dermatitidis* (67% homology to *AspRedAm*), *Aspergillus nidulans* (75% homology to *AspRedAm*) and *Aspergillus terreus* (75% homology to *AspRedAm*) respectively. Through sequence analysis and structural characterisation, the fungal RedAms had certain residues in common with each other (**Figure 114**), however certain residues were noticeably different compared to IREDs that were unable to catalyse reductive amination.

```

AdRedAm      ----MANSPVSVFGLGAMGTALATQFLRKGHKTTVWNRTPAKAQ--PLIAIGASHAPTID
AnRedAm      -----MSKISLFLGLGSMGQALAHRYIDTGYTTTWNRTPEKAQSSGLIQGAHQALTVA
AtRedAm      MATTTTTLTKLTFGLGAMGTAMATQFLKQGHPTTWNRTAAKAN--PLVEQGAHLAATIP
AspRedAm     -----MSKHIGIFGLGAMGTALAAKYLEHGYKTSVWNRRTAKAI--PLVEQGAKLASTIS
              . : :****:* *:* *:: : . :****: ** *:* * * : *

AdRedAm      SAAAASSLLIICQLDKASVMQTLQOAPTAWAAKTIVDLTNGTPAHARETADWALAHGARY
AnRedAm      EGLEAADMVILCLLDNASVRETLSKAVTSLSGKTI VNLNTPPTQARDLSEWAGAHGAEY
AtRedAm      AAI AASPLLI FCLLDNAAVEQTLAAGPPSLAGKTI LNLTNGTPSQARRLATLASARGARY
AspRedAm     EGVNANDLIIICLLNQVVEDALRDALQTLPSKTI VNLNTPPNQARKLADFTVSHGARY
              . * :*: * *:: * :* . .****:***** :* * : :*: *

AdRedAm      IHGGIMAVPFMI--GQPDAMILYSGP--AEVFEGVKDTLSVLGTNTYVGEDVGLASLHDL
AnRedAm      IHGGIMAVPDMISSGSPHSILLYSGK-SNEVFTRIEPDLAHLGAAKFLGTDPGSASLHDL
AtRedAm      FHGGIMATPDMI--GAPHAVILLYSGGSAETYASVEGVLAVLGSGKYLGDAGSASLHDL
AspRedAm     IHGGIMAVPTMI--GSPHAVLLYSGE-SLELFQSIESHLSLLGMSKYLGTDAGSASLHDL
              .*****.* ** * * :*:***** : : : *:* . :*: * * *****

AdRedAm      ALLSGMYGLFSGFTHAVALVQSAN---IPAAGFVATQILPWLMTAMQHLLNLLATQVDEKD
AnRedAm      ALLSGMYGLFSGFFHATALVRSQSD--TTSEGFL-ELLIPWLGAMTHYLGLLAKQIDSGD
AtRedAm      ALLSGMYGLFAGFLHATALVRSEGE-GVSATEFL-GLLAPWLQAMTGYLGLLARQIDDG
AspRedAm     ALLSGMYGLFSGFLHAVALIKSGQDSTTTATGLL-PLLTPWLSAMTGYLSSIAKQIDGD
              *****:* * * *:*:* : : : * * * * * : : * * * .

AdRedAm      YDGGSSLDMQAKAAPNILEASQAQGVSVELIQPIFKLIERRVEEGKSGEGLAALVGMIM
AnRedAm      YTTQGSNIAMQVAGLENI IAASEEVGVTTPAFILPIQRLMKRAADEGHGNTDISSLIQYLR
AtRedAm      YTAQTSNLEMQLVALENACAASREQGVSAEVMLPLKGLVERAVREGRGGHDISSLIQYFR
AspRedAm     YATQGSNLGMQLAGVENIIRAGEEQRVSSQMILPIKALIEQAVGEGHGGEDLSALIEYFK
              * * .: * * . * *.. * : .: * : :*: . *:* :*: .

AdRedAm      KGGTKDSV
AnRedAm      TRE-----
AtRedAm      NASV----
AspRedAm     VGKNVD--

```

Figure 114 – Sequence alignment between the fungal RedAms, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹

Two key residues include Asp169 and Tyr177, both present in *AspRedAm* and have been determined to affect activity when mutated to D169A and Y177A, as performed by Aleku and co-workers (personal communication). Another possibly significant residue is Asn93, which has been observed to be within hydrogen bonding distance of Asp169. However mutagenesis studies have not yet been performed on Asn93. A possible effect of Asn93 on Asp169 maybe a lowering of the pK_a , which would allow Asp169 more likely to be deprotonated at pH 9.0. With

the information currently known, a working hypothesis for a mechanism can be deduced (**Figure 115**).

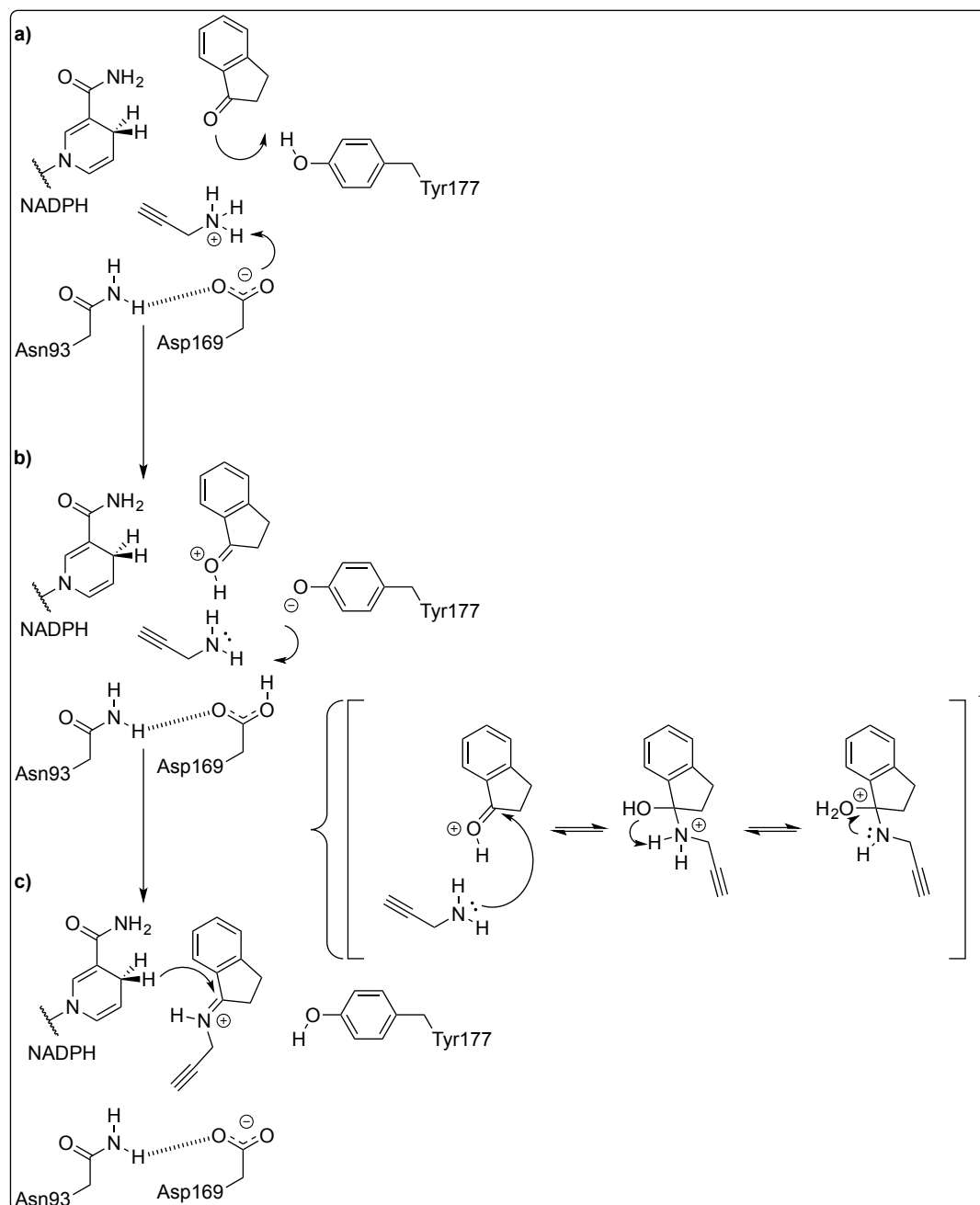


Figure 115 – Proposed mechanism for reductive amination with RedAms. a) Activation of the carbonyl by Tyr177 and deprotonation of the ammonium substrate by Asp169. b) Nucleophilic attack of the carbonyl by the amine. c) Imine reduction by NADPH. Residue numbering is based on the sequence numbering of AspRedAm.

For reductive amination to be catalysed, the carbonyl is bound and activated by Tyr177, the ammonium substrate is subsequently bound to allow the enzyme to close up into a tighter complex. Due to the hydrogen bonding of Asn93 to Asp169, Asp169 is more likely to be deprotonated, which allows for the

ammonium substrate to be deprotonated by Asp169 (**Figure 115a**). With both substrates now activated and in closer proximity, would enable a classical reductive amination mechanism for which the amine substrate, a nucleophile, attacks the activated carbonyl (**Figure 115b**). After the removal of water, the imine bond is formed, which is attacked by NADPH to form the chiral amine product (**Figure 115c**).

AspRedAm is an incredibly versatile IRED with its strengths being an RedAm capable of catalysing certain reductive amination reactions at a carbonyl-to-amine ratio of 1:1. This makes *AspRedAm* an invaluable tool for catalysis and biocatalysis as the biotransformations are performed with minimal solvent in water-based buffer systems. However, as *AspRedAm* is natively from fungal origin, *E. coli* may not be its ideal host as seen with its instability during and after purification. Despite this, *AspRedAm* could be a potential candidate for large-scale industrial biocatalysis once it has been further evolved for stability.

4.4. The bacterial RedAm from *Streptomyces* sp. PRh5 (*StrepRedAm*)

RedAms have been demonstrated to be incredibly useful and powerful tools as catalysts. As such, they are potentially highly desirable biocatalysts, however only a handful of RedAm sequences are currently known. Therefore genome mining was performed with limited information from the recent RedAm investigations, in an attempt to find bacterial RedAms with a more diverse sequence and substrate spectrum.

Eukaryotic sequences were omitted from the sequence search and bacterial sequences were targeted with the 'triad' of Asn93-Asp169-Tyr177 (*AspRedAm*), in an attempt to find a bacterial RedAm. In non-RedAm IREDs, Asn93-*AspRedAm* has been observed to be serine, and Tyr177 is often tryptophan. A potential IRED with RedAm activity from *Streptomyces* sp. PRh5 (*StrepRedAm*) was found (**Figure 116**) with the desired criteria. From analysis by sequence similarity, the sequence was predicted by the server to be a 6-phosphogluconate dehydrogenase.

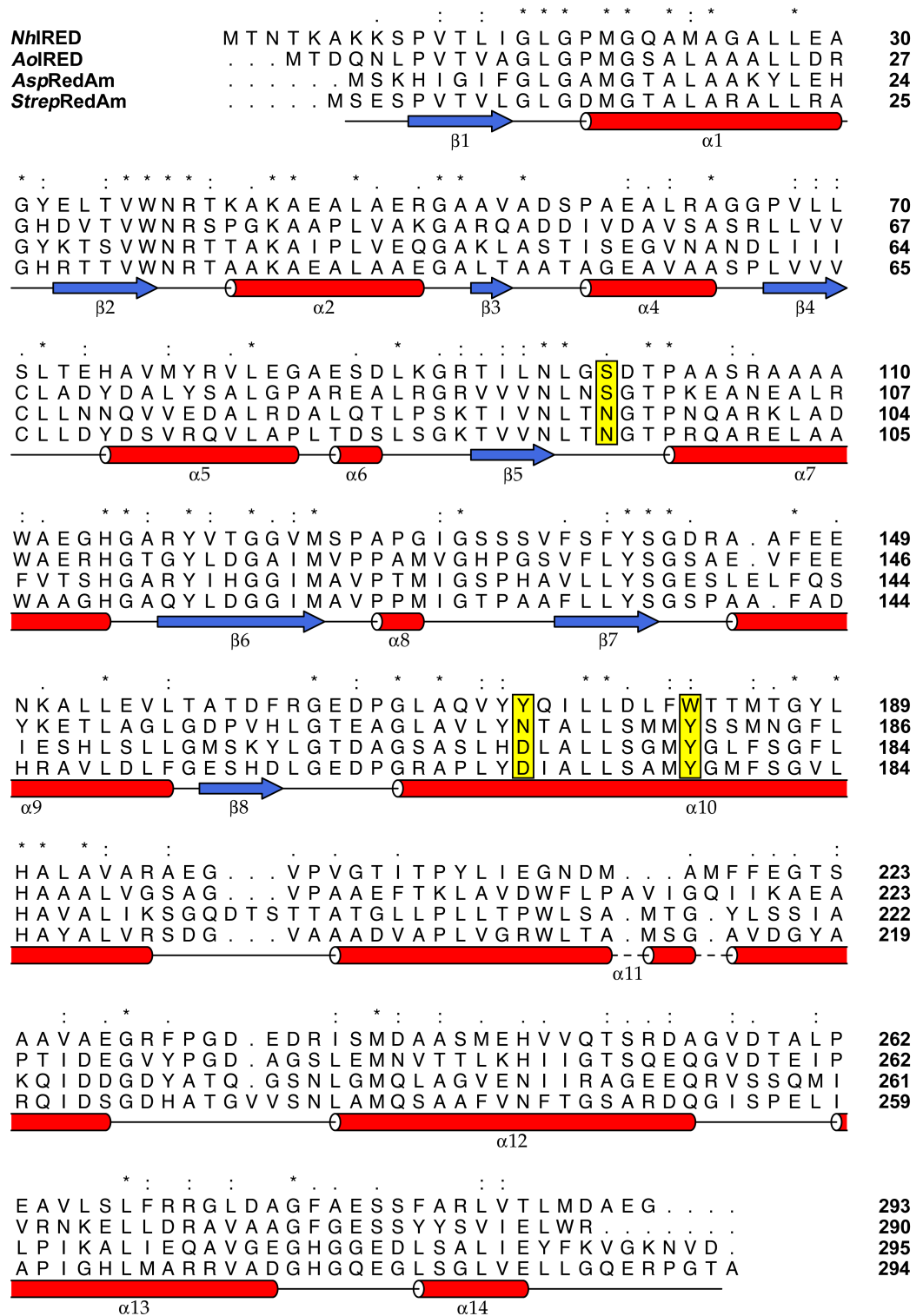


Figure 116 – Sequence alignment between IREDs and RedAms, where “*” are homologous residues, “:” are strongly similar residues and “.” are weakly similar residues. The DSSP is of the ternary complex of *Asp*RedAm. Residues of interests are highlighted in yellow with red font. Protein sequence alignment was performed using Clustal Omega.¹⁴⁹ *Asp*RedAm shares 69.9%, 64.2% and 61.1% sequence homology with *Nh*IRED, *Aoi*IRED and *Strep*RedAm respectively.

A gene sequence for *StrepRedAm* was generated by reverse translating the protein sequence (**Figure 117**) followed by subsequent codon optimisation for expression in *E. coli* (**Figure 118**).

```
>StrepRedAm
MSESPVTVLGLGDMGTALARALLRAGHRTTVWNR TAAKAEALAAEGALTA
ATAGEAVAASPLVVVCLLDYDSVRQVLAPLTDLSL SGKTVVNL TNGTPRQA
RELAAWAAGHGAQYLDGGIMAVPPMIGTPAAFLLYSGSPAAFADHRAVLD
LFGESHDLGEDPGRAPLYDIALLSAMYGMFSGVLHAYALVRS DGVAADV
APLVGRWLTAMSGAVDGYARQIDSGDHATGVVSNLAMQSAAFVNFTGSAR
DQGISPELIAPIGHLMARRVADGHGQEGLSGLVELL GQERPETA
```

Figure 117 – The protein sequence for *StrepRedAm* in fasta format.

The synthetic gene for *StrepRedAm* was ordered from GeneArt Gene Synthesis (ThermoFisher Scientific).

```
>StrepRedAm
ATGAGCGAAAGTCCGGTTACCGTTCTGGGTCTGGGTGATATGGGCACCGCACTGGCACGTGCAC T GCTGCGTG CAGGTCATCGT
ACCACCGTTTGGAATCGTACCGCAGCAAAAGCAGAAGCACTGGCAGCAGAAGGTGCACTGACCGCAGCCACCGCAGGCGAAGCA
GTTGCAGCAAGTCCGCTGGTTGTTGTTTGTCTGCTGGATTATGATAGCGTTCGTCAGGTTCTGGCACCGCTGACCGATAGCCTG
AGCGGTA AAAACCGTTGTTAATCTGACCAATGGTACACCGCTCAGGCACGTGA ACTGGCAGCCTGGGCAGCAGGTCATGGTGCA
CAGTATCTGGATGGTGGTATTATGGCAGTTC CGCCTATGATTGGCACACCGGCAGCATTCTGCTGTATAGCGGTAGTCCGGCA
GCCTTTCAGATCATCGTGCAGTTCTGGACCTGTTTGGTGAAAGCCATGATCTGGGTGAAGATCCGGGTCGTGCACCGCTGTAT
GATATTGCCCTGCTGAGCGCAATGTATGGTATGTTTAGCGGTGTTCTGCATGCATATGCCCTGGTTCGTAGTGATGGTGTGCA
GCAGCAGATGTTGCACCTCTGGTTGGTCTGGCTGACCGCAATGAGCGGTGCAGTTGATGGTTATGCACGTCAGATTGATAGC
GGTGATCATGCAACCGGTGTTGTTAGCAATCTGGCAATGCAGAGCGCAGCATTTGTGAATTTACCGGTAGCGCACGTGATCAG
GGTATTAGTCCGGAAC TGATTGCACCGATTGGTCATCTGATGGCACGTGCTGTTGCAGATGGTCATGGTCAAGAAGGTCTGAGC
GGTCTGGTTGAACTGCTGGGTCAAGAACGTCCGGGTACAGCATAA
```

Figure 118 – The DNA sequence for *StrepRedAm*, codon optimised for *E. coli*, in fasta format.

Primers for *StrepRedAm* were manually designed (**3.1.1**), such that the gene product could be annealed into the pET-YSBLIC3C vector by In-Fusion® cloning (**3.1.6**). The gene was amplified by TD-PCR (**3.1.3**) using KOD HotStart DNA polymerase (**3.1.2**) with the primers listed in **Table 43** and conditions outlined in **Table 42**.

Table 42 – Primers and PCR conditions for the amplification of *StrepRedAm* to be annealed by In-Fusion® into the pET-YSB LIC3C vector.

Step	Time	Temperature (°C)	Cycles
Denature	2 min	95	1
Denature	20 s	95	15
Anneal	10 s	65 ^[a]	
Extend	15 s	70	
Denature	20 s	95	20
Anneal	10 s	55	
Extend	15 s	70	
Extend	5 min	70	1
Cool	∞	10	

[a] Temperature was reduced by -0.5°C per cycle

The gene product was purified by gel extraction and was annealed into the pET-YSB LIC3C vector. A cloning strain of *E. coli*, NovaBlue Singles™, was transformed (3.1.8) by the nicked plasmid. Starter cultures were prepared from colonies of the transformants. The repaired plasmids were extracted and purified by mini prep extraction kit. The plasmids were subjected to double digest by restriction enzymes (3.1.9) to confirm the presence of the gene of interest prior to sequencing. The gene of interest had an expected gene size of 885 bp, however there is a cut site for *NdeI* at 560 bp, therefore the presence of two fragments were expected (Figure 119).

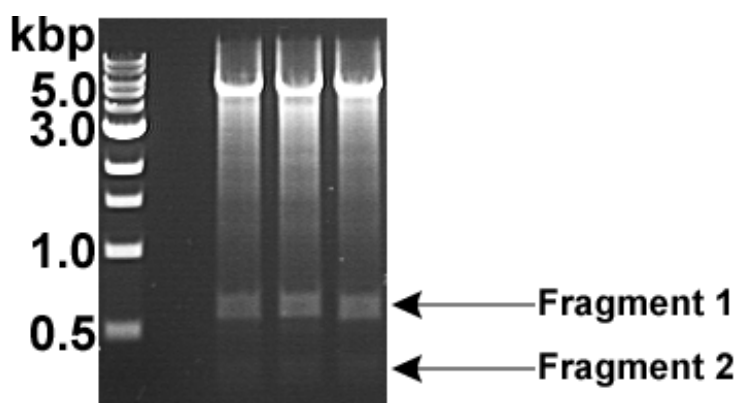


Figure 119 – Fragment analysis of double digest by restriction enzyme, confirming the presence of positive gene insertions.

Once sequencing had confirmed the gene of interest had been inserted into the vector correctly, the plasmid was prepared for expression. An expression strain of *E. coli*, BL21(DE3), was transformed by the plasmid encoding *StrepRedAm*. Small-

scale expression was performed to assess the optimal expression temperature (Figure 120).

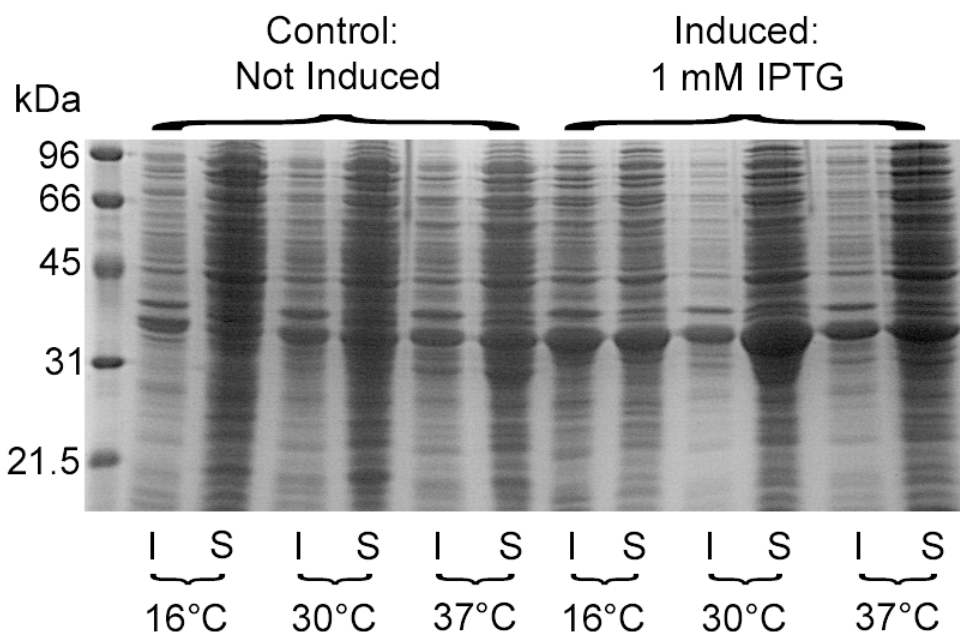


Figure 120 – The expression test for *StrepRedAm*, where “I” are insoluble fractions from the pellet and “S” are soluble fractions of cell lysate. The expected molecular weight of *StrepRedAm* is 33.5 kDa, including the 6His-tag and HRV3C protease precision site.

Despite induced expression being greater at 30°C, 16°C was chosen as the ideal expression temperature. There is often overall elevated levels of expression of all genes in *E. coli*, typically generating more unspecific binding for the nickel charged columns, which results in a more difficult purification.

Using the ideal expression temperature obtained from small-scale expression, gene expression was scaled up for protein purification. *StrepRedAm* was initially purified nickel-affinity, followed by anion exchange (Figure 121) and SEC (Figure 122). However, despite an extra purification step, the SDS-PAGE analysis (Figure 123) indicates increasing quantities of degradation over the course of the three purification steps and the final purified protein was not sufficiently pure for crystallographic or kinetic assay studies.

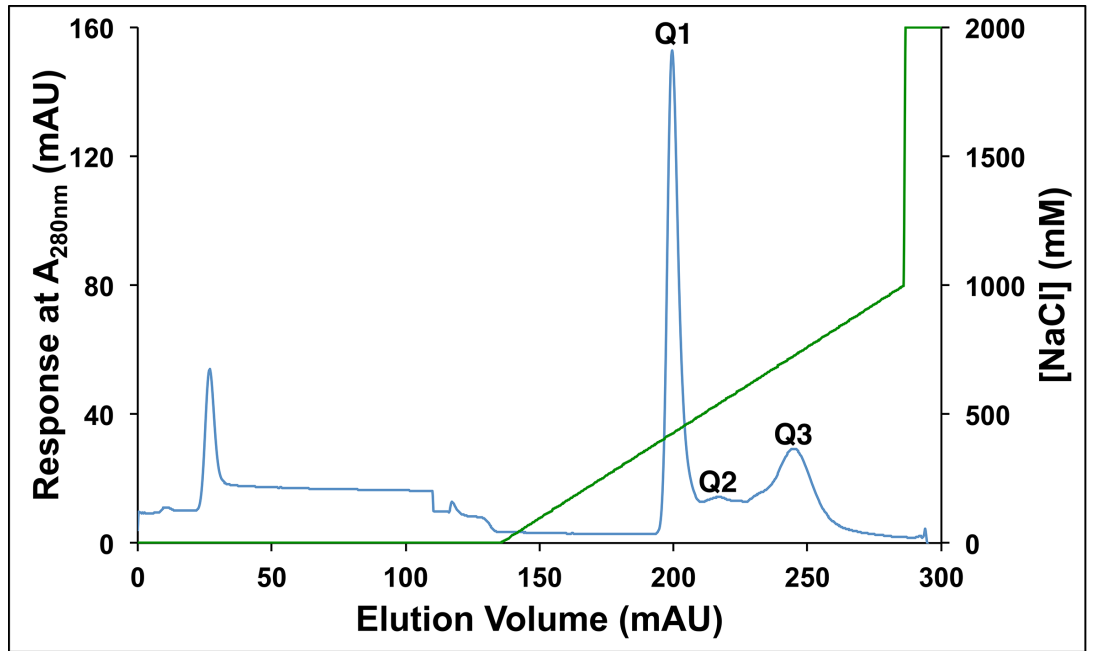


Figure 121 – The chromatogram for the anion exchange purification for *StrepRedAm*. Main peak elution occurs at approximately 500 mM NaCl.

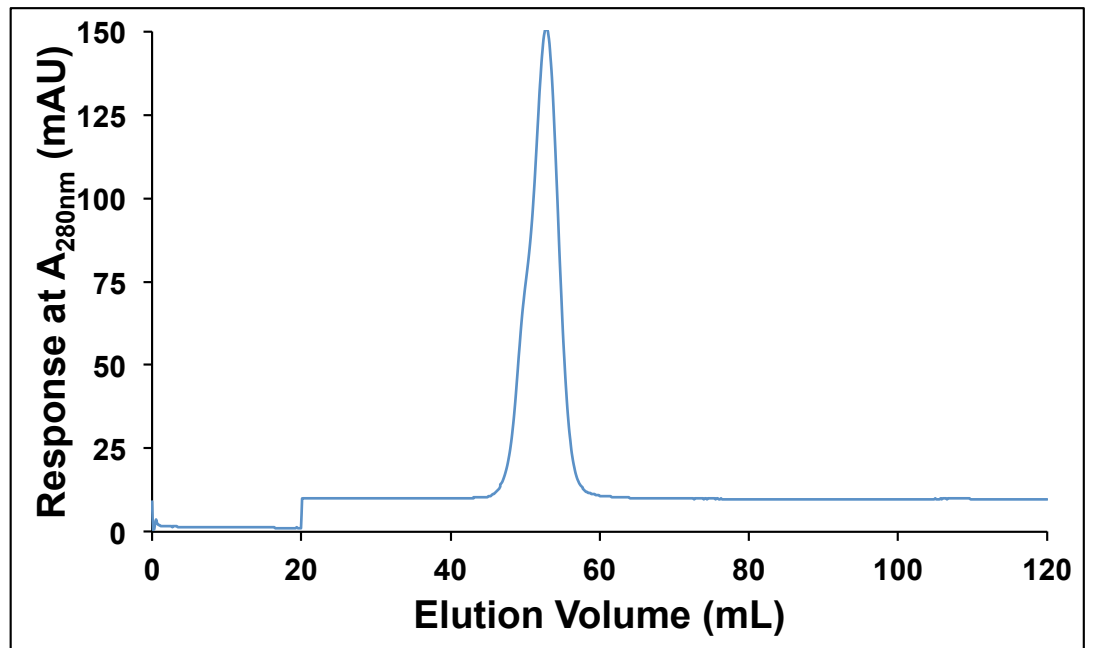


Figure 122 – The chromatogram of the SEC for *StrepRedAm*. The peak eluting between 50-60 mL indicates a species eluting with a molecular weight of approximately 60 kDa.

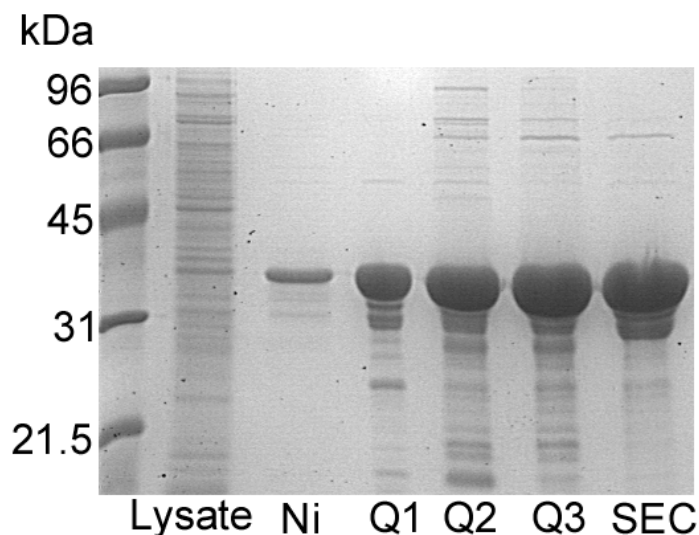


Figure 123 – SDS-PAGE analysis for the purification of *StrepRedAm* by nickel affinity, anion exchange (Q) and SEC. Samples Q1, Q2 and Q3 correspond to samples taken from the anion exchange purification in Figure 120. Over the course of the purification, degradation was increased. After SEC the purity of the protein is not sufficient for crystallographic and kinetic assay studies.

To counter the degradation, 10% (v/v) glycerol was added to all buffers and protease inhibitors (3.3) were added to the cells suspension prior to cell lysis. Only the last fractions (Figure 124) from the nickel purification were taken to be further purified by SEC (Figure 125).

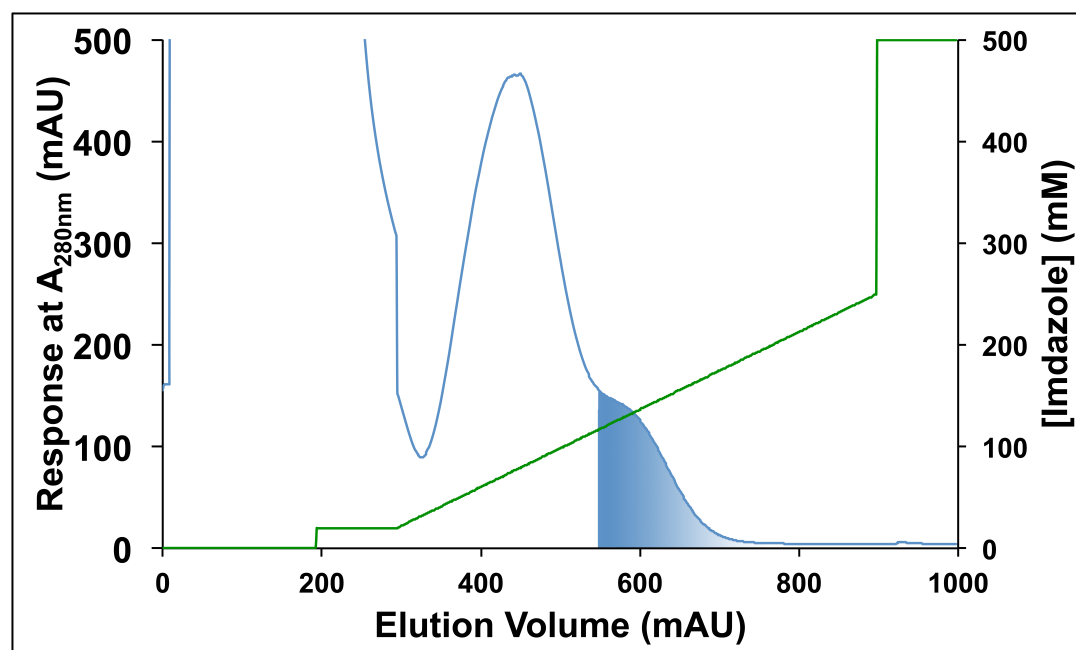


Figure 124 – The chromatogram for the nickel purification of *StrepRedAm* with protease inhibitors. The peak area highlighted in blue was the fractions taken for further purification.

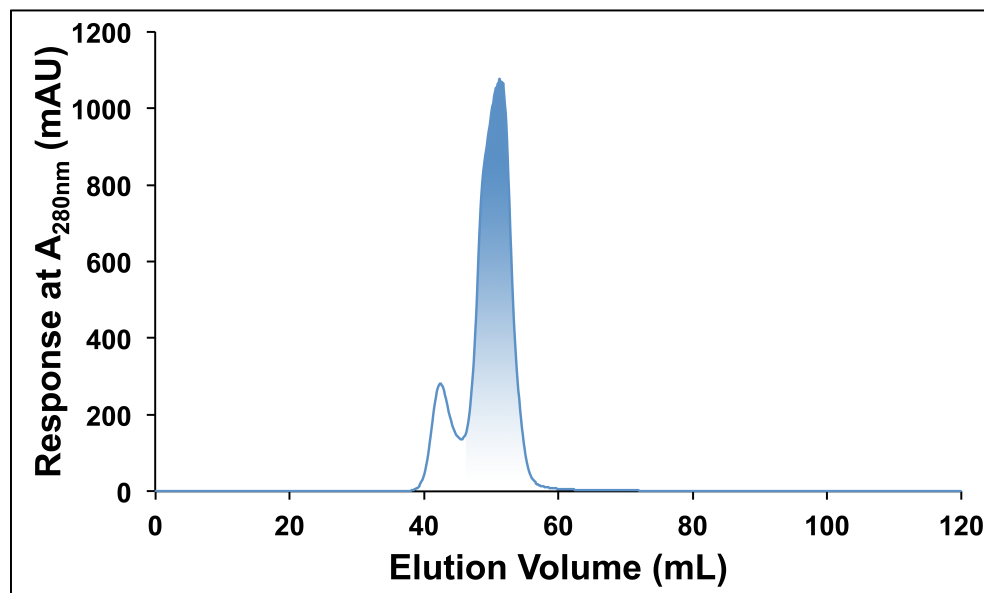


Figure 125 – The chromatogram of SEC for *StrepRedAm*. The main peak eluting between 50-60 mL indicates a species eluting with a molecular weight of approximately 60 kDa. The peak area highlighted in blue was the fractions taken for further characterisation

The SDS-PAGE (Figure 126) analysis of the nickel purification indicates a significant contaminant with a similar molecular weight to *StrepRedAm*. Despite using protease inhibitors, the contaminant could still be degradation products of *StrepRedAm*. If the contaminant is degradation of *StrepRedAm*, the SDS-PAGE analysis (Figure 126) indicates C-terminal proteolysis as the 6His-tag is fused onto the N-terminus and is still binding to the nickel charged resin.

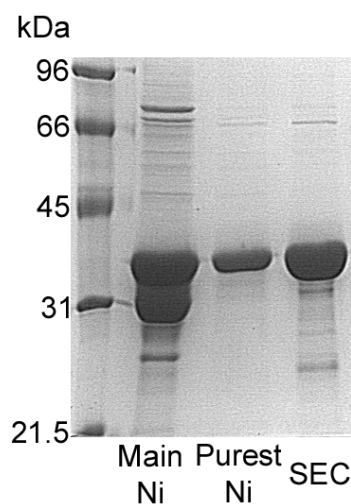


Figure 126 – SDS-PAGE analysis of the purification of *StrepRedAm*. Main Ni is a sample from the main peak of the nickel purification. Purest Ni are the fractions towards the end of the peak elution which were taken for further purification. SEC is the protein obtained after SEC.

Throughout all purification steps, *StrepRedAm* did not precipitate whilst stored at room temperature or at 4°C for over one month. As such, *StrepRedAm* is noticeably more stable in solution compared to *AspRedAm*.

Once the *StrepRedAm* was purified (**Figure 125**), *StrepRedAm* was further characterised with reductive amination biotransformations (**Figure 127**).


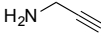
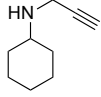

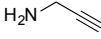
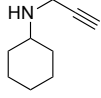

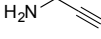
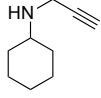

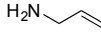
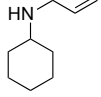

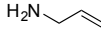
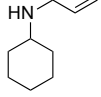

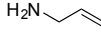
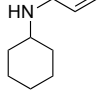

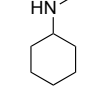

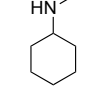

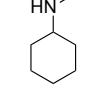

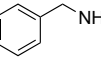
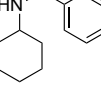

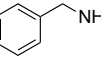
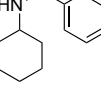

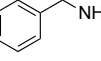
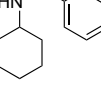
Entry	Ketone	Amine	Ketone:Amine ratio	Product	Conversion (%)
1			1:1		88
2			1:4		97
3			1:50		15
4			1:1		81
5			1:4		>99
6			1:50		>99
7		—NH_2	1:1		50
8		—NH_2	1:4		95
9		—NH_2	1:50		>99
10			1:1		19
11			1:4		20
12			1:50		0

Figure 127 – Achiral reductive amination biotransformations performed with *StrepRedAm* at carbonyl-to-amine ratios of 1:1, 1:4 and 1:50.

From the initial achiral screening, it appears *StrepRedAm* is capable of performing reductive aminations at a carbonyl-to-amine ratio of 1:1 with comparable activity to 1:4, which results in the highest conversion in nearly all cases (**Figure 127**). Interestingly the best conversions are not necessarily obtained at a carbonyl-to-amine ratio of 1:50 as seen with entries 3 and 12 (**Figure 127**), this could be due to high concentrations of amine altering the pH to more alkaline conditions which is capable of unstabilising and inactivating the enzyme.

To assess the enantioselectivity of *StrepRedAm*, chiral reductive amination biotransformations were performed at carbonyl-to-amine ratios of 1:1, 1:4: 1:20 and 1:50 (**Figure 128**). For phenylbutanone (entries 4-9) and *para*-fluorophenylacetone (entries 10-13), *StrepRedAm* is (*R*)-selective. However, despite good enantioselectivity, overall conversions are moderate. Enantioselectivity for biotransformations with ethyl levulinate (entries 1-3), at a carbonyl-to-amine ratio of 1:20, resulted in higher conversions compared to the other chiral reductive amination biotransformations performed, however enantioselectivity was generally poor.

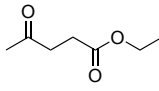
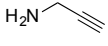
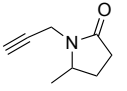
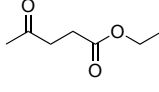

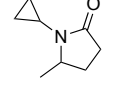
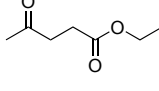
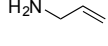
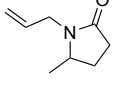
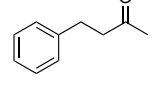
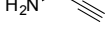
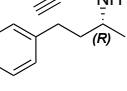
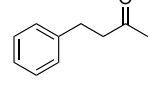
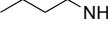
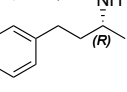
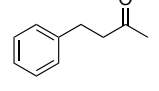

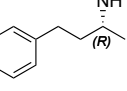
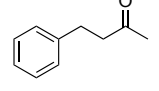

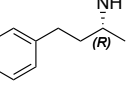
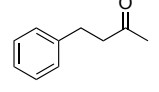

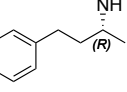
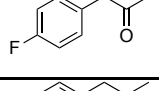

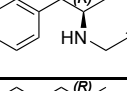
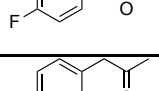
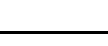
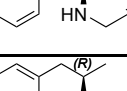
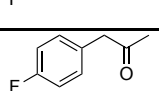
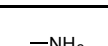
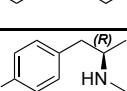
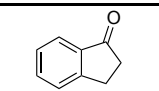

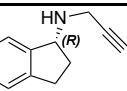
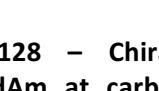
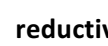
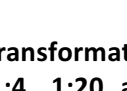
Entry	Ketone	Amine	Ketone:Amine Ratio	Product	Conversion (%)	e.e (%)
1			1:20		85	36 ^[a]
2			1:20		86	58 ^[a]
3			1:20		68	57 ^[a]
4			1:20		84	82 (R)
5			1:20		36	92 (R) ^[b]
6			1:1		<5	99 (R)
7			1:4		8	99 (R)
8			1:50		43	99 (R)
9			1:1		18	97 (R)
10			1:4		39	97 (R)
11			1:50		20	97 (R)
12			1:50		25	n.d
13			1:50		52	80 (R)

Figure 128 – Chiral reductive amination biotransformations performed with *StrepRedAm* at carbonyl-to-amine ratios of 1:1, 1:4, 1:20 and 1:50. [a] Absolute configuration not yet assigned. [b] Absolute configuration assigned based on N-methyl analogues.

Structural characterisation would be required if mutations were to be performed to improve enantioselectivity, and also to determine any structural differences between the fungal RedAms and *StrepRedAm*, the bacterial RedAm. Therefore purified *StrepRedAm* was subjected to crystal trial screening with commercially available screens. In an attempt to obtain a ternary complex,

reactions with different cofactor combinations and varying substrate ratios were added as an additive to the mother liquor of the dispensed commercially available screens. The reactions in **Figure 129** were chosen, as both selegiline (**Figure 129a**) and rasagiline (**Figure 129b**) are notable drugs for the treatment for Parkinson's disease and the reaction with cyclohexanone and propargylamine (**Figure 129c**) is a good model reaction to assess RedAm activity.

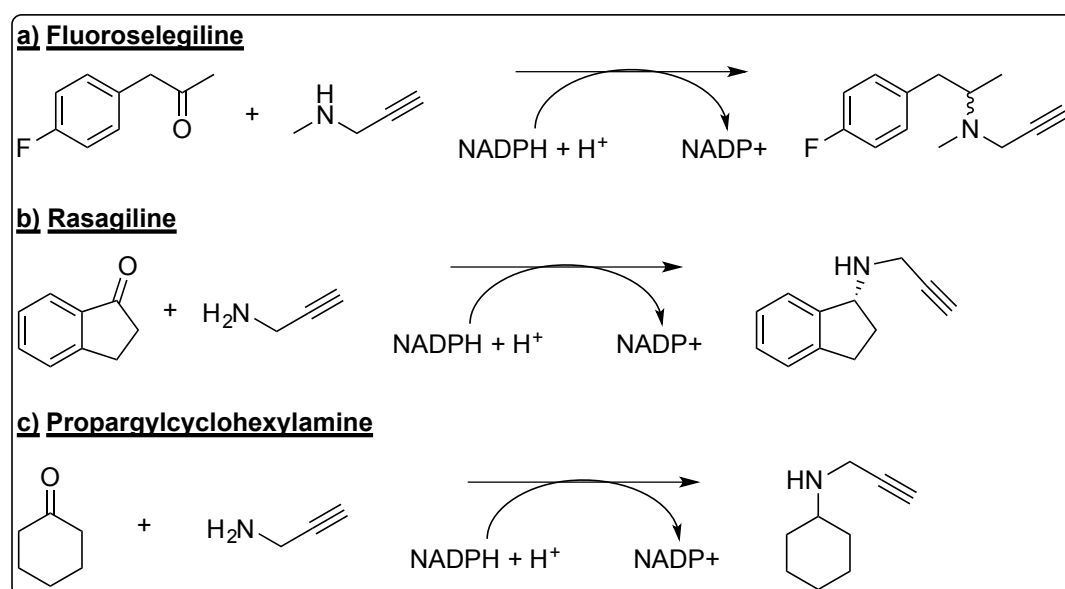


Figure 129 – Reactions attempted in the crystallisation of *StrepRedAm*, in an attempt to trap a ternary complex.

However, no crystal hits were observed from the crystal trial screening. The stability of *StrepRedAm* may be hindering the crystallisability of the enzyme.

4.4.1. Discussion

Potential IRED sequences were aligned against *AspRedAm*. Residues of interest were Asn93, Asp169 and Tyr177 (*AspRedAm*), with mutation studies performed on Asp169 and Tyr177 having already demonstrated these residues affect on activity. Mutation studies have not yet been performed on Asn93 to clarify its catalytic relevance, however its role is like to be to suppress the pK_a of Asp169. From the phylogenetic tree analysis of IREDs and RedAms recently investigated by the Grogan group with the IRED library from Roche,¹⁰³ the closest

homologues to the fungal cluster (**Figure 130**) is indeed *StrepRedAm* as well as IR21. IR9, IR19, IR21 and IR23 could also exhibit some RedAm activity.

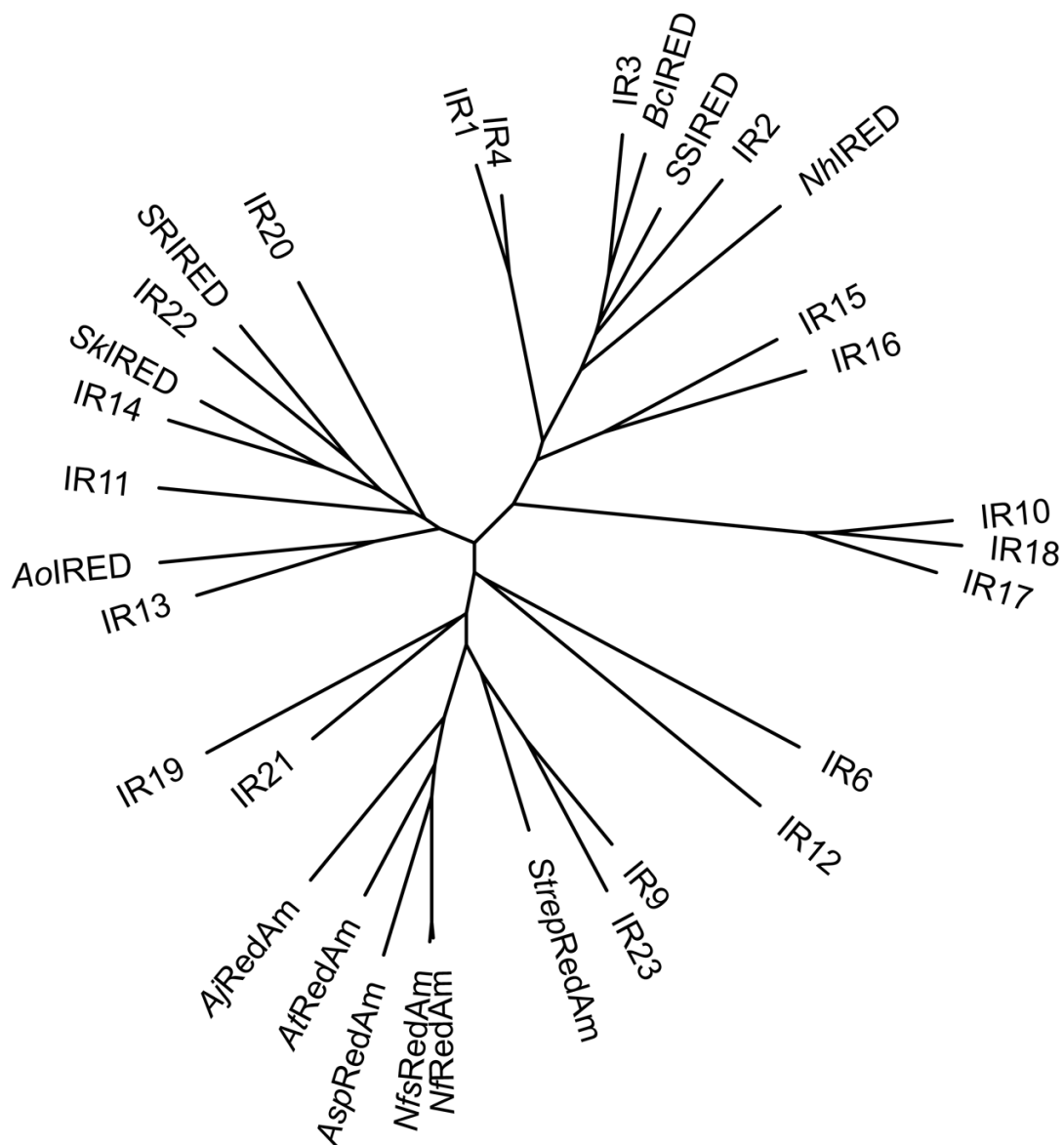


Figure 130 – The phylogenetic tree for recently investigated IREDs and RedAms, including the IRED library from Roche.¹⁰³ The phylogenetic tree analysis was performed with POWER¹⁵⁴ and displayed with iTOL.¹⁵⁵ The phylogenetic tree shows the RedAms clustered together neighbouring IREDs that may exhibit RedAm activity. IREDs further away from the RedAm cluster such as the previously investigated *SkIRED*, *SRiRED*, *BciRED*, *SSIRED* and *NhiRED* are unlikely to exhibit RedAm activity.

The purification of *StrepRedAm* was initially challenging. However this was solved with the addition of 10% (v/v) glycerol in buffers, the addition of protease inhibitors to the cell suspension prior to lysis and by collecting only the purest fractions from the nickel purification. Despite having obtained pure protein, the crystal structure for *StrepRedAm* were not obtained. This could be due to the

stability of *StrepRedAm*, in that it is a very stable protein in solution, which would hinder protein crystallisation.

RedAm activity was confirmed for *StrepRedAm* by performing biotransformations. Overall yields were moderate to good, however the final enzyme concentrations used were approximately 0.5 mg mL^{-1} , therefore to achieve higher yields, higher enzyme loadings could be used.

When comparing RedAm activity between *StrepRedAm* and *AspRedAm* (**Figure 131**), in the majority of scenarios, *AspRedAm* is the superior enzyme. *AspRedAm* has a clear advantage over *StrepRedAm* with reactions listed in entries 4 to 8 (**Figure 131**), whereas *StrepRedAm* achieves greater enantioselectivity with the reaction listed in entry 9 (**Figure 131**). Higher conversions would be obtained if higher enzyme loadings were used with the biotransformations of *StrepRedAm*. However, as the activity of *StrepRedAm* is comparable to *AspRedAm*, *StrepRedAm* could be used as a more stable substitute for reductive amination and *StrepRedAm* could be engineered to improve its activity.



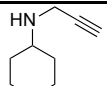


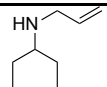


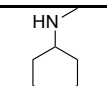

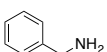
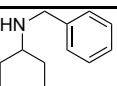


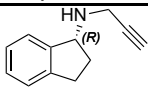
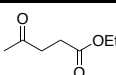

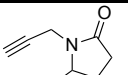
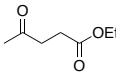

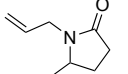
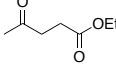

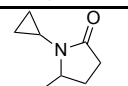
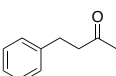

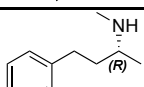
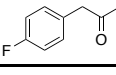

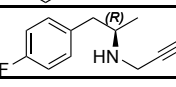
Entry	Ketone	Amine	Ketone : Amine ratio	Product	Conversion (%), e.e (%) and absolute configuration	
					<i>StrepRedAm</i>	<i>AspRedAm</i>
1			1:1		88	94
2			1:1		81	73
3			1:4		95	>99
4			1:1		19	84
5			1:50		52, 80 (R)	64, 95 (R)
6			1:20		85, 36 ^[a]	>99, 90 ^[a]
7			1:20		68, 57 ^[a]	>99, 68 ^[a]
8			1:20		86, 58 ^[a]	>99, 82 ^[a]
9			1:50		43, 99 (R)	>98, 85 (R)
10			1:50		20, 97 (R)	>97, 97 (R)

Figure 131 – Direct comparison of reductive amination biotransformations performed with *StrepRedAm* and *AspRedAm*. [a] Absolute configuration not yet assigned.

StrepRedAm has been demonstrated as the first bacterial IRED that is capable of catalysing reductive amination with a carbonyl-to-amine ratio of 1:1. Therefore *StrepRedAm* is a bacterial RedAm displaying similar activity and product enantioselectivity as *AspRedAm*. However, the major benefit of using *StrepRedAm* over *AspRedAm* is that *StrepRedAm* is less likely to precipitate during purification resulting in a greater protein yield. In most scenarios, *AspRedAm* is a more enantioselective enzyme than *StrepRedAm*, however *StrepRedAm* could have its active site mutated to create a variant that retains the stability of the wild type enzyme with the activity and enantioselectivity of *AspRedAm*.

5. Conclusion

IREDs are a family of dehydrogenase enzymes that are capable of reducing imines to afford chiral amines. A subset of IREds, the RedAms, have evolved to be capable of catalysing reductive amination, in which the RedAms catalyse the formation of the imine intermediate from carbonyl and amine substrates, which is then subsequently reduced to form the chiral amine. The major benefit of using RedAms over other enzymes capable of reductive aminations is that the RedAms have a significantly broader substrate spectrum and are not limited to ammonia as an amine donor.

RedAms are a subgroup of IREds, therefore share the the same structural fold, the IRED fold. The IRED fold is composed of an N-terminal domain composed of alternating β -sheets and α -helices forming a Rossmann fold domain for nucleotide binding, followed by a long interdomain helix which joins the C-terminal helical bundle domain, which is involved in catalysis and dimerising.

The development of the addition of ligands added as an additive to the mother liquor of commercially available screens (**3.5.2.3**) has provided an alternative method for obtaining crystal structures, where traditional native-protein screening and co-crystallisation approaches were unsuccessful. This method is more useful for obtaining biologically and catalytically relevant binary and ternary complexes. As seen with the xenobiotic ligand complexes of AoIRED, retrospectively soaking *apo*-crystals may result in crystallographic artifacts. The major benefits of this method of co-crystallising ligands as an additive over traditional approaches of ligand co-crystallisation where the ligand is added to a protein solution is that the protein solution is not harshly exposed to the protein in a large concentration which can often cause protein aggregation. Also, when the ligand is added to the dilute protein solution, once concentrated it is difficult to quantify the amount of ligand retained which can cause difficulties with reproducing the same crystals, hence, the addition of ligands as an additive allows quantification of ligand concentration and increases chances of reproducibility.

Having obtained a series of IRED and RedAm structures with a variety of ligands, the active sites and mechanisms for both IREds and RedAms have been

proposed. Along with the active sites, a potential motif for RedAms has been identified that include the residues, using residue numbers from *AspRedAm*, Asn93, Asp169 and Tyr177. From sequence and phylogenetic tree analysis, potential bacterial RedAms were identified: IR9, IR19, IR21 and IR23 from the Roche IRED library, and the RedAm from *Streptomyces* sp. PRh5, *StrepRedAm*.

Without further investigations, IR9, IR19, IR21 and IR23 are only putative RedAms. However, *StrepRedAm* has been characterised to exhibit similar RedAm activity to *AspRedAm* and, most importantly, it is capable of catalysing certain reductive aminations at a carbonyl-to-amine ratios of 1:1. Further kinetic investigations are required to fully characterise *StrepRedAm*. Despite an unsuccessful attempt at structural characterisation, it can be assumed that *StrepRedAm* will adopt the IRED fold and homology models can be built using RedAm structures as templates, which will be useful in engineering *StrepRedAm* for improved activity. Despite *AspRedAm* being the superior RedAm in terms of activity, *StrepRedAm* may still be used as a more stable replacement.

IREDs and RedAms are incredibly useful biocatalysts for the production of chiral amines. Recent large-scale biotransformations performed by France and co-workers (personal communication) have shown the potential for RedAms to be used in industrial biocatalysis. The basis of the work performed in this thesis was to provide structural information to elucidate the mechanism of IREDs and RedAms. With significant progress towards these objectives, the crystal structures can be applied for further investigations into engineering IREDs and RedAms, which can provide an alternative route for the asymmetric synthesis of chiral amines.

6. Appendix

6.1. List of Primers

Table 43 – List of primers.

Entry	Primer	Sequence (5' - 3')
1	<i>Bc</i> IRED_For	CCAGGGACCAGCAATGCTGAAAAGCAATAGCCAGAACGAAAAAATGGTA
2	<i>Bc</i> IRED_Rev	GAGGAGAAGGCGCGTTATTAGCTGCGGATATCGCTATTTTTGAAAATCTC
3	<i>Nh</i> IRED_For	CCAGGGACCAGCAATGACCAACACCAAAGCAAAAAAAGTCCGGTTACC
4	<i>Nh</i> IRED_Rev	GAGGAGAAGGCGCGTTATTAGCCTTCGGCATCCATCAGGGTCAC
5	<i>Strep</i> RedAm_For	GAAGTTCTGTTCCAGGGACCAGCAAGCGAAAGTCCGGTTACCG
6	<i>Strep</i> RedAm_Rev	TCCACCAGTCATGCTAGCCATATGTTATGCTGTACCCGGACGTTTC

Abbreviations

A_{280} - absorbance at 280 nm

AdRedAm - Reductive aminase from *Ajellomyces dermatitidis*

AnRedAm - Reductive aminase from *Aspergillus nidulans*

AoIRED - Imine reductase from *Amycolatopsis orientalis*

AspRedAm - Reductive aminase from *Aspergillus oryzae*

AtRedAm - Reductive aminase from *Aspergillus terreus*

bp - Base pair

BclIRED - Imine reductase from *Bacillus cereus* BAG3X2

BT1 - biopterin transporter

Bu - Butyl

CIP – Cahn-Ingold-Prelog

DAM - deoxyadenosine methylase

DCM - dichloromethane

DF_c - F_c weighted with the Sigma-A weighting factor (D)

ΔG - Gibbs free energy

DHFR - dihydrofolate reductase

DMF - dimethylformamide

DMSO - dimethylsulfoxide

DNA - Deoxyribonucleic acid

dsDNA - Double stranded deoxyribonucleic acid

ϵ - Molar absorption extinction coefficient

E. coli – Escherichia coli

Et - Ethyl

$[E]_T$ - Catalyst site concentration

FAD - Flavin adenine dinucleotide

F_c - Calculated structure factors

FMN - Flavin mononucleotide

F_o - Observed structure factors

FPLC -Fast protein liquid chromatography

FT1 - Folate transporter

g - Gram
g - Gravity
GC - Gas chromatography
h - Hours
HIBDH - Hydroxyisobutyrate dehydrogenase
6His - Hexa-histadine
HMWP - High-molecular-weight protein
HPLC - High performance liquid chromatography
i - Iso
IPTG - Isopropyl- β -D-1-thiogalactopyranoside
IRED - Imine Reductase
kbp - Kilobase pair
 k_{cat} - Turnover number
kg - Kilogram
 K_M - Michaelis-Menten constant
L - Litre
LCP - Lipid cubic phase
LIC - Ligation independent cloning
M - Molar
m - Meta substituted
MCS - Multiple cloning site
MDQ - 1-Methyl-3,4-dihydroisoquinoline
Me - Methyl
 $mF_o - F_o$ weighted with the Figure of Merit (*m*)
2MPN - 2-Methyl-1-pyrroline
min - Minutes
mL - Millilitre
mM - Millimolar
MS - Mass spectrometry
MTQ - 1-Methyl-1,2,3,4-tetrahydroisoquinoline
NADH - Nicotinamide adenine dinucleotide
NADPH - Nicotinamide adenine dinucleotide phosphate
NDSB - Non-detergents sulfobetaines
NfRedAm - Reductive aminase from *Neosartorya fumigata*

NhIRED - Imine reductase from *Nocardiopsis halophila*

Ni - Nickel

nL - Nanolitre

nM - Nanomolar

NMR - Nuclear magnetic resonance

o - Ortho substituted

OBDG - *n*-octyl- β -D-glucoside

OD₆₀₀ - optical density at 600 nm

OpDHs - Opine dehydrogenases

p - Para substituted

PCR - polymerase chain reaction

PDB - protein data bank

PEG - polyethyleneglycol

PLP - Pyridoxal phosphate

Pr - Propyl

PTR1 - Pteridine reductase

RedAm - Reductive aminase

RMSD - Root mean squared deviation

RSA - Retrosynthetic analysis

s - Seconds

S. bayanus - *Saccharomyces bayanus*

SDM - Site directed mutagenesis

SDS-PAGE - Sodium dodecyl sulphate polyacrylamide
electrophoresis

SEC - Size exclusion chromatography

SkIRED - Imine reductase from *Streptomyces kanamyceticus*

SLIC - Sequence and ligation independent cloning

SPINE - Structural proteomics in Europe

SRiRED - (R)-selective imine reductase from *Streptomyces* sp.

GF3587

ssDNA - Single stranded deoxyribonucleic acid

SSiRED - (S)-selective imine reductase from *Streptomyces* sp.

GF3546

SSM – Secondary structure matching

STORR - (S)- to (R)-Reticuline

StrepRedAm - Reductive aminase from *Streptomyces* sp. PRh5

T_a - annealing temperature

TD-PCR - Touchdown polymerase chain reaction

THBP - Thyroid-hormone-binding protein

TLC - Thin layer chromatography

T_m - melting temperature

μL - microlitre

μM - micromolar

V_{max} - maximum velocity of the reaction (substrate saturation)

(v/v) - volume per volume

(w/v) - weight per volume

References

1. Noyori, R.; Ohkuma, T.; Kitamura, M.; Takaya, H.; Sayo, N.; Kumobayashi, H.; Akutagawa, S., Asymmetric hydrogenation of .beta.-keto carboxylic esters. A practical, purely chemical access to .beta.-hydroxy esters in high enantiomeric purity. *Journal of the American Chemical Society* **1987**, *109* (19), 5856-5858.
2. Knowles, W. S.; Sabacky, M. J.; Vineyard, B. D., Catalytic asymmetric hydrogenation. *Annals of the New York Academy of Sciences* **1973**, *214*, 119-24.
3. Katsuki, T.; Sharpless, K. B., The first practical method for asymmetric epoxidation. *Journal of the American Chemical Society* **1980**, *102* (18), 5974-5976.
4. Funk, T. W.; Berlin, J. M.; Grubbs, R. H., Highly active chiral ruthenium catalysts for asymmetric ring-closing olefin metathesis. *Journal of the American Chemical Society* **2006**, *128* (6), 1840-1846.
5. Uematsu, N.; Fujii, A.; Hashiguchi, S.; Ikariya, T.; Noyori, R., Asymmetric transfer hydrogenation of imines. *Journal of the American Chemical Society* **1996**, *118* (20), 4916-4917.
6. Nolin, K. A.; Ahn, R. W.; Kobayashi, Y.; Kennedy-Smith, J. J.; Toste, F. D., Enantioselective Reduction of Ketones and Imines Catalyzed by (CN-Box)Re(V)-Oxo Complexes. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2010**, *16* (31), 9555-9562.
7. Yang, X.; Toste, F. D., Direct Asymmetric Amination of α -Branched Cyclic Ketones Catalyzed by a Chiral Phosphoric Acid. *Journal of the American Chemical Society* **2015**, *137* (9), 3205-3208.
8. Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D., Reductive Amination of Aldehydes and Ketones with Sodium Triacetoxyborohydride. Studies on Direct and Indirect Reductive Amination Procedures1. *The Journal of Organic Chemistry* **1996**, *61* (11), 3849-3862.
9. List, B., Proline-catalyzed asymmetric reactions. *Tetrahedron* **2002**, *58* (28), 5573-5590.
10. Halland, N.; Hansen, T.; Jorgensen, K. A., Organocatalytic asymmetric Michael reaction of cyclic 1,3-dicarbonyl compounds and alpha,beta-unsaturated ketones - A highly atom-economic catalytic one-step formation of optically active warfarin anticoagulant. *Angew Chem Int Edit* **2003**, *42* (40), 4955-4957.
11. Iwasaki, F.; Onomura, O.; Mishima, K.; Kanematsu, T.; Maki, T.; Matsumura, Y., First chemo- and stereoselective reduction of imines using trichlorosilane activated with N-formylpyrrolidine derivatives. *Tetrahedron letters* **2001**, *42* (13), 2525-2527.
12. Kanemitsu, T.; Umehara, A.; Haneji, R.; Nagata, K.; Itoh, T., A simple proline-based organocatalyst for the enantioselective reduction of imines using trichlorosilane as a reductant. *Tetrahedron* **2012**, *68* (20), 3893-3898.
13. Friedman, D., Masciangioli, T., Olson, S., Replacing Critical Materials with Abundant Materials. In *The Role of the Chemical Sciences in Finding Alternatives to*

Critical Resources: A Workshop Summary., Washington (DC): National Academies Press (US): 2012; pp 21-28.

14. Lee, S.; MacMillan, D. W. C., Organocatalytic Vinyl and Friedel–Crafts Alkylations with Trifluoroborate Salts. *Journal of the American Chemical Society* **2007**, *129* (50), 15438-15439.
15. Faber, K., *Biotransformations in Organic Chemistry*. Springer: 1992.
16. Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J., Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **2010**, *329* (5989), 305-309.
17. Pierucci, O., Chromosome Replication and Cell-Division in Escherichia-Coli at Various Temperatures of Growth. *Journal of Bacteriology* **1972**, *109* (2), 848-&.
18. Dalby, P. A., Strategy and success for the directed evolution of enzymes. *Curr Opin Struc Biol* **2011**, *21* (4), 473-480.
19. Höhne, M.; Bornscheuer, U. T., Biocatalytic Routes to Optically Active Amines. *ChemCatChem* **2009**, *1* (1), 42-51.
20. Leipold, F.; Hussain, S.; Ghislieri, D.; Turner, N. J., Asymmetric Reduction of Cyclic Imines Catalyzed by a Whole-Cell Biocatalyst Containing an (S)-Imine Reductase. *Chemcatchem* **2013**, *5* (12), 3505-3508.
21. Hernaiz, M. J.; Sanchezmontero, J. M.; Sinisterra, J. V., Comparison of the Enzymatic-Activity of Commercial and Semipurified Lipase of Candida-Cylindracea in the Hydrolysis of the Esters of (R,S) 2-Aryl Propionic Acids. *Tetrahedron* **1994**, *50* (36), 10749-10760.
22. Uenishi, J.; Hiraoka, T.; Hata, S.; Nishiwaki, K.; Yonemitsu, O.; Nakamura, K.; Tsukube, H., Chiral Pyridines: Optical resolution of 1-(2-pyridyl)- and 1-[6-(2,2'-bipyridyl)]ethanols by lipase-catalyzed enantioselective acetylation. *J Org Chem* **1998**, *63* (8), 2481-2487.
23. Gonzalez-Sabin, J.; Gotor, V.; Rebolledo, F., CAL-B-catalyzed resolution of some pharmacologically interesting beta-substituted isopropylamines. *Tetrahedron-Asymmetr* **2002**, *13* (12), 1315-1320.
24. Takayama, S.; Lee, S. T.; Hung, S. C.; Wong, C. H., Designing enzymatic resolution of amines. *Chemical communications* **1999**, (2), 127-128.
25. Shimomura, K.; Harami, H.; Matsubara, Y.; Nokami, T.; Katada, N.; Itoh, T., Lipase-mediated dynamic kinetic resolution (DKR) of secondary alcohols in the presence of zeolite using an ionic liquid solvent system. *Catal Today* **2015**, *255*, 41-48.
26. Karl, U.; Simon, A., BASF's ChiPros (R) chiral building blocks The cornerstones of your API syntheses! *Chim Oggi* **2009**, *27* (5), 66-69.
27. Turner, N. J. F., I.; Speight, R., Novel biocatalyst technology for the preparation of chiral amines. *Innovations in Pharmaceutical Technology* **2004**, *4*, 114-116.
28. Carr, R.; Alexeeva, M.; Enright, A.; Eve, T. S. C.; Dawson, M. J.; Turner, N. J., Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. *Angew Chem Int Edit* **2003**, *42* (39), 4807-4810.

29. Carr, R.; Alexeeva, M.; Dawson, M. J.; Gotor-Fernandez, V.; Humphrey, C. E.; Turner, N. J., Directed evolution of an amine oxidase for the preparative deracemisation of cyclic secondary amines. *Chembiochem* **2005**, *6* (4), 637-639.
30. Ghislieri, D.; Houghton, D.; Green, A. P.; Willies, S. C.; Turner, N. J., Monoamine Oxidase (MAO-N) Catalyzed Deracemization of Tetrahydro-beta-carbolines: Substrate Dependent Switch in Enantioselectivity. *Acs Catalysis* **2013**, *3* (12), 2869-2872.
31. Li, T.; Liang, J.; Ambrogelly, A.; Brennan, T.; Gloor, G.; Huisman, G.; Lalonde, J.; Lekhal, A.; Mijts, B.; Muley, S.; Newman, L.; Tobin, M.; Wong, G.; Zaks, A.; Zhang, X. Y., Efficient, Chemoenzymatic Process for Manufacture of the Boceprevir Bicyclic [3.1.0]Proline Intermediate Based on Amine Oxidase-Catalyzed Desymmetrization. *Journal of the American Chemical Society* **2012**, *134* (14), 6467-6472.
32. Grunwald, P., *Industrial Biocatalysis*. CRC Press: 2011; Vol. 1.
33. Hwang, B. Y.; Cho, B. K.; Yun, H.; Koteswar, K.; Kim, B. G., Revisit of aminotransferase in the genomic era and its application to biocatalysis. *J Mol Catal B-Enzym* **2005**, *37* (1-6), 47-55.
34. Stewart, J. D., Dehydrogenases and transaminases in asymmetric synthesis. *Curr Opin Chem Biol* **2001**, *5* (2), 120-129.
35. Koszelewski, D.; Tauber, K.; Faber, K.; Kroutil, W., omega-Transaminases for the synthesis of non-racemic alpha-chiral primary amines. *Trends Biotechnol* **2010**, *28* (6), 324-332.
36. Chassin, C., A biotechnological process for the production of nicotinamide. *Chim Oggi* **1996**, *14* (1-2), 9-12.
37. Meyer, H. P.; Robins, K. T., Large scale bioprocess for the production of optically pure L-carnitine. *Monatsh Chem* **2005**, *136* (8), 1269-1277.
38. Hohne, M.; Kuhl, S.; Karen, R.; Bornscheuer, U. T., Efficient asymmetric synthesis of chiral Amines by combining transaminase and pyruvate decarboxylase. *Chembiochem* **2008**, *9* (3), 363-365.
39. Abrahamson, M. J.; Vázquez-Figueroa, E.; Woodall, N. B.; Moore, J. C.; Bommaris, A. S., Development of an Amine Dehydrogenase for Synthesis of Chiral Amines. *Angewandte Chemie International Edition* **2012**, *51* (16), 3969-3972.
40. Abrahamson, M. J.; Wong, J. W.; Bommaris, A. S., The Evolution of an Amine Dehydrogenase Biocatalyst for the Asymmetric Production of Chiral Amines. *Advanced Synthesis & Catalysis* **2013**, *355* (9), 1780-1786.
41. Ye, L. J.; Toh, H. H.; Yang, Y.; Adams, J. P.; Snajdrova, R.; Li, Z., Engineering of Amine Dehydrogenase for Asymmetric Reductive Amination of Ketone by Evolving Rhodococcus Phenylalanine Dehydrogenase. *ACS Catalysis* **2015**, *5* (2), 1119-1122.
42. Itoh, N.; Yachi, C.; Kudome, T., Determining a novel NAD⁺-dependent amine dehydrogenase with a broad substrate range from *Streptomyces virginiae* IFO 12827: purification and characterization. *Journal of Molecular Catalysis B: Enzymatic* **2000**, *10* (1-3), 281-290.

43. Hussain, S.; Leipold, F.; Man, H.; Wells, E.; France, S. P.; Mulholland, K. R.; Grogan, G.; Turner, N. J., An (R)-Imine Reductase Biocatalyst for the Asymmetric Reduction of Cyclic Imines. *ChemCatChem* **2015**, *7* (4), 579-583.
44. Schwizer, F.; Köhler, V.; Dürrenberger, M.; Knörr, L.; Ward, T. R., Genetic Optimization of the Catalytic Efficiency of Artificial Imine Reductases Based on Biotin–Streptavidin Technology. *ACS Catalysis* **2013**, *3* (8), 1752-1755.
45. Muñoz Robles, V.; Vidossich, P.; Lledós, A.; Ward, T. R.; Maréchal, J.-D., Computational Insights on an Artificial Imine Reductase Based on the Biotin–Streptavidin Technology. *ACS Catalysis* **2014**, *4* (3), 833-842.
46. Robles, V. M.; Dürrenberger, M.; Heinisch, T.; Lledós, A.; Schirmer, T.; Ward, T. R.; Maréchal, J.-D., Structural, Kinetic, and Docking Studies of Artificial Imine Reductases Based on Biotin–Streptavidin Technology: An Induced Lock-and-Key Hypothesis. *Journal of the American Chemical Society* **2014**, *136* (44), 15676-15683.
47. Quinto, T.; Schwizer, F.; Zimbron, J. M.; Morina, A.; Köhler, V.; Ward, T. R., Expanding the Chemical Diversity in Artificial Imine Reductases Based on the Biotin–Streptavidin Technology. *ChemCatChem* **2014**, *6* (4), 1010-1014.
48. Puvvada, M. S.; Forrow, S. A.; Hartley, J. A.; Stephenson, P.; Gibson, I.; Jenkins, T. C.; Thurston, D. E., Inhibition of bacteriophage T7 RNA polymerase in vitro transcription by DNA-binding pyrrolo[2,1-c][1,4]benzodiazepines. *Biochemistry* **1997**, *36* (9), 2478-2484.
49. Li, W.; Chou, S. C.; Khullar, A.; Gerratana, B., Cloning and Characterization of the Biosynthetic Gene Cluster for Tomaymycin, an SJG-136 Monomeric Analog. *Appl Environ Microb* **2009**, *75* (9), 2958-2963.
50. Taylor, M.; Scott, C.; Grogan, G., F-420-dependent enzymes - potential for applications in biotechnology. *Trends Biotechnol* **2013**, *31* (2), 63-64.
51. Pol, A.; Vanderdrift, C.; Vogels, G. D.; Cuppen, T. J. H. M.; Laarhoven, W. H., Comparison of Coenzyme F-420 from Methanobacterium-Bryantii with 7-Hydroxy-10-Methyl-5-Deazaisoalloxazine and 8-Hydroxy-10-Methyl-5-Deazaisoalloxazine. *Biochem Bioph Res Co* **1980**, *92* (1), 255-260.
52. Dimarco, A. A.; Bobik, T. A.; Wolfe, R. S., Unusual Coenzymes of Methanogenesis. *Annu Rev Biochem* **1990**, *59*, 355-394.
53. Draper, R. D.; Ingraham, L. L., A Potentiometric Study of Flavin Semiquinone Equilibrium. *Arch Biochem Biophys* **1968**, *125* (3), 802-&.
54. Finn, R. D.; Basran, J.; Roitel, O.; Wolf, C. R.; Munro, A. W.; Paine, M. J. I.; Scrutton, N. S., Determination of the redox potentials and electron transfer properties of the FAD- and FMN-binding domains of the human oxidoreductase NR1. *Eur J Biochem* **2003**, *270* (6), 1164-1175.
55. Bhattacharyya, S.; Stankovich, M. T.; Truhlar, D. G.; Gao, J. L., Combined quantum mechanical and molecular mechanical simulations of one- and two-electron reduction potentials of flavin cofactor in water, medium-chain acyl-CoA dehydrogenase, and cholesterol oxidase. *J Phys Chem A* **2007**, *111* (26), 5729-5742.

56. Jacobson, F.; Walsh, C., Properties of 7,8-Didemethyl-8-Hydroxy-5-Deazaflavins Relevant to Redox Coenzyme Function in Methanogen Metabolism. *Biochemistry* **1984**, *23* (5), 979-988.
57. Walsh, C., Naturally-Occurring 5-Deazaflavin Coenzymes - Biological Redox Roles. *Accounts of chemical research* **1986**, *19* (7), 216-221.
58. Li, W.; Khullar, A.; Chou, S.; Sacramo, A.; Gerratana, B., Biosynthesis of Sibiromycin, a Potent Antitumor Antibiotic. *Appl Environ Microb* **2009**, *75* (9), 2869-2878.
59. Lapalikar, G. V.; Taylor, M. C.; Warden, A. C.; Onagi, H.; Hennessy, J. E.; Mulder, R. J.; Scott, C.; Brown, S. E.; Russell, R. J.; Easton, C. J.; Oakeshott, J. G., Cofactor promiscuity among F-420-dependent reductases enables them to catalyse both oxidation and reduction of the same substrate. *Catal Sci Technol* **2012**, *2* (8), 1560-1567.
60. Bloemendal, H.; de Jong, W.; Jaenicke, R.; Lubsen, N. H.; Slingsby, C.; Tardieu, A., Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Bio* **2004**, *86* (3), 407-485.
61. Wistow, G., The human crystallin gene families. *Hum Genomics* **2012**, *6*.
62. Hallen, A.; Cooper, A. J. L.; Jamie, J. F.; Karuso, P., Insights into Enzyme Catalysis and Thyroid Hormone Regulation of Cerebral Ketimine Reductase/mu-Crystallin Under Physiological Conditions. *Neurochem Res* **2015**, *40* (6), 1252-1266.
63. Francelle, L.; Galvan, L.; Gaillard, M. C.; Guillermier, M.; Houitte, D.; Bonvento, G.; Petit, F.; Jan, C.; Dufour, N.; Hantraye, P.; Elalouf, J. M.; De Chaldee, M.; Deglon, N.; Brouillet, E., Loss of the thyroid hormone-binding protein Crym renders striatal neurons more vulnerable to mutant huntingtin in Huntington's disease. *Hum Mol Genet* **2015**, *24* (6), 1563-1573.
64. Mistry, M.; Gillis, J.; Pavlidis, P., Genome-wide expression profiling of schizophrenia using a large combined cohort. *Mol Psychiatr* **2013**, *18* (2), 215-225.
65. Boehm, C.; Newrzella, D.; Herberger, S.; Schramm, N.; Eisenhardt, G.; Schenk, V.; Sonntag-Buck, V.; Sorgenfrei, O., Effects of antidepressant treatment on gene expression profile in mouse brain: cell type-specific transcription profiling using laser microdissection and microarray analysis. *Journal of neurochemistry* **2006**, *97*, 44-49.
66. Ijlst, L.; de Kromme, I.; Oostheim, W.; Wanders, R. J. A., Molecular cloning and expression of human L-pipecolate oxidase. *Biochem Biophys Res Co* **2000**, *270* (3), 1101-1105.
67. Suzuki, S.; Mori, J. I.; Hashizume, K., mu-crystallin, a NADPH-dependent T-3-binding protein in cytosol. *Trends Endocrin Met* **2007**, *18* (7), 286-289.
68. Hallen, A.; Cooper, A. J.; Smith, J. R.; Jamie, J. F.; Karuso, P., Ketimine reductase/CRYM catalyzes reductive alkylamination of alpha-keto acids, confirming its function as an imine reductase. *Amino acids* **2015**.
69. Vasil, M. L.; Ochsner, U. A., The response of *Pseudomonas aeruginosa* to iron: genetics, biochemistry and virulence. *Mol Microbiol* **1999**, *34* (3), 399-413.
70. Pollack, J. R.; Neilands, J. B., Enterobactin, an iron transport compound from *Salmonella typhimurium*. *Biochem Biophys Res Co* **1970**, *38* (5), 989-992.

71. Wendenbaum, S.; Demange, P.; Dell, A.; Meyer, J. M.; Abdallah, M. A., The Structure of Pyoverdine Pa, the Siderophore of *Pseudomonas-Aeruginosa*. *Tetrahedron letters* **1983**, 24 (44), 4877-4880.
72. Neilands, J. B., Siderophores - Structure and Function of Microbial Iron Transport Compounds. *Journal of Biological Chemistry* **1995**, 270 (45), 26723-26726.
73. Karpishin, T. B.; Raymond, K. N., The First Structural Characterization of a Metal-Enterobactin Complex: [V(enterobactin)]²⁻. *Angewandte Chemie International Edition in English* **1992**, 31 (4), 466-468.
74. Osorio, C. R.; Juiz-Rio, S.; Lemos, M. L., A siderophore biosynthesis gene cluster from the fish pathogen *Photobacterium damsela* subsp *piscicida* is structurally and functionally related to the *Yersinia* high-pathogenicity island. *Microbiol-Sgm* **2006**, 152, 3327-3341.
75. Pelludat, C.; Rakin, A.; Jacobi, C. A.; Schubert, S.; Heesemann, J., The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: Organization and siderophore-dependent regulation. *Journal of Bacteriology* **1998**, 180 (3), 538-546.
76. Meneely, K. M.; Lamb, A. L., Two structures of a thiazolanyl imine reductase from *Yersinia enterocolitica* provide insight into catalysis and binding to the nonribosomal peptide synthetase module of HMWP1. *Biochemistry* **2012**, 51 (44), 9002-13.
77. Cunningham, M. L.; Beverley, S. M., Pteridine salvage throughout the *Leishmania* infectious cycle: implications for antifolate chemotherapy. *Molecular and Biochemical Parasitology* **2001**, 113 (2), 199-213.
78. Rajagopalan, P. T. R.; Zhang, Z. Q.; McCourt, L.; Dwyer, M.; Benkovic, S. J.; Hammes, G. G., Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99 (21), 13481-13486.
79. Goodsell, D. S., The molecular perspective: Methotrexate. *Stem Cells* **1999**, 17 (5), 314-315.
80. Eguchi, T.; Oshiro, T.; Kuge, Y.; Mochida, K.; Uwajima, T., Process for producing L(-)tetrahydrofolic acid. Google Patents: 1990.
81. Huang, Y.; Liu, N.; Wu, X. R.; Chen, Y. J., Dehydrogenases/Reductases for the Synthesis of Chiral Pharmaceutical Intermediates. *Curr Org Chem* **2010**, 14 (14), 1447-1460.
82. Khor, G. K.; Uzir, M. H., *Saccharomyces cerevisiae*: a potential stereospecific reduction tool for biotransformation of mono- and sesquiterpenoids. *Yeast* **2011**, 28 (2), 93-107.
83. Rainieri, S.; Zambonelli, C.; Kaneko, Y., *Saccharomyces sensu stricto*: Systematics, genetic diversity and evolution. *J Biosci Bioeng* **2003**, 96 (1), 1-9.
84. Barnett, J. A., The Taxonomy of the Genus *Saccharomyces* Meyen Ex Rees - a Short Review for Non-Taxonomists. *Yeast* **1992**, 8 (1), 1-23.
85. Espinoza-Moraga, M.; Petta, T.; Vasquez-Vasquez, M.; Laurie, V. F.; Moraes, L. A. B.; Santos, L. S., Bioreduction of beta-carboline imines to amines

employing *Saccharomyces bayanus*. *Tetrahedron-Asymmetr* **2010**, *21* (16), 1988-1992.

86. Mirabal-Gallardo, Y.; Soriano, M. D. C.; Santos, L. S., Stereoselective bioreduction of beta-carboline imines through cell-free extracts from earthworms (*Eisenia foetida*). *Tetrahedron-Asymmetr* **2013**, *24* (8), 440-443.

87. Winzer, T.; Kern, M.; King, A. J.; Larson, T. R.; Teodor, R. I.; Donninger, S. L.; Li, Y.; Dowle, A. A.; Cartwright, J.; Bates, R.; Ashford, D.; Thomas, J.; Walker, C.; Bowser, T. A.; Graham, I. A., Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. *Science* **2015**, *349* (6245), 309-312.

88. Farrow, S. C.; Hagel, J. M.; Beaudoin, G. A. W.; Burns, D. C.; Facchini, P. J., Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. *Nat Chem Biol* **2015**, *11* (9), 728-+.

89. Galanie, S.; Thodey, K.; Trenchard, I. J.; Interrante, M. F.; Smolke, C. D., Synthetic Biology Complete Biosynthesis of Opioids in Yeast. *Science* **2015**, *349* (6252), 1095-1100.

90. Kries, H.; O'connor, S. E., Biocatalysts from alkaloid producing plants. *Curr Opin Chem Biol* **2016**, *31*, 22-30.

91. Vladimirov, I. A.; Matveeva, T. V.; Lutova, L. A., Opine biosynthesis and catabolism genes of *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. *Russ J Genet+* **2015**, *51* (2), 121-129.

92. Moore, L. W.; Chilton, W. S.; Canfield, M. L., Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. *Appl Environ Microb* **1997**, *63* (1), 201-207.

93. van Os, N.; Smits, S. H. J.; Schmitt, L.; Grieshaber, M. K., Control of D-octopine formation in scallop adductor muscle as revealed through thermodynamic studies of octopine dehydrogenase. *J Exp Biol* **2012**, *215* (9), 1515-1522.

94. Smits, S. H. J.; Meyer, T.; Mueller, A.; van Os, N.; Stoldt, M.; Willbold, D.; Schmitt, L.; Grieshaber, M. K., Insights into the Mechanism of Ligand Binding to Octopine Dehydrogenase from *Pecten maximus* by NMR and Crystallography. *Plos One* **2010**, *5* (8).

95. Kato, Y.; Yamada, H.; Asano, Y., Stereoselective synthesis of opine-type secondary amine carboxylic acids by a new enzyme opine dehydrogenase - Use of recombinant enzymes. *J Mol Catal B-Enzym* **1996**, *1* (3-6), 151-160.

96. Smitst, S. H. J.; Mueller, A.; Schmitt, L.; Grieshaber, M. K., A structural basis for substrate selectivity and stereoselectivity in octopine dehydrogenase from *Pecten maximus*. *Journal of Molecular Biology* **2008**, *381* (1), 200-211.

97. Chen, H.; Moore, J.; Collier, S. J.; Smith, D.; Nazor, J.; Hughes, G.; Janey, J.; Huisman, G.; Novick, S.; Agard, N., Engineered imine reductases and methods for the reductive amination of ketone and amine compounds. Google Patents: 2013.

98. Mitsukura, K.; Suzuki, M.; Tada, K.; Yoshida, T.; Nagasawa, T., Asymmetric synthesis of chiral cyclic amine from cyclic imine by bacterial whole-cell catalyst of enantioselective imine reductase. *Org Biomol Chem* **2010**, *8* (20), 4533-5.

99. Mitsukura, K.; Suzuki, M.; Shinoda, S.; Kuramoto, T.; Yoshida, T.; Nagasawa, T., Purification and Characterization of a Novel (<I>R</I>)-Imine Reductase from

Streptomyces sp. GF3587. *Bioscience, Biotechnology, and Biochemistry* **2011**, 75 (9), 1778-1782.

100. Mitsukura, K.; Kuramoto, T.; Yoshida, T.; Kimoto, N.; Yamamoto, H.; Nagasawa, T., A NADPH-dependent (S)-imine reductase (SIR) from *Streptomyces* sp. GF3546 for asymmetric synthesis of optically active amines: purification, characterization, gene cloning, and expression. *Appl Microbiol Biotechnol* **2013**, 97 (18), 8079-86.

101. Huber, T.; Schneider, L.; Präg, A.; Gerhardt, S.; Einsle, O.; Müller, M., Direct Reductive Amination of Ketones: Structure and Activity of S-Selective Imine Reductases from *Streptomyces*. *ChemCatChem* **2014**, 6 (8), 2248-2252.

102. Scheller, P. N.; Fademrecht, S.; Hofelzer, S.; Pleiss, J.; Leipold, F.; Turner, N. J.; Nestl, B. M.; Hauer, B., Enzyme Toolbox: Novel Enantiocomplementary Imine Reductases. *Chembiochem* **2014**, 15 (15), 2201-2204.

103. Wetzl, D.; Berrera, M.; Sandon, N.; Fishlock, D.; Ebeling, M.; Muller, M.; Hanlon, S.; Wirz, B.; Iding, H., Expanding the Imine Reductase Toolbox by Exploring the Bacterial Protein-Sequence Space. *Chembiochem* **2015**, 16 (12), 1749-56.

104. Man, H.; Wells, E.; Hussain, S.; Leipold, F.; Hart, S.; Turkenburg, J. P.; Turner, N. J.; Grogan, G., Structure, Activity and Stereoselectivity of NADPH-Dependent Oxidoreductases Catalysing the S-Selective Reduction of the Imine Substrate 2-Methylpyrroline. *ChemBioChem* **2015**, 16 (7), 1052-1059.

105. Fademrecht, S.; Scheller, P. N.; Nestl, B. M.; Hauer, B.; Pleiss, J., Identification of imine reductase-specific sequence motifs. *Proteins* **2016**, 84 (5), 600-610.

106. Li, H.; Zhang, G. X.; Li, L. M.; Ou, Y. S.; Wang, M. Y.; Li, C. X.; Zheng, G. W.; Xu, J. H., A Novel (R)-Imine Reductase from *Paenibacillus lactis* for Asymmetric Reduction of 3H-Indoles. *Chemcatchem* **2016**, 8 (4), 724-727.

107. Aleku, G. A.; Man, H.; France, S. P.; Leipold, F.; Hussain, S.; Toca-Gonzalez, L.; Marchington, R.; Hart, S.; Turkenburg, J. P.; Grogan, G.; Turner, N. J., Stereoselectivity and Structural Characterization of an Imine Reductase (IRED) from *Amycolatopsis orientalis*. *ACS Catalysis* **2016**, 6 (6), 3880-3889.

108. Rodriguez-Mata, M.; Frank, A.; Wells, E.; Leipold, F.; Turner, N. J.; Hart, S.; Turkenburg, J. P.; Grogan, G., Structure and activity of NADPH-dependent reductase Q1EQE0 from *Streptomyces kanamyceticus*, which catalyses the R-selective reduction of an imine substrate. *Chembiochem* **2013**, 14 (11), 1372-9.

109. Holm, L.; Rosenström, P., Dali server: conservation mapping in 3D. *Nucleic Acids Res* **2010**, 38 (suppl 2), W545-W549.

110. Lokanath, N. K.; Ohshima, N.; Takio, K.; Shiromizu, I.; Kuroishi, C.; Okazaki, N.; Kuramitsu, S.; Yokoyama, S.; Miyano, M.; Kunishima, N., Crystal Structure of Novel NADP-dependent 3-Hydroxyisobutyrate Dehydrogenase from *Thermus thermophilus* HB8. *Journal of Molecular Biology* **2005**, 352 (4), 905-917.

111. Novagen, *KOD Hot Start DNA Polymerase*. 2011.

112. ThermoScientific, *Pfu DNA Polymerase*. 2014.

113. Korbie, D. J.; Mattick, J. S., Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* **2008**, 3 (9), 1452-1456.

114. Stuart, D. I.; Jones, E. Y.; Wilson, K. S.; Daenke, S., SPINE: Structural Proteomics IN Europe - the best of both worlds. *Acta Crystallographica Section D* **2006**, *62* (10), ii-i.
115. Alzari, P. M.; Berglund, H.; Berrow, N. S.; Blagova, E.; Busso, D.; Cambillau, C.; Campanacci, V.; Christodoulou, E.; Eiler, S.; Fogg, M. J.; Folkers, G.; Geerlof, A.; Hart, D.; Haouz, A.; Herman, M. D.; Macieira, S.; Nordlund, P.; Perrakis, A.; Quevillon-Cheruel, S.; Tarandeu, F.; van Tilbeurgh, H.; Unger, T.; Luna-Vargas, M. P. A.; Velarde, M.; Willmanns, M.; Owens, R. J., Implementation of semi-automated cloning and prokaryotic expression screening: the impact of SPINE. *Acta Crystallographica Section D* **2006**, *62* (10), 1103-1113.
116. Fogg, M. J.; Wilkinson, A. J., Higher-throughput approaches to crystallization and crystal structure determination. *Biochem Soc Trans* **2008**, *36* (Pt 4), 771-5.
117. Zhu, B. G.; Cai, G. F.; Hall, E. O.; Freeman, G. J., In-Fusion (TM) assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *Biotechniques* **2007**, *43* (3), 356-359.
118. Sleight, S. C.; Bartley, B. A.; Lieviant, J. A.; Sauro, H. M., In-Fusion BioBrick assembly and re-engineering. *Nucleic Acids Res* **2010**, *38* (8), 2624-2636.
119. ClontechLaboratories, I., *In-Fusion® HD Cloning Kit User Manual*. 2014.
120. Marinus, M. G.; Morris, N. R., Isolation of Deoxyribonucleic Acid Methylase Mutants of Escherichia-Coli K-12. *Journal of Bacteriology* **1973**, *114* (3), 1143-1150.
121. Geier, G. E.; Modrich, P., Recognition Sequence of the Dam Methylase of Escherichia-Coli-K12 and Mode of Cleavage of Dpn-I Endonuclease. *Journal of Biological Chemistry* **1979**, *254* (4), 1408-1413.
122. D'arcy, A.; Mac Sweeney, A.; Stihle, M.; Haber, A., The advantages of using a modified microbatch method for rapid screening of protein crystallization conditions. *Acta Crystallogr D* **2003**, *59*, 396-399.
123. Simanshu, D. K.; Murthy, M. R. N., Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of propionate kinase (TdcD) from Salmonella typhimurium. *Acta Crystallogr F* **2005**, *61*, 52-55.
124. Fu, T. M.; Zhang, X. Y.; Li, L. F.; Liang, Y. H.; Su, X. D., Preparation, crystallization and preliminary X-ray analysis of the methionine synthase (MetE) from Streptococcus mutans. *Acta Crystallogr F* **2006**, *62*, 984-985.
125. Feil, S. C.; Tang, J. L.; Hansen, G.; Gorman, M. A.; Wikteliuss, E.; Stenberg, G.; Parker, M. W., Crystallization and preliminary X-ray analysis of glutathione transferases from cyanobacteria. *Acta Crystallogr F* **2009**, *65*, 475-477.
126. Newman, J.; Egan, D.; Walter, T. S.; Meged, R.; Berry, I.; Ben Jelloul, M.; Sussman, J. L.; Stuart, D. I.; Perrakis, A., Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr D Biol Crystallogr* **2005**, *61* (Pt 10), 1426-31.
127. Brzozowski, A. M.; Walton, J., Clear strategy screens for macromolecular crystallization. *J Appl Crystallogr* **2001**, *34* (2), 97-101.
128. Hassell, A. M.; An, G.; Bledsoe, R. K.; Bynum, J. M.; Carter, H. L.; Deng, S. J.; Gampe, R. T.; Grisard, T. E.; Madauss, K. P.; Nolte, R. T.; Rocque, W. J.; Wang, L.

- P.; Weaver, K. L.; Williams, S. P.; Wisely, G. B.; Xu, R.; Shewchuk, L. M., Crystallization of protein-ligand complexes. *Acta Crystallogr D* **2007**, *63*, 72-79.
129. Senda, M.; Hayashi, T.; Hatakeyama, M.; Takeuchi, K.; Sasaki, A. T.; Senda, T., Use of Multiple Cryoprotectants to Improve Diffraction Quality from Protein Crystals. *Cryst Growth Des* **2016**, *16* (3), 1565-1571.
130. Nave, C.; Garman, E. F., Towards an understanding of radiation damage in cryocooled macromolecular crystals. *Journal of Synchrotron Radiation* **2005**, *12* (3), 257-260.
131. Ravelli, R. B. G.; McSweeney, S. M., The 'fingerprint' that X-rays can leave on structures. *Struct Fold Des* **2000**, *8* (3), 315-328.
132. Weik, M.; Ravelli, R. B. G.; Kryger, G.; McSweeney, S.; Raves, M. L.; Harel, M.; Gros, P.; Silman, I.; Kroon, J.; Sussman, J. L., Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97* (2), 623-628.
133. Murray, J.; Garman, E., Investigation of possible free-radical scavengers and metrics for radiation damage in protein cryocrystallography Presented at the 'Second International Workshop on Radiation Damage to Crystalline Biological Samples' held at Advanced Photon Source, Chicago, USA, in December 2001. *Journal of Synchrotron Radiation* **2002**, *9* (6), 347-354.
134. Ravelli, R. B. G.; Theveneau, P.; McSweeney, S.; Caffrey, M., Unit-cell volume change as a metric of radiation damage in crystals of macromolecules Presented at the 'Second International Workshop on Radiation Damage to Crystalline Biological Samples' held at Advanced Photon Source, Chicago, USA, in December 2001. *Journal of Synchrotron Radiation* **2002**, *9* (6), 355-360.
135. Flot, D.; Mairs, T.; Giraud, T.; Guijarro, M.; Lesourd, M.; Rey, V.; van Brussel, D.; Morawe, C.; Borel, C.; Hignette, O.; Chavanne, J.; Nurizzo, D.; McSweeney, S.; Mitchell, E., The ID23-2 structural biology microfocus beamline at the ESRF. *Journal of Synchrotron Radiation* **2010**, *17* (1), 107-118.
136. Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr* **2010**, *43* (1), 186-190.
137. Arvai, A., *Adxv - A Program to Display X-ray Diffraction Images*. <http://www.scripps.edu/tainer/arvai/adxv.html>, 2013; p <http://www.scripps.edu/tainer/arvai/adxv.html>.
138. Powell, H. R., The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr D* **1999**, *55*, 1690-1695.
139. Leslie, A. G. W.; Powell, H. R., Processing diffraction data with MOSFLM. *Nato Sci Ser II Math* **2007**, *245*, 41-51.
140. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D* **2011**, *67*, 271-281.
141. Kabsch, W., XDS. *Acta Crystallographica Section D* **2010**, *66* (2), 125-132.
142. Evans, P. R.; Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr D* **2013**, *69*, 1204-1214.

143. Vagin, A.; Teplyakov, A., Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 1), 22-5.
144. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J Appl Crystallogr* **2007**, *40*, 658-674.
145. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (Pt 12 Pt 1), 2126-32.
146. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **1997**, *53* (3), 240-255.
147. Schuttelkopf, A. W.; van Aalten, D. M., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (Pt 8), 1355-63.
148. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M., Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Crystallogr* **1993**, *26*, 283-291.
149. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W. Z.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **2011**, *7*.
150. Joosten, R. P.; Beek, T. A. H. T.; Krieger, E.; Hekkelman, M. L.; Hooft, R. W. W.; Schneider, R.; Sander, C.; Vriend, G., A series of PDB related databases for everyday needs. *Nucleic Acids Res* **2011**, *39*, D411-D419.
151. Kabsch, W.; Sander, C., Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577-2637.
152. Bond, C. S.; Schuttelkopf, A. W., ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D* **2009**, *65*, 510-512.
153. Scheller, P. N.; Lenz, M.; Hammer, S. C.; Hauer, B.; Nestl, B. M., Imine Reductase-Catalyzed Intermolecular Reductive Amination of Aldehydes and Ketones. *Chemcatchem* **2015**, *7* (20), 3239-3242.
154. Lin, C.-Y.; Lin, F.-K.; Lin, C. H.; Lai, L.-W.; Hsu, H.-J.; Chen, S.-H.; Hsiung, C. A., POWER: Phylogenetic WEb Repeater—an integrated and user-optimized framework for biomolecular phylogenetic analysis. *Nucleic Acids Res* **2005**, *33* (Web Server issue), W553-W556.
155. Letunic, I.; Bork, P., Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **2016**, *44* (W1), W242-W245.