Integrated Catalysis Opens New C-C Bond Formation Selectivity

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science and Engineering.

2017

Jonathan Latham

School of Chemistry

List of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

Å	Angstrom
α-KG	α-ketoglutarate
Ac	Acetyl
ACN	Acetonitrile
ADH	Alcohol dehydrogenase
AIBN	Azobisisobutyronitrile
AmDH	Amine dehydrogenase
APCI	Atmospheric pressure chemical ionisation
Ar	Aryl
BLAST	Basic local alignment search tool
Bu	Butyl
c	Concentration
Cat.	Catalyst
CALB	*Candida Antarctica* Lipase B
calcd	Calculated
CAR	Carboxylic acid reductase
CD	Circular dichroism
CDI	1,1'-carbonyldiimidazole
CL	Column length
CLE	Chan-Lam-Evans
CLEA	Cross-linked enzyme aggregate
CMPhos	2-[2-(dicyclohexylphosphino)phenyl]-1-methyl-1*H*-indole
CoA	Coenzyme A
COSY	Correlation spectroscopy
CuTC	Copper(I) thiophene-2-carboxylate
Cu-TMEDA	Di-μ-hydroxo-bis[(*N*,*N*,*N*′,*N*′-tetramethylethylenediamine)copper(II)] chloride
CV	Column volume
Cy	Cyclohexyl
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	*N*,*N'*-dicyclohexylcarbodiimide
dec	Decomposed
DEPT	Distortionless enhancement by polarisation transfer
DIPA	Diisopropylamine
DKR	Dynamic kinetic resolution
DMF	*N*,*N*-dimethylformamide
DMG	*N*,*N*-dimethylglycine
DMSO	Dimethyl sulfoxide
DUF	Domain of unknown function
δ	Chemical shift
E. coli	*Escherichia coli*

ee	Enantiomeric excess
EI	Electron impact
eq	Equivalent(s)
ESI	Electrospray ionisation
Et	Ethyl
FAD	Flavin adenine dinucleotide (oxidised)
$FADH_2$	Reduced flavin adenine dinucleotide
FDA	5'-deoxy-5'-fluoroadenosine
FlHal	Flavin dependent halogenase
Fre	Flavin reductase
FTIR	Fourier transform infrared spectroscopy
GDH	Glucose dehydrogenase
HILIC	Hydrophilic interaction chromatography
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HPLC	High-pressure liquid chromatography
HRMS	High-resolution mass spectrometry
hrs	Hours
HSQC	Heteronuclear single quantum correlation spectroscopy
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
IRED	Imine reductase
J	NMR scalar coupling constant
kDa	Kilodalton
KPi	Potassium phosphate
LAAD	L-amino acid deaminase
LAAO	L-amino acid oxidase
LB	Lysogeny broth
LBA	Lysogeny broth agar
LC-HRMS	Liquid-chromatography high-resolution mass spectrometry
LC-MS	Liquid-chromatography mass spectrometry
LRMS	Low-resolution mass spectrometry
m/z	Mass to charge ratio
MAO-N	Monoamine oxidase
Me	Methyl
MES	2-(*N*-morpholino)ethanesulfonic acid
MHz	Megahertz
MIDA	*N*-methyliminodiacetic acid
min	Minute(s)
mol %	Mole percent
mp	Melting point

MS	Molecular sieves
MWCO	Molecular weight cut-off
mwt	Molecular weight
N	Normality (eq/L, concentration)
N.C.	No conversion
N.D.	Not detected
NAD^+	Nicotinamide adenine dinucleotide (oxidised)
NADH	Reduced nicotinamide adenine dinucleotide
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NBS	*N*-bromo succinimide
NHase	Nitrile hydratase
Ni-NTA	Nickel nitrilotriacetic acid
NMP	*N*-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NRPS	Nonribosomal peptide synthetase
OD	Optical density
P450	Cytochrome P450
PCP	Peptidyl carrier protein
PDMS	Poly(dimethylsiloxane)
PE	Petroleum ether
PET	Positron emission tomography
Ph	Phenyl
Pin	Pinacol ester
PKS	Polyketide synthase
PLP	Pyridoxal 5'-phosphate
PNP	Purine nucleoside phosphorylase
PyNP	Pyrimidine-nucleoside phosphorylase
RCY	Radiochemical yield
Rh.	*Rhodococcus*
rpm	Revolutions per minute
rt	Room temperature
SAM	*S*-adenosyl methionine
^{s}Bu	Secondary butyl
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	Second(s)
SMC	Suzuki-Miyaura coupling
sSphos	Sodium 2′-dicyclohexylphosphino-2,6-dimethoxy-1,1′-biphenyl-3-sulfonate
STY	Space-time yield
TA	Transaminase
TB	Terrific broth

TBATB	Tetrabutylammonium tribromide
^{t}Bu	Tertiary butyl
^{t}BuXPhos	2-Di-*tert*-butylphosphino-2′,4′,6′-triisopropylbiphenyl
^{t}BuXPhos-Pd-G3	[(2-Di-*tert*-butylphosphino-2′,4′,6′-triisopropyl-1,1′-biphenyl)-2-(2′-amino-1,1′-biphenyl)] palladium(II) methanesulfonate
Tf	Trifluoromethanesulfonate (triflate)
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin-layer chromatography
T_m	Melting temperature
TMEDA	Tetramethylethylenediamine
tppts	Triphenylphosphine-3,3′,3″-trisulfonic acid trisodium salt
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
TS	Tryptophan synthase
UHPLC	Ultra-high pressure liquid chromatography
UV	Ultraviolet
V	Volts
v/v	Volume/volume concentration
W	Watts
wt/v	Weight/volume concentration
λ	Wavelength

List of Abbreviations (Amino Acids)

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Cystine-sulfenic acid	CSO	
Cystine-sulfinic acid	CSD	
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Abstract

Integrated Catalysis Opens New C-C Bond Formation Selectivity.
Jonathan Latham
The University of Manchester
Doctor of Philosophy - 2017

The combination of multiple catalytic events into chemical transformations can allow a significant amount of molecular complexity to be introduced in an efficient manner. Of particular interest is the combination of biocatalysis with chemocatalysis, since such combinations can allow the selectivity inherent to enzymes to be exploited beyond the repertoire of reactions which enzymes can directly catalyse. Although an attractive prospect, differences in operating conditions and mutual deactivation often preclude the efficient combination of bio- and chemo-catalysts.

This thesis focuses on the combination of biocatalytic halogenation reactions with palladium-catalysed cross-coupling reactions. Aryl halides are useful compounds in their own right, with a significant number of pharmaceuticals and agrochemicals containing halogen atoms, as well as being versatile synthetic intermediates for Pd-catalysed cross-coupling reactions. Traditional halogenation methods are typically fraught with poor regioselectivity, often affording mixtures, and require deleterious reagents. The flavin-dependent halogenase (FlHal) enzymes have been demonstrated to be a viable alternative – catalysing the regioselective halogenation of a range of substrates using benign inorganic halide sources. It was therefore envisioned that combination of these enzymes with Pd-catalysed reactions could allow regioselective cross-coupling reactions.

The utility of FlHals in synthesis is currently limited by poor enzyme productivity, thought to result from slow halogenation reactions combined with low enzyme stability. To address these issues, this thesis characterised a proposed FlHal from a thermophilic *Streptomyces* strain in the hope of finding more productive biocatalysts. In addition to exploring the stability and substrate scope of this enzyme, its crystal structure was also obtained and studied to rationalise the structural features which contribute to increased stability.

A number of FlHals were also combined with Pd-catalysis to allow the regioselective arylation, vinylation and heteroarylation of several aromatic scaffolds. To overcome the mutual deactivation of each of the catalytic systems, a number of compartmentalisation strategies were explored to allow efficient combination of the two reactions into single pot transformations. Notably, halogenases of different regioselectivity were used to allow the selective activation of the indole C7, C6 and C5 positions – reactions which were beyond the scope of traditional CH activation approaches at the time. This rationale was subsequently extended to allow the regioselective cyanation of the indole nucleus and, by introduction of additional catalysts, the selective amidation of the indole C7 position.

A number of attempts were also made to combine nitrile-hydrating enzymes with Cu- and Pd-catalysed C-N bond formation reactions to allow the atom economic synthesis of secondary amides and enamides.

Declaration

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

Copyright Statement

Acknowledgements

I would like to thank my supervisors, Professors Jason Micklefield and Michael Greaney, for the opportunity to work on this project in their groups as well as their intellectual input and general tutelage in Science. I am also grateful to the Faculty of Science and Engineering and the School of Chemistry for funding my time here.

I would also like to thank the entirety of the Micklefield and Greaney groups, past and present, for their help and ideas throughout the course of this work. Much of this work would not have been possible without the work, training and guidance of Drs Sarah Shepherd, Binuraj Menon and Anna Winona-Struck. Specifically, work on the thermophilic tryptophan halogenase would not have been possible without Binuraj who was instrumental in the identification and characterisation of this enzyme. Additionally much of the work on integrated catalysis would not have happened without the input and hard work of Jean-Marc Henry and Humera Sharif – who also managed to somehow brave dealing with me during various rushes towards deadlines. The contribution of the rest of Team Halogenase, Eileen Brandenburger and Heidi Fisk, was also key to much of this work. Without the work of Daniel Francis, the use of these methods for the modification of tryptophan analogues would not have been possible. I am also grateful to Dr Michael Winn, Jean-Marc Henry, Dr Sarah Shepherd and Dr Binuraj Menon for proofreading this document.

Besides specific contributions, the 2013 PhD Cohort (Dan, Eileen, Heidi and Vicki) and members of the “Cool Bay” have made this experience of Research and enjoyable one.

Finally thanks to my family and friends for support throughout this time.

Publications

Portions of the work presented herein has been published in peer-reviewed journals during the course of this PhD. These publications are attached for convenience:

- Latham, J.; Henry, J.-M.; Sharif, H. H.; Menon, B. R. K.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J. Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation. *Nature Communications* **2016,** *7*, 11873.
- Menon, B. R. K.; Latham, J.; Dunstan, M. S.; Brandenburger, E.; Klemstein, U.; Leys, D.; Karthikeyan, C.; Greaney, M. F.; Shepherd, S. A.; Micklefield, J. Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes. *Organic & Biomolecular Chemistry* **2016,** *14* (39), 9354.
- Francis, D.; Winn, M.; Latham, J.; Greaney, M. F.; Micklefield, J. An Engineered Tryptophan Synthase Opens New Enzymatic Pathways to β-Methyltryptophan and Derivatives. *ChemBioChem* **2017,** *18* (4), 382.
- Latham, J.; Brandenburger, E.; Menon, B. R.; Shepherd, S. A.; Micklefield, J. Development of Halogenase Enzymes for Use in Synthesis. *Chemical Reviews* **2017,** *ASAP*, DOI: 10.1021/acs.chemrev.7b00032.

The author also contributed to the following publications during the course of this PhD:

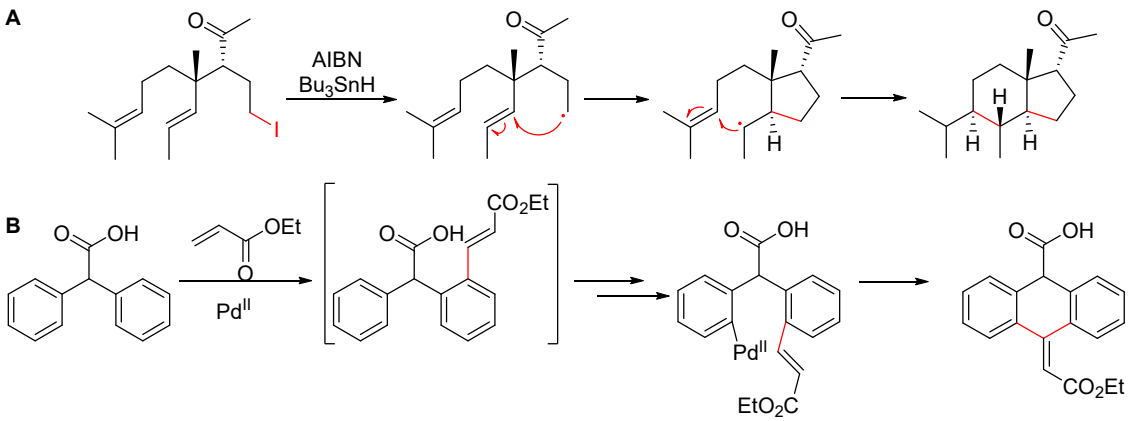
- Shepherd, S. A.; Karthikeyan, C.; Latham, J.; Struck, A. W.; Thompson, M. L.; Menon, B. R. K.; Styles, M. Q.; Levy, C.; Leys, D.; Micklefield, J. Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. *Chemical Science* **2015,** *6* (6), 3454.

1. Introduction

The combination of multiple catalysts into single transformations can allow the assembly of complex molecular scaffolds which would be inaccessible by using a single catalyst in isolation.[1-4] Conducting such reactions as single-pot transformations, whereby intermediates are not isolated or purified between reactions, represents a potential step-change in the efficiency of chemical synthesis.[3-5] Such intermediary processing reduces the overall space-time yield of a process (i.e. product per volume per time) by not only requiring additional time and reactor capacity, but oftentimes by necessitating the addition of auxiliary chemicals and organic solvents. By reducing the solvent consumption and waste of a process, in addition to the energy demand associated with the removal of solvents, progress towards the goal of green chemical synthesis can be achieved.[3,5,6]

1.1 Domino Reactions

Domino reactions, whereby multiple transformations are carried out following a single mechanism, are classified as one-pot transformations.[1] Early examples include radical domino reactions whereby a radical intermediate, after forming a single C-C bond, leads to the generation of a second radical, which then reacts in the same manner to form additional C-C bonds.[7] This approach, as exemplified in **Scheme 1A**, can allow significant complexity to be introduced from relatively simple starting materials.[8] A similar rationale is possible using transition-metal catalysed chemistry, whereby metallated intermediates are capable of catalysing subsequent bond-formations in primed substrates (**Scheme 1B**).[9]

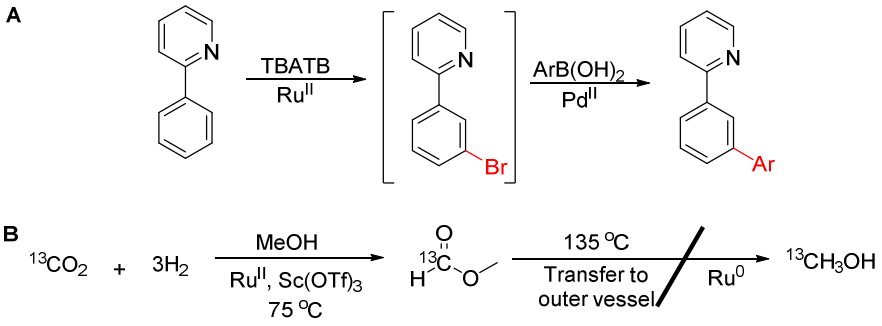


Scheme 1: Examples of radical and transition metal-catalysed domino reactions.[8,9] (AIBN = azobisisobutyonitrile)

1.2 Orthogonal Tandem Catalysis

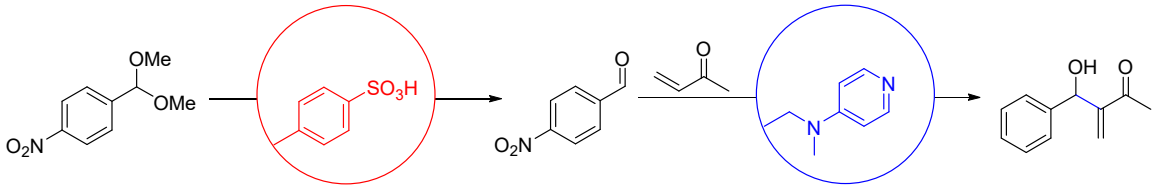
Transformations where multiple catalysts operate via distinct mechanisms are termed orthogonal tandem catalysis.[1] The combination of multiple mechanisms can afford transformations of increased complexity. Recent examples of this rationale involve the manipulation of functional groups which direct the key bond formation step.[10,11] **Scheme 2A** for example shows how one metal can be used to install a reactive functionality, which is then utilised by the second metal for subsequent reactions.[11] In many examples, the multiple catalysts used for a cascade are

incompatible, either in terms of operating conditions or through deactivation. In these cases, some form of compartmentalisation is often used.[1] For example whilst developing a process for the conversion of carbon dioxide to methanol, it was found that the Lewis acid ($Sc(OTf)_3$), used to promote esterification of formic acid, could poison the Ru^0 catalyst employed to catalyse hydrogenation of methyl formate to methanol (**Scheme 2B**).[12] Since methyl formate is significantly more volatile than any of the catalysts however, the intermediate could be transferred to a second vessel, containing the Ru^0 catalyst, by heating.[12]



Scheme 2: Examples of orthogonal tandem cascade catalysis using different combinations of transition metal catalysts.[11,12] (TBATB = tetrabutylammonium tribromide)

The compartmentalisation of incompatible reagents is oftentimes not as straightforward however, and more complex means are required. One approach is to immobilise the incompatible reagents in such a way that they can no longer directly interact with each other, but still contact substrates and intermediates in the cascade. An early example involves the specific isolation of general acid and base catalysts into separate styrene beads (**Scheme 3**).[13] By preventing direct contact of the acid and base catalysts, their mutual deactivation through salt formation was prevented – allowing the two fundamentally different catalysts operate in the same pot.[13] This idea of specifically locating or encapsulating different incompatible chemocatalysts, as heterogeneous catalysts, has been expanded to a number of transformations.[14]

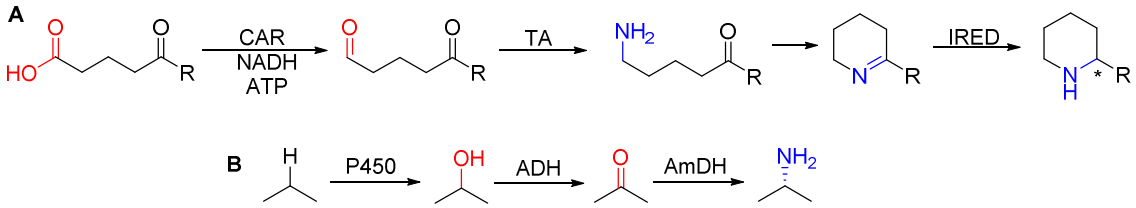


Scheme 3: Immobilisation of acid and base catalysts into polymeric beads to prevent their mutual deactivation.[13]

Biocatalysis, the use of enzymes to perform preparative chemical transformations, is an attractive prospect in synthetic chemistry.[2,15] Typically enzymes operate in aqueous media (a non-toxic and inexpensive solvent) under mild reaction conditions (thereby reducing energy consumption) and are sourced from renewable feedstocks. The stereo- and regio-selectivity which is often exhibited by enzymes is also key to their green credentials, as this can significantly improve the efficiency of synthesis by circumventing the need for tedious purification or resolution processes, in addition to

improving material efficiency.[15] The mild operating conditions and chemoselectivity of enzymes also means that protection and de-protection steps are usually not required, improving the overall atom and step economy of synthesis.[15,16] Cascades involving multiple biocatalysts have therefore been employed to allow complex, but selective, transformations in a single pot.[2]

One such example involves the combination of a carboxylic acid reductase (CAR) enzyme and a transaminase (TA) (**Scheme 4A**).[17] The tandem action of these two enzymes on a keto-acid substrate selectively produces an amino-aldehyde which can spontaneously cyclise to an imine. Chemoselectivity of the CAR is key to the production of this intermediate, as over-reduction of the acid or reaction of the ketone would preclude the required cyclisation. An imine reductase (IRED) is then employed for the stereoselective reduction of this intermediate to afford chiral piperidines and pyrrolidines.[17] Multi-enzyme cascades have also been used to afford powerful transformations such as the conversion of alcohols, or inert C-H bonds, to chiral amines.[18,19] In one such cascade an alcohol dehydrogenase (ADH) is used to produce a planar ketone, which then acts as substrate for an amine dehydrogenase (AmDH), to ultimately afford chiral amines.[18] The same strategy has also been employed to allow the transformation of benzylic C-H bonds to C-NH_2 groups stereoselectively, through the incorporation of an additional cytochrome P450 (**Scheme 4B**).[19]



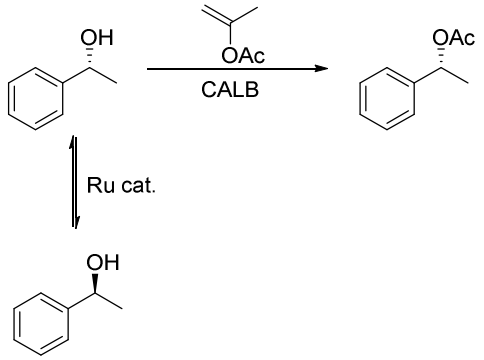
Scheme 4: Combination of multiple biocatalysts into orthogonal tandem cascade reactions to afford chiral products.[17-19] (CAR = carboxylic acid reductase, TA = transaminase, IRED = imine reductase, ADH = alcohol dehydrogenase, AmDH = amine dehydrogenase).

1.3 Integrated Catalysis

Given the ability of biocatalysis to carry out such selective transformations and the versatility of metal-catalysed reactions, it is unsurprising that numerous attempts have been made to combine these two types of catalyst into single pot transformations.[3,4] Such combinations, termed integrated catalysis, allow the selectivity of biocatalysis to be exploited beyond transformations which enzymes can directly catalyse. The earliest examples of such combinations are dynamic kinetic resolution (DKR) reactions, whereby a transition metal catalyst is used to catalyse the racemisation of starting material and an enzyme is employed to selectively convert one enantiomer of the mixture to a separable compound (**Scheme 5**).[20-23] Because of the racemisation, the maximum theoretical yield of enantiopure product is 100 %.

These examples are possible because the robust nature and cofactor-free operation of the biocatalysts in these cases allows them to function in organic solvents. The markedly different operation conditions of biocatalysis and chemocatalysis however, has limited the development of additional transformations using similar combinations. Additionally, the mutual deactivation of

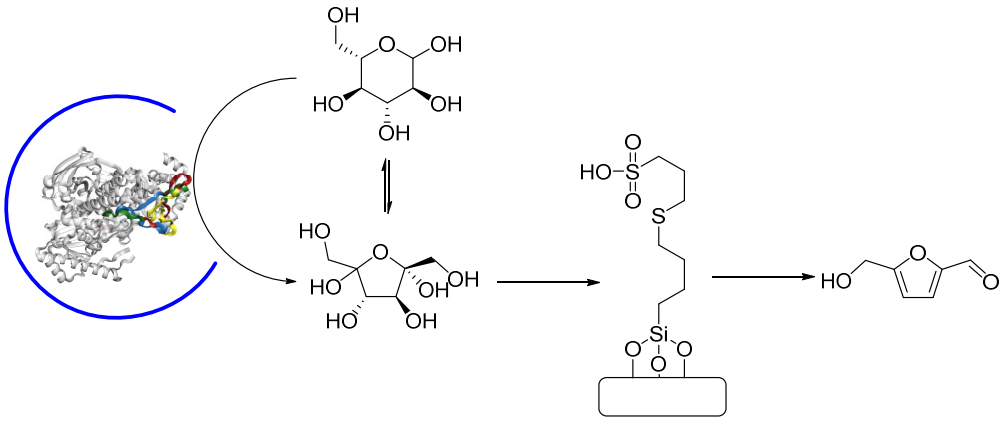
biocatalysts and chemocatalysts oftentimes means that some form of compartmentalisation or immobilisation is required for productive cascades.[3] A number of examples have therefore employed the removal of certain reagents or a change of reaction media between the two catalytic events.[24-28] Although useful transformations can be achieved, the need for intermediary processing is undesirable – consequently compartmentalisation is often used.



Scheme 5: Example of a dynamic kinetic resolution process using a ruthenium catalyst for racemisation and and *S*-selective lipase for the resolution of racemic alcohols.[22] (CALB = *Candida Antarctica* Lipase B).

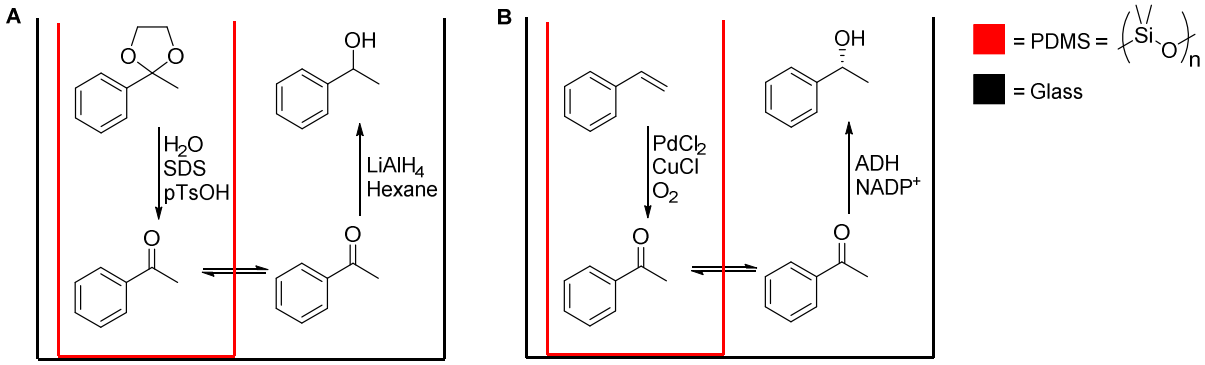
In a similar vein to the example shown in **Scheme 3**, compartmentalisation has been achieved by the immobilisation of the catalysts onto separate solid supports. In one example, a glucose isomerase is immobilised inside silica beads, which allows efficient operation of the enzyme in mixtures of water and organic solvent, whilst the acid catalyst is immobilised onto a solid support (**Scheme 6**).[29] Such an approach prevents direct contact of the enzyme and the acid, thereby preventing inhibition, which improves the practicability of this process compared to previous two-pot approaches.[30,31] Such spatial separation has also been achieved by using polymeric beads to encapsulate the biocatalyst.[32] The use of these beads allows a biocatalytic decarboxylation to occur in the enzyme's native aqueous environment, whilst suspended in a largely organic reaction medium. The size of the beads then allows filtration, prior to removal of residual water and addition of the ruthenium metathesis catalyst to its preferred organic phase.[32] Although intermediary processing is minimal, the need for filtration and removal of residual water is not ideal.

Transition metal catalysts have also been immobilised onto solid supports to allow similar cascades.[28,32,33] In a number of examples the solid-supported transition metal catalyst is immobilised inside a flow-cell, whilst products are allowed to continually flow into a second biocatalyst-containing reactor, thereby separating the chemocatalyst from the biocatalyst.[33,34] Similar examples of using laccases to generate radicals, which then re-combine with chemical reagents, did not require similar compartmentalisation due to the lack of mutual deactivation.[35] Other examples employ palladium nano-particles immobilised onto fluorous phase silica.[28] Although the palladium catalyst is immobilised in this instance, the other reagents required for the palladium-catalysed coupling mean that significant modification of the reaction is required before addition of the biocatalyst.[28]



Scheme 6: Immobilisation of a glucose isomerase and acid catalyst onto separate solid supports to allow their tandem action without inactivation of the biocatalyst.[29]

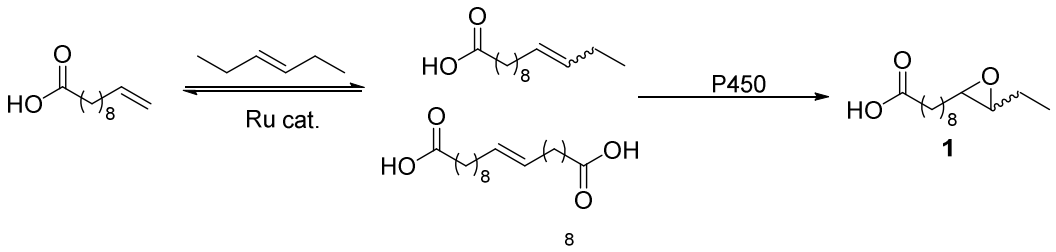
An approach to site-isolate all incompatible reagents, rather than selected immobilised compounds, is to use semi-permeable membranes.[36-39] In a number of examples poly-dimethylsiloxane (PDMS), a hydrophobic polymer,[36] is used to prevent passage of polar components between two compartments of the same reactor (**Scheme 7**).[37-39] Initially this was used to prevent the deactivation of $LiAlH_4$ by H_2O (**Scheme 7A**).[38] The polar H_2O and $LiAlH_4$ are unable to pass through the membrane, whilst intermediary ketone is sufficiently non-polar to do so. More recently, this was employed to separate the palladium catalyst required for the Wacker oxidation of styrenes from the biocatalytic asymmetric reduction of the subsequent acetophenones (**Scheme 7B**). As the $PdCl_2$ and ADH catalysts, as well as the nicotinamide cofactors, are charged they are incapable of passing through the membrane of the reactor and therefore cannot deactivate each other.[39] The non-polar intermediate can freely pass, however, allowing an effective cascade to occur.[39]



Scheme 7: Use of poly-dimethylsiloxane (PDMS) to compartmentalise incompatible reagents and therefore allow cascade reactions to occur.[38,39] (ADH = alcohol dehydrogenase, $NADP^+$ = nicotinamide adenine dinucleotide phosphate).

While many cascades exploit the stereo-selectivity of biocatalysts, others have exploited their substrate specificity to influence equilibria established by chemocatalytic reactions, in a similar vein to DKR.[40,41] In one such example (**Scheme 8**) a ruthenium-catalysed metathesis is used to establish an equilibrating mixture of cross and homo-metathesis products. The substrate specificity of the P450 then means that only a single component of this mixture is oxidised to the

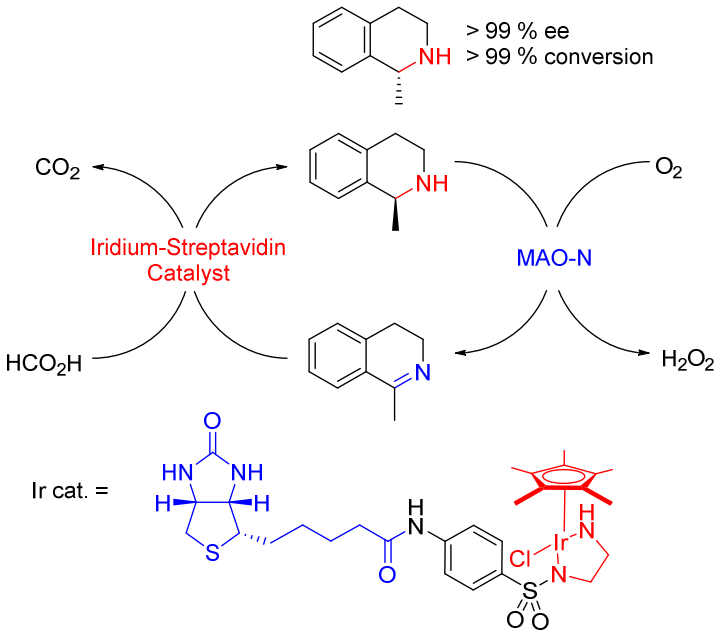
corresponding epoxide. Importantly, it was demonstrated that the cascade process afforded greater amounts of **1** than when the two reactions were carried out sequentially - demonstrating the influence of the biocatalyst's selectivity on the equilibrium.[40,41] For such an outcome, termed co-operative catalysis, it is essential that the two catalytic systems are operating in concert.[41] In these examples concurrent activity of the two catalysts is achieved by the compartmentalisation of the enzyme and ruthenium catalyst into separate aqueous and organic phases respectively.



8

Scheme 8: Co-operative catalysis using a Ruthenium-catalysed metathesis and P450-catalysed epoxidation.[41]

An elegant approach to preventing the deactivation of transition metal catalysts by biocatalytic components is to encapsulate the catalytic metal inside the binding site of a protein.[42-44] The encapsulation of metals into styrene cores has allowed metal-catalysed chemistry to occur *in vivo* without deactivation of the metal catalyst.[45] The same approach has also been employed to reduce the mutual deactivation of transition metal catalysts and biocatalysts to allow one-pot deracemisation cascades.[44] In **Scheme 9** for example, the iridium complex responsible for reduction of intermediate imine is encapsulated into streptavidin by the conjugation of a biotin adduct onto one of the ligands of the catalyst. When combined with an *S*-selective MAO-N, the deracemisation of amines was possible, with high *ee*, by eliminating mutual inactivation of the transition metal catalyst and the biocatalyst.[44]



Scheme 9: Use of a streptavidin-encapsulated iridium complex for compartmentalisation from a biocatalytic reduction (MAO-N = monoamine oxidase).

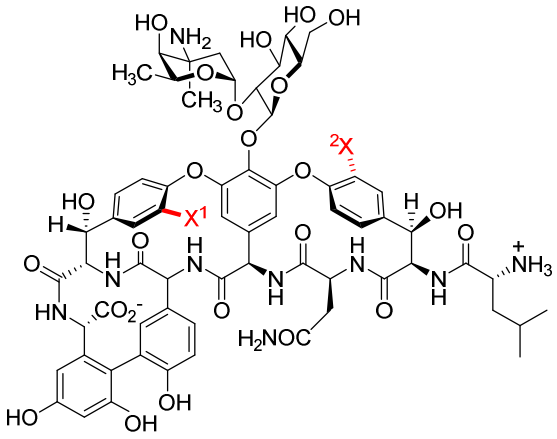
1.4 Conclusions

These examples demonstrate the powerful transformations which can be realised through the combination of multiple catalysts into single pot reactions – specifically the combination of highly selective biocatalysts with metal catalysts which catalyse reactions beyond the scope of biocatalysis alone. This thesis uses this rationale to develop new, selective, methodologies for a number of key bond formations. Chapter 2 will describe the study of a naturally-occurring thermophilic tryptophan halogenase which was employed for the regioselective halogenation of aromatic substrates containing multiple, similar, C-H bonds. Chapters 3 and 4 then utilise this regio-controlled biocatalytic halogenation to template C-C bond formations through the integration of these enzymes with palladium-catalysed cross-coupling chemistry. Finally Chapter 5 attempts to employ chemoselective nitrile hydratases for the mild hydration of nitriles, followed by the integration of transition-metal catalysed C-N bond formations, to produce enamides.

2. Halogenase-Catalysed Regioselective Halogenation

2.1 Introduction

Organohalogen compounds are of importance in all sectors of the chemical industry. Transition-metal catalysed cross-coupling reactions, for example, have become key tools in the synthesis of complex molecules because of the plethora of C-C, C-N, C-F and other C-heteroatom couplings which are possible.[46,47] The ability to metallate the C-halogen bond is key to many of these reactions, meaning that organohalogen compounds have become ubiquitous synthetic intermediates. Additionally, the effect that introducing a halogen atom can have upon the bioactivity and physicochemical properties of small molecules means that this strategy has been exploited by medicinal chemistry, resulting in a large proportion of all drugs in clinical trials or on the market containing halogen atoms (**Figure 2**).[48,49] In the case of the antibiotic vancomycin, for example, the dechloro derivatives have significantly reduced bioactivity towards their biological target than the chlorinated counterparts (**Figure 1**).[50]



Vancomycin, X^1 = Cl, X^2 = Cl, K_a = 6.4
Monochlorovancomycin, X^1 = H, X^2 = Cl, K_a = 2.4
Dechlorovancomycin, X^1 = H, X^2 = H, K_a = 1

Figure 1: Effect of halogen atoms upon bioactivity of of vancomycin. K_a values are association constants to a peptidoglycan mimic of the biological target.[50]

The effect of halogen atoms upon biological activity and bioavailability is due in part to the modulation of lipophilicity, allowing increased bioavailbility, and the formation of non-specific hydrophobic interactions with the biological target. Recently, however, it has been shown that the halogen atom in carbon-halogen bonds (C-X) can form directional intermolecular interactions to protein targets in an analogous manner to hydrogen bonds - termed halogen bonds.[48] These bonds come about because of the electron-deficient “sigma hole” of the C-X, which allows its interaction with the lone pairs of electrons on atoms such as P, S, O and N. The C-F bond is of particular interest in medicinal chemistry, with fluorine being the most prevalent halogen found in pharmaceuticals.[49] This is due in part to fluorine’s significantly increased electronegativitiy, but similar size to, hydrogen. These two factors allow the introduction of increased lipophilicity to a target compound, without significantly altering sterics. Additionally, metabolically-labile C-H bonds can be replaced with C-F bonds with minimal effect upon steric binding modes. This effect is not

limited to the development of pharmaceuticals, with a number of bestselling agrochemicals also containing halogen atoms (**Figure 2**).[51,52] Additionally, advantageous properties of organohalogens have been found in polymers, meaning that they are becoming of increased importance for the development of the next generation of materials.[53,54]

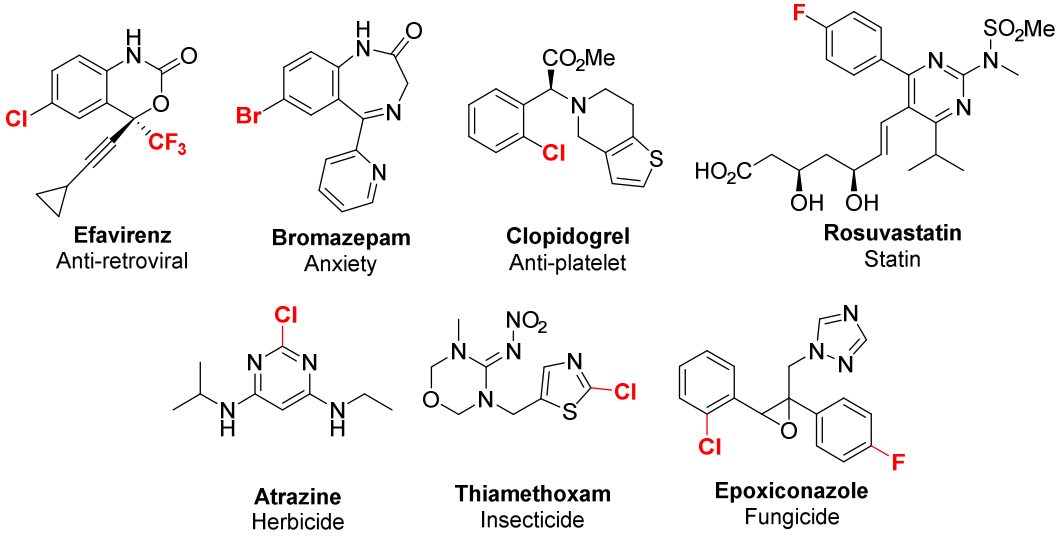
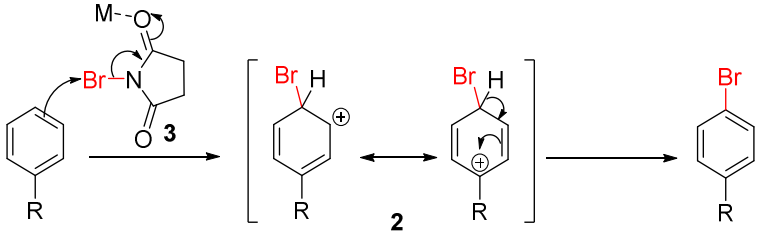


Figure 2: Halogenated compounds of medicinal and agrochemical importance.

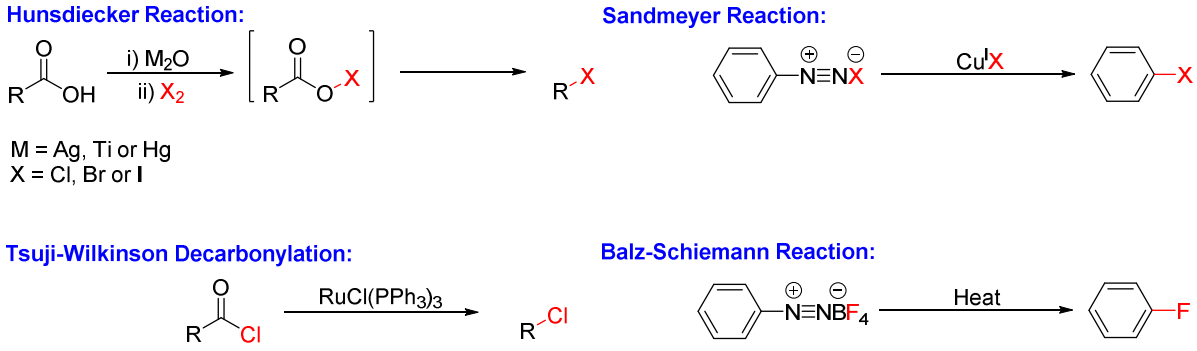
2.1.1 Traditional Halogenation Methods

The central role that halogenated compounds play in many areas of chemistry, including as synthetic intermediates, means that methods for their facile, economical, ecological and selective synthesis are required. Many of the traditional methods for such transformations fall short of these criteria, however. The archetypal method for halogenating aromatic substrates involves the generation of electrophilic halogen sources, usually from elemental halogens using a Lewis acid, which then directly interact with the π-system of substrate (**Scheme 10**).[55] The intermediate carbocation **2** is then deprotonated to restore aromaticity. More recent developments allow the use of less toxic halogen sources such as N-halo succinimides (**3**),[56,57] with Lewis acids often times employed also to increase the electrophilicity of **3**.[55,58] Regardless of which specific reagents are used, the regioselectivity of these transformations is influenced largely by the stabilisation of intermediate **2**. As such, these classical methods often times afford mixtures of regioisomers and render some regioselectivities inaccessible. The production of these unwanted regioisomers means that material efficiency of the overall process is reduced, and necessitates the separation and disposal of unwanted halogenated products, which can be toxic to and persistent in the environment.[59]



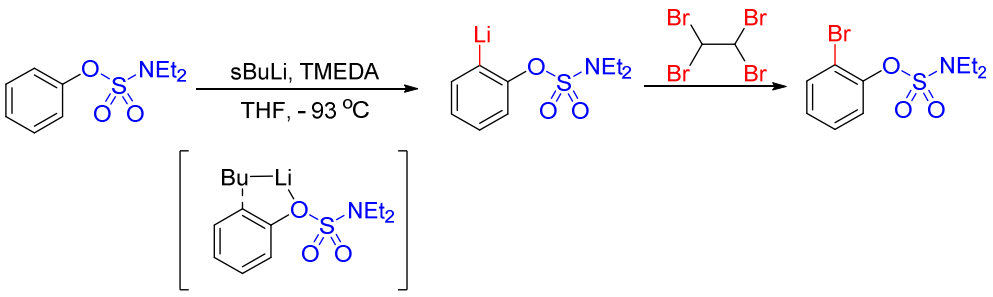
Scheme 10: Typical mechanism of electrophilic aromatic substation showing stabilisation of the Wheland intermediate (**2**) and coordination of Lewis acids to NBS (**3**).

A number of reactions are known which involve the conversion of functional groups into halogen substituents (**Figure 11**).[60] The Hunsdiecker reaction for example involves the generation of a carbon radical from a carboxylic acid, which then recombines with elemental halogen to afford alkyl halides.[61] A similar rationale is employed in the Tsuji-Wilkinson decarbonylation, whereby a ruthenium catalyst is involved in the decarbonylation and halogenation of acyl halides to halides.[62] The Sandmeyer and Balz-Schiemann reactions involve the conversion of alkyl or aryl amines into the associated halides via intermediate diazonium halides or tetrafluoroborates.[63,64] In each of these examples, a functional group must first be installed to act as the reactive handle for halogen installation.[60] Many of the same issues of selectivity exist with the installation of these groups as with the direct halogenation of C-H bonds, however.



Scheme 11: Selected methods for the preparation of organohalogen compounds.[60]

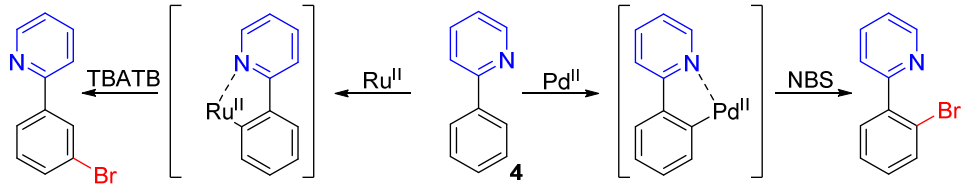
A more recent strategy for the direct halogenation of C-H bonds is *ortho*-metallation. Here, coordinating effects between the substrate and a stoichiometric organometallic species are used to control the position of deprotonation (**Scheme 12**). Electrophilic sources of halide can then be attacked by the subsequent carbanion to afford organohalogen product.[65,66] The substrate scope of these methods is limited however, with a limited number of motifs capable of the required coordination, and the control of certain positions is not possible. Additionally, the use of such strong organometallic bases precludes the presence of unstable moieties.



Scheme 12: Example of an *ortho*-lithiation approach to bromination of an aromatic substrate.[65] (TMEDA = tetramethylethylenediamine).

2.1.2 C-H Activation for Halogenation

An alternative to *ortho*-metallation is to use catalytic amounts of a transition metal to metallate C-H bonds, generating reactive organometallic intermediates, which can then catalyse subsequent functionalisation reactions – termed C-H activation.[55,67,68] In a similar vein to *ortho*-metallation approaches, coordinating effects within the substrate typically stabilise a particular organometallic intermediate, thereby allowing regiocontrolled functionalisations through further catalysis by the metal. A common substrate for these regioselective metallation steps is 2-phenyl pyridine (**4**). Here the nitrogen of the pyridine moiety is ideally situated to stabilise the cyclo-metallated intermediate formed from insertion of metal across the *ortho*-CH bond (**Scheme 13**). By employing different metal catalysts, alternative positions of the phenyl ring can then be activated for subsequent halogenations (**Scheme 13**). Palladium for example activates *ipso*- and therefore treatment with N-halo succinimides selectively affords *ortho*-halo aryls.[69] Ruthenium catalysts on the other hand have been found to activate *para*- to the metallated position and therefore can afford *meta*-brominated aryls when treated with TBATB.[11] Other directing groups such as esters, nitriles and triazoles are also capable of templating metallation in this way to allow regioselective halogenation.[70-72] Whilst these methods can allow impressive regio-control, the requirement for such directing groups can limit their substrate scope and mean that certain regioisomers are still inaccessible.



Scheme 13: Examples of C-H activation using the 2-phenyl pyridine moiety (**4**) for either *ortho*- or *meta*-bromination using Pd^{II} or Ru^{II} respectively.[11,69] (NBS = *N*-bromo succinamide (**3**). TBATB = tetrabutylammonium tribromide).

2.1.3 Biocatalytic Halogenation

Whilst the methods of C-H activation for halogenation discussed above are promising avenues towards selective methodologies for the synthesis of organohalogens, complementary methods may be found using biocatalysis. Given the impact that the installation of a halogen atom can have upon the bioactivity of small molecules, it is unsurprising that halogenated natural products have been isolated from a diverse range of microorganisms.[73] Study of the pathways responsible for the biosynthesis of these compounds has identified a number of distinct classes of halogenase.[74-78]

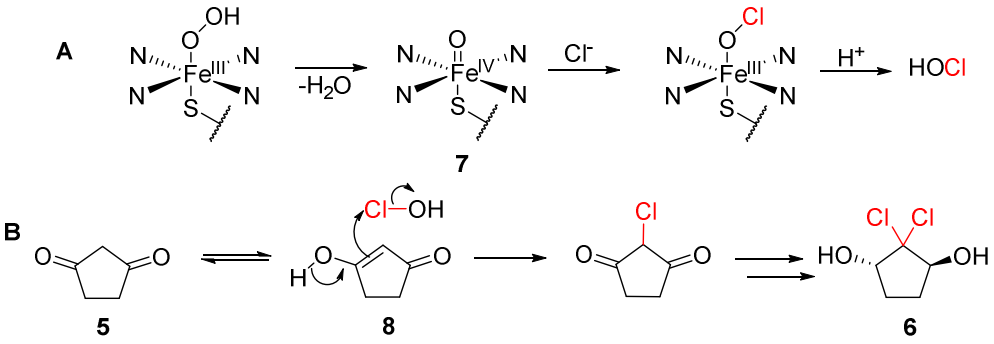
Using these enzymes as biocatalysts for halogenation has a number of potential advantages. By virtue of evolving to utilise benign bio-available sources of halide, the need for toxic sources of halogens could be eliminated. The operation of enzymes in aqueous solvents under ambient conditions could afford processes with lower energy demand and higher functional group tolerance that the current methods, in addition to reducing reliance on toxic organic solvents. Most importantly the potential selectivity of these enzymes could provide more efficient syntheses of organohalogen compounds, as well as halogenation patterns not currently directly accessible. The

main distinctions between different classes of halogenase are based upon their mechanism for halide installation, as well as their substrate preference.[74-78]

2.1.3.1 Haloperoxidases

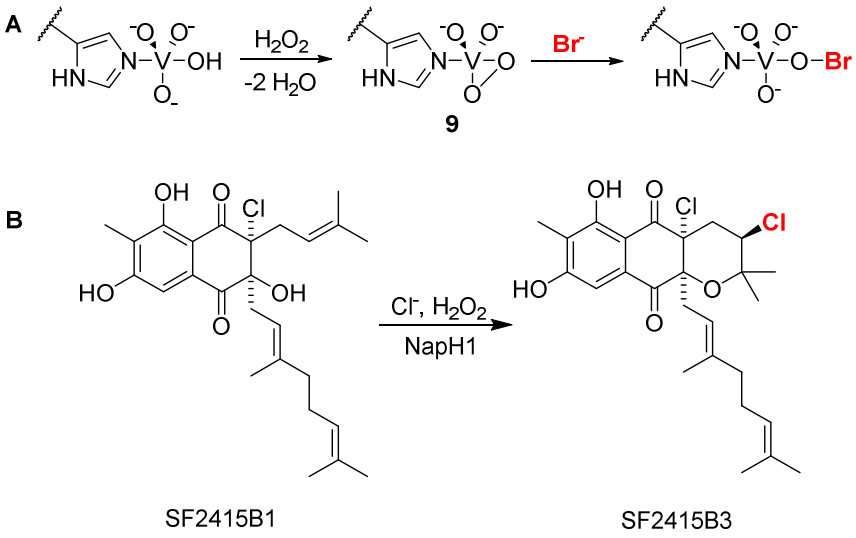
Haloperoxidases are the earliest class of halogenase to be identified and characterised *in vitro*.[76] These metalloenzymes use either iron or vanadium to assist in the generation of oxidised halide, hypohalous acids, which ultimately act as the halogenation species.

Chloroperoxidase from *Caldariomyces fumago* was found to be responsible for the dichlorination of 1,3-cyclopentanedione (**5**) *en route* to calariomycin (**6, Scheme 14B**).[79] Subsequent crystallographic and mechanistic studies found hypohalous acid generation to occur via catalysis at the Fe^{III} center of a heme-cofactor, through generation of the highly activated Fe^{IV}-oxo species (**7**, **Scheme 14A**).[80,81] Reaction of chloride with this species then generates the electrophilic chlorine source HOCl.[82] The lack of selectivity observed in transformations involving these enzymes suggests that HOCl is likely to be freely-diffusing, and therefore react at the most electron-rich position of most substrates.[76] Indeed, attack at the 2- position of **5** is likely promoted by increased stabilisation of the enone intermediate **8** (**Scheme 14B**).



Scheme 14: Proposed mechanism of hypochlorous acid generation by Fe-heme haloperoxidases and involvement of CPO from *C. fumago* in the biosynthesis of calariomycin (**6**).

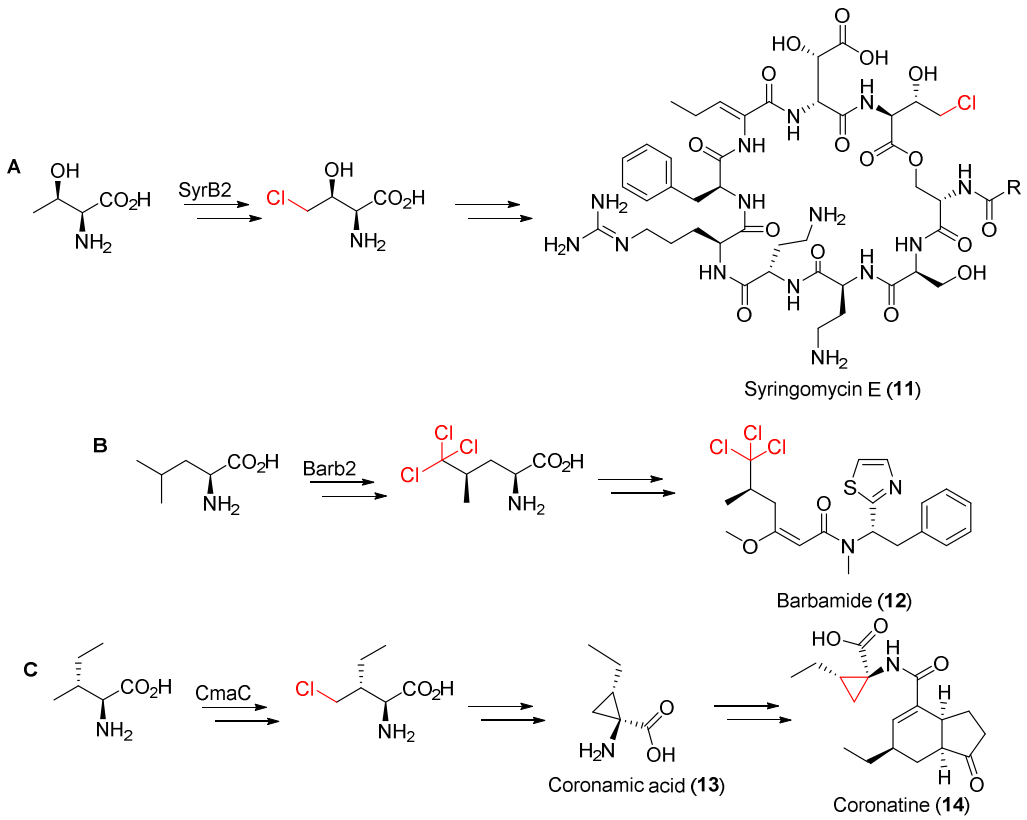
Unlike the Fe-dependent haloperoxidases, the vanadium-dependent halogenases are not thought to require a change in oxidation state of the metal centre to generate hypohalous acid. Oxidation of halide is instead thought to occur via the peroxo-vanadate intermediate **9**, generated by binding of peroxide to V^V and subsequent loss of H_2O (**Scheme 15A**).[83,84] Many enzymes of this class have also been found to exhibit poor regioselectivity, suggesting that HOCl is freely diffusing.[76,85,86] A small number of examples which do exhibit some regio- and stereo-selectivity with some substrates have recently been discovered however, suggesting that these may bind substrate and position it towards the halogenation species in some manner (**Scheme 15B**).[87-89] Given the lack of selectivity with most of the haloperoxidases, however, the application of these enzymes as biocatalysts for selective halogenation has been quite limited.



Scheme 15: (**A**) Proposed mechanism of hypobromus acid (HOBr) generation by vanadium-dependent haloperoxidases. (**B**) example of a vanadium chloroperoxidase reaction found to proceed with selectivity.[87]

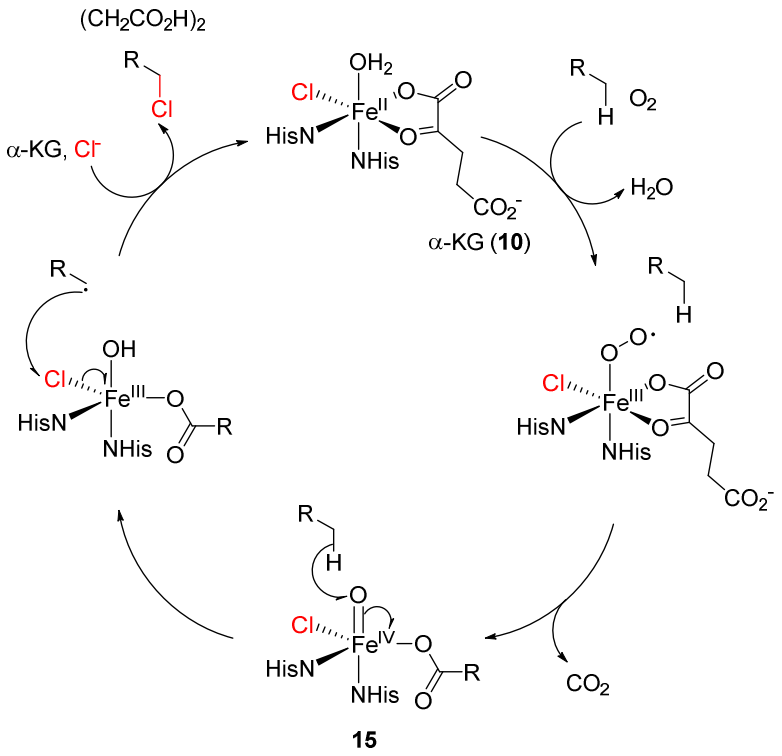
2.1.3.2 Fe^{II}/α-ketoglutarate-dependent Halogenases

A distinct class of iron-dependent halogenase has been discovered which are capable of halogenating unactivated aliphatic C-H bonds. Rather than heme, this class utilise α-ketoglutarate (**10**, αKG) as the ligand for coordination to Fe^{II}.[76,78] The first enzyme of this class to be characterised was SyrB2, found to be responsible for the halogenation of carrier protein-tethered threonine during the biosynthesis of Syringomycin E (**11**) by *Pseudomonas syringae* (**Scheme 16A**).[90] A number of Fe^{II}/α-KG halogenases have subsequently been found to be involved in the biosynthesis of other natural products. For example the trichloromethyl group of barbamide (**12**) was found to be derived from trichlorination of a carrier protein-tethered L-leucine by two Fe^{II}/α-KG halogenases (BarB1 and BarB2, **Scheme 16B**).[91] Biosynthesis of the cyclopropyl ring of coronamic acid (**13**) by *P. syringae* was found to involve CmaB and CmaC, a carrier protein and a Fe^{II}/α-KG halogenase, responsible for the halogenation of L-allo-leucine prior to cyclisation (**Scheme 16C**) *en route* to coronatine (**14**).[92]



Scheme 16: Known involvement of Fe^{II}/α-ketoglutarate dependent halogenases in the biosynthesis of a number of natural products.

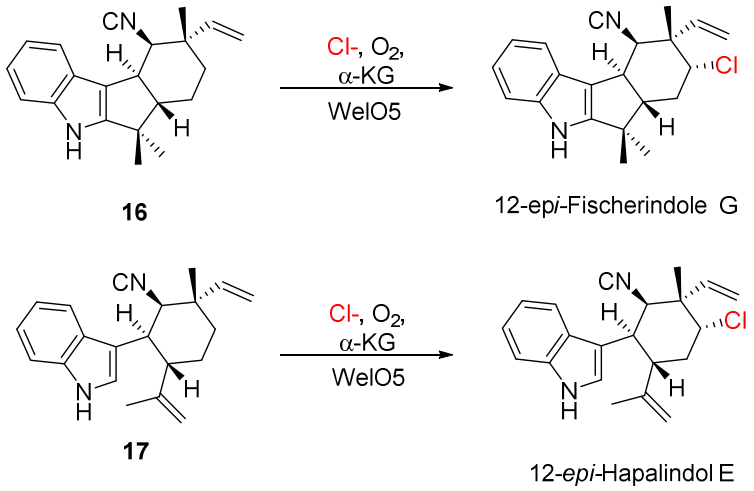
Halogenation of such unactivated positions suggests a mechanism distinct from the haloperoxidases. Analogy to the mechanism of α-KG hydroxylases, in combination with crystallography of SyrB2 and the observed requirement for chloride, α-KG and oxygen for halogenation activity, led to the postulation of a radical mechanism - which would allow activation of inert C-H bonds (**Scheme 17**).[93-95] After binding of chloride to the Fe^{II} centre, it is believed that oxygen binding and decarboxlation of the α-KG ligand affords a Fe^{IV}-oxo species (**15**). This highly reactive intermediate is then thought to be responsible for radical abstraction of a hydrogen atom. Recombination of the carbon-centred radical with Fe-bound $Cl^{\cdot}$ is then thought to regenerate Fe^{II} and afford chlorinated product.[96,97] Coordination of the active halogenation species to the Fe-centre throughout explains why selectivity is observed with these transformations, but not the haloperoxidases.[97]



Scheme 17: Proposed mechanism of halogenation by Fe^{II}/α-ketoglutarate dependent halogenases.

Interestingly a Fe^{II}/α-KG halogenase was recently identified which is capable of catalysing the regio- and stereo-selective halogenation of C-H bonds in freely-diffusing substrates (**Scheme 18**).[98,99] WelO5, from the weltindoline biosynthetic pathway in *Hapalosiphon welwitschii*, is capable of catalysing the chlorination of **16** and **17** to 12-*epi*-Fischerindole G and 12-*epi*-Hapalindol E.[100] The chlorination of free-standing substrates means that this particular enzyme is of greater interest for biocatalytic applications, in preference to those which halogenate tethered substrates.

Given the analogous mechanism of the Fe^{II}/α-KG halogenases and hydroxylases, it seems reasonable that the halogenation and hydroxylation activity of these enzymes could be interchangeable.[97] Crystallography of WelO5 revealed that one particular active site residue of the halogenase, G116, was replaced by an aspartate residue at the analogous position in the hydroxylases. Indeed mutation of the glycine residue of WelO5 to aspartate afforded a complete switch in selectivity from halogenation to hydroxylation.[100] The same rationale has been applied in reverse also, with a Fe^{II}/α-KG hydroxylase being mutated to a halogenase.[101] There is therefore the potential to expand this approach to create new halogenase biocatalysts from Fe^{II}/α-KG hydroxylases - therefore providing regio- and stereo-controlled halogenations of unactivated C-H bonds.



Scheme 18: WelO5 catalysed halogenation of indolinone natural products.

2.1.3.3 Fluorinases

The enzymes involved in the biosynthesis of fluorinated natural products have remained relatively elusive, partially due to their rarity compared to naturally-occuring organochlorine and organobromine compounds.[77] Only five fluorinated natural products have been definitively identified to date (**Figure 3**), with fluoroacetate (**18**) being the most ubiquitous.[77] The first fluorinated metabolite to be identified was the antibiotic nucleocidin (**19**) from *Streptomyces calvus*, although elucidation of the biosynthetic pathway responsible was hampered by poor fluorometabolite production under laboratory conditions.[102] The fluoroacetate (**18**) and 4-fluorothreonine (**20**) producing *Streptomyces cattleya* however, proved to be sufficiently tractable and allowed the identification of the fluorinase responsible for the fluorination of *S*-adenosyl methionine (**21**, SAM) to afford 5'-deoxy-5'fluoroadenosine (**22**, FDA, **Figure 4**).[103] Since this seminal work, a number of other fluorinases have been identified with high sequence identity to the fluroinase from *S. cattelya*, and some have had fluorination activity verified *in vitro*.[104-106]

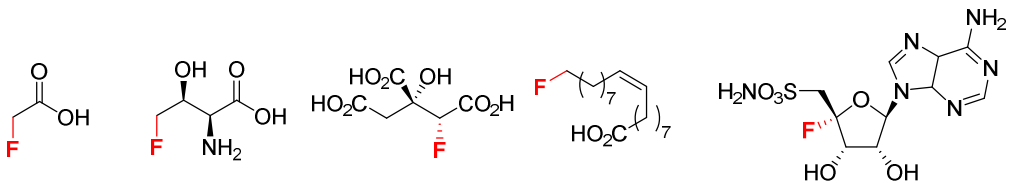


Figure 3: Fluorinated natural products identified to date.

The *S. cattelya* fluorinase has subsequently informed much of the functional and mechanistic understanding of this class. The mechanism of enzymatic fluorination is particularly intriguing because the high redox potential of fluoride would prelude function by any of the oxidative mechanisms previously mentioned herein. Additionally, the tight solvation of fluoride in water would create a great energetic barrier for its direct use as a nucleophile from an aqueous source.[77] Crystallography of the *S. cattelya* fluorinase with FDA (**22**) product bound suggested that fluoride could be bound within the active site through electrostatic interactions (**Figure 4**).[107] These

interactions are thought to position fluoride at the ideal trajectory for an S_N2 displacement at the ribose 5'-C – consistent with the observation that fluorination occurs with inversion of stereochemistry using isotopically labelled substrates.[108,109] The energetic penalty for desolvation of fluoride is thought to be compensated for by the high binding affinity of SAM to the active site, in addition to the retention of four hydrogen-bonded waters within the active site.[108,110]

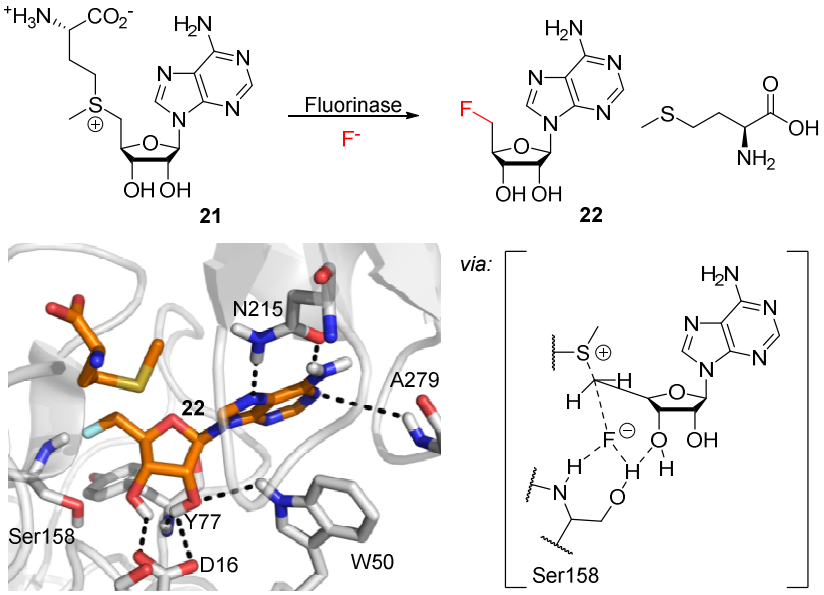
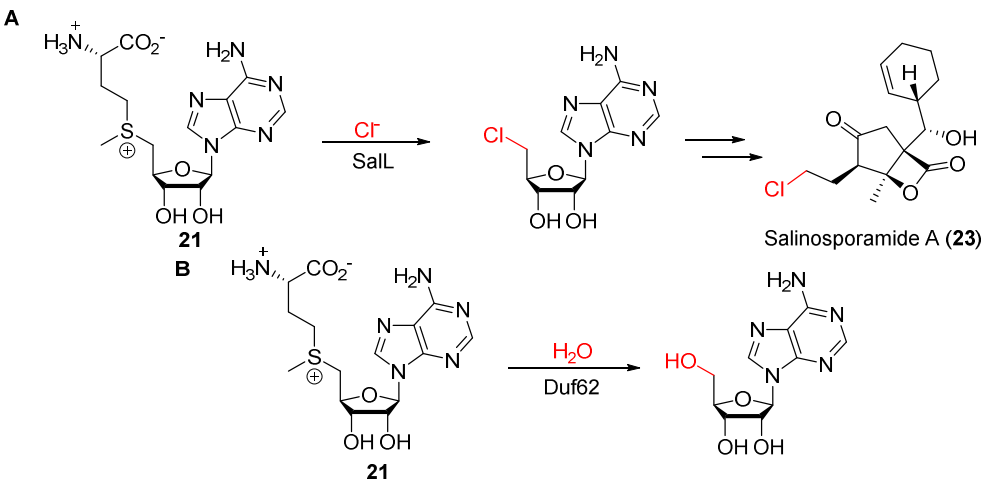


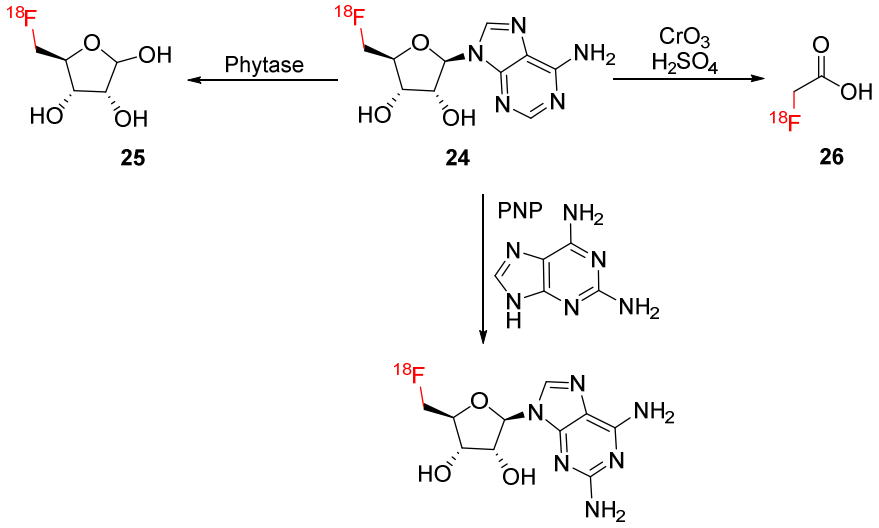
Figure 4: Fluorinase-mediate production of 5'-fluoro adenosine (**22**, FDA) showing some of the interactions residues to contribute to the high binding affinity of SAM (PDB: 1RQR) and the proposed arrangement of fluoride relative to SAM substrate for S_N2 displacement.

A mechanistically-related chlorinase (SalL) has been identified from *Salinispora tropica*, responsible for chlorination during the biosynthesis of salinosporamide A (**23**, **Scheme 19A**).[111] Unlike the *S. cattelya* fluorinase which can utilise either chloride or fluoride as nucleophile,[112] the *S. tropica* chlorinase can only utilise chloride. Similarly the putative fluorinase from *Pyrococcus horikoshii* was found to act as a hydroxylase, utilising hydroxide (from water) as nucleophile, rather than a chlorinase or fluorinase (**Scheme 19B**).[113]



Scheme 19: Identified chlorinases and hydroxylases related to fluorinases.

The biocatalytic applications of fluorinases have mainly involved the ^{18}F labelling of small molecules (**Scheme 20**). The radioactive properties of ^{18}F mean that such labelled compounds are commonly used in positon-emission tomography (PET). As ^{18}F is typically produced from the bombardment of $H_2{}^{18}O$, most sources of ^{18}F are aqueous solutions containing $^{18}F^-$.[77] A method for the direct use of such sources, without drying or coordination to specialist ligands, would therefore be desirable because of the potential for expedient routes with higher radiochemical yields (RCY) than are currently possible. It was found that $^{18}F^-$ equivalents could indeed by utilised by the *S. cattelya* fluorinase to produce ^{18}F-FDA (**24**).[114] To allow access to additional ^{18}F-labelled nucleotides, this reaction was combined with purine nucleotide phosphorylase (PNP) and pyrimidine nucleotide phosphorylase (PyNP) enzymes to enable "base swapping".[115] Use of a phytase enzyme also allowed the nucleobase to be replaced with a hydroxyl group to afford 5'-^{18}F ribose (**25**).[116] A Kuhn-Roth oxidation was also employed to allow conversion of ^{18}F-FDA to ^{18}F fluoroacetate (**26**) – a common PET agent in neurology and oncology.[117] The conjugation of these small ^{18}F-nucleotides onto large biological molecules has been achieved by the use of functionalised nucleobases, which allow orthogonal conjugation reactions to occur with tagged peptides.[118]



Scheme 20: Use of fluorinase-produced 5'-^{18}FDA (**24**) the synthesis of other ^{18}F small molecules. (PNP = purine nucleotide phosphorylase).

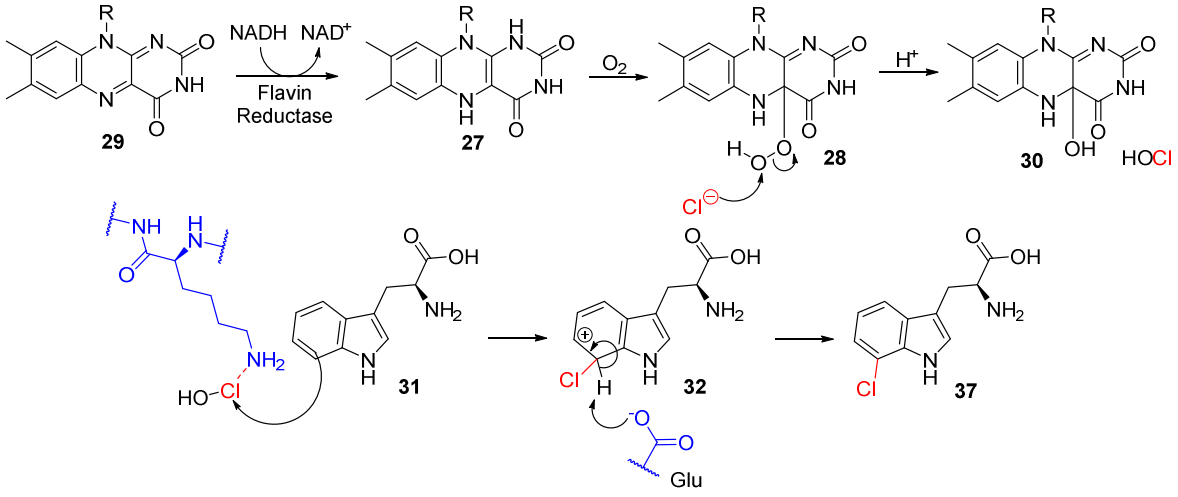
Given that fluorination can have a profound effect upon biological activity, and that natural products are common sources of medicinal compounds, methods for the facile preparation of fluorinated natural products which do not require laborious total synthesis are desirable. In an attempt to achieve such methods, fluorinase genes have been included into the biosynthetic pathways of natural products.[119] In this work the fluorinase gene from *S. cattelya* was inserted into the SalL encoding gene of the salinosporamide (**23**) producer *Salinospora tropica*. The resulting mutant strain therefore produced fluorosalinosporamide directly by fermentation.[119]

Direct application of fluorinases to the production of other fluorinated natural products, or fluorinated small molecules, is limited by the substrate scope of these enzymes. Although fluorination ability is thus far unique to this class, their nucleophilic mechanism is dependent upon a methionine leaving group, in addition to the high binding affinity of the natural substrate SAM to allow desolvation of fluoride. Such requirements mean that the fluorinases currently known are unlikely to find any real application as biocatalysts beyond what has already been realised.

2.1.3.4 Flavin-Dependent Halogenases

The flavin-dependent halogenases (FlHals) belong to the superfamily of flavin-dependent monooxygenases. This class activate molecular oxygen through reaction with reduced FAD (**27**, $FADH_2$) to generate C4a-hydroperoxy flavin (**28**).[120,121] This highly activated intermediate then allows reactions such as hydroxylation, epoxidation and Baeyer-Villiger oxidations.[120,121] The flavin-dependent halogenases are classified as two-component monooxygenases because $FADH_2$ (**27**) is usually produced by a second flavin-reductase enzyme from FAD (**29**), rather than the halogenase enzyme itself.[121] Whilst many FlHals halogenate freely-diffusing substrates, a significant number require substrate to be activated or tethered onto an ancillary protein in a similar way to most of the Fe^{II}/α-KG halogenases.[76,78]

It was originally suggested that the FlHals operate via a mechanism analogous to other flavin-dependent monooxygenases, whereby substrate interacts directly with either C4a-hydroperoxy flavin or a related chlorinated flavin-moiety.[122,123] Crystallography of the tryptophan (**31**) 7-halogenase PrnA subsequently revealed, however, that $FADH_2$ and substrate were bound in spatially distinct sites, separated by a tunnel, which would prevent direct interaction of substrate and **28**.[124,125] Stopped-flow experiments subsequently demonstrated that reaction of $FADH_2$ with molecular oxygen occurred prior to halogenation, followed by generation of C4a-hydroxy flavin (**30**), suggesting that chloride attacks **28** to form hypochlorous acid (**Scheme 21**) .[126]



Scheme 21: Generation of C4a-hydroperoxy flavin (**28**) and proposed mechanism of chlorination by the flavin-dependent halogenases showing one of the potential interactions of active site lysine with the electrophilic halogenation species.

Radiolabelling experiments subsequently implied the formation of a long-lived enzyme-chloride adduct in the absence of substrate.[127] The exact nature of this adduct is still debated, however the positioning of an active site lysine relative to the C-H position of substrate which is halogenated, and mutagenesis studies demonstrating that this lysine is essential for halogenation activity, suggest that the lysine is involved in some way to direct or activate oxidised halide for attack by the aromatic substrate.[127] As with a typical electrophilic aromatic substitution, this would then lead to generation of a Wheland intermediate (**32**), which is then thought to be deprotonated by an active site general base - conserved in most FlHals.[127-130] Comparison of the crystal structures of the tryptophan 7- and 5-halogenases PrnA and PyrH, respectively (**Figure 5**), showed that in both cases there is a lysine closest to the C-H position of substrate which is halogenated - suggesting that positioning of substrate relative to this lysine is critical in control of FlHal regioselectivity.[124]

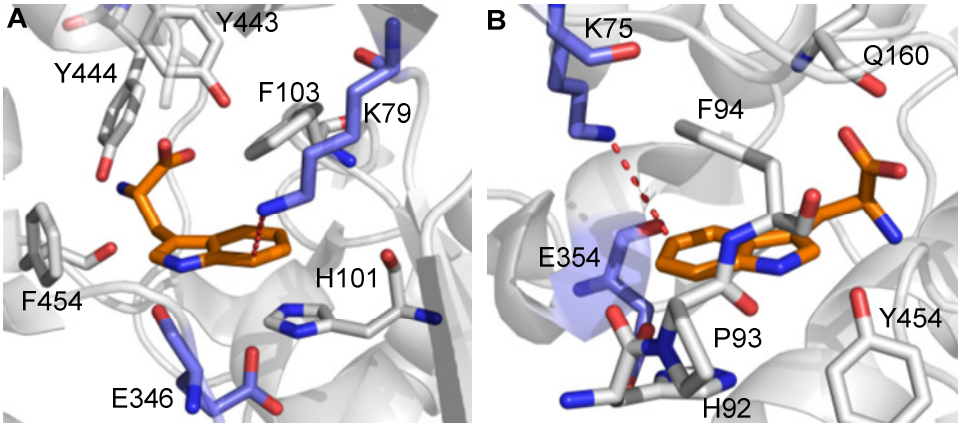
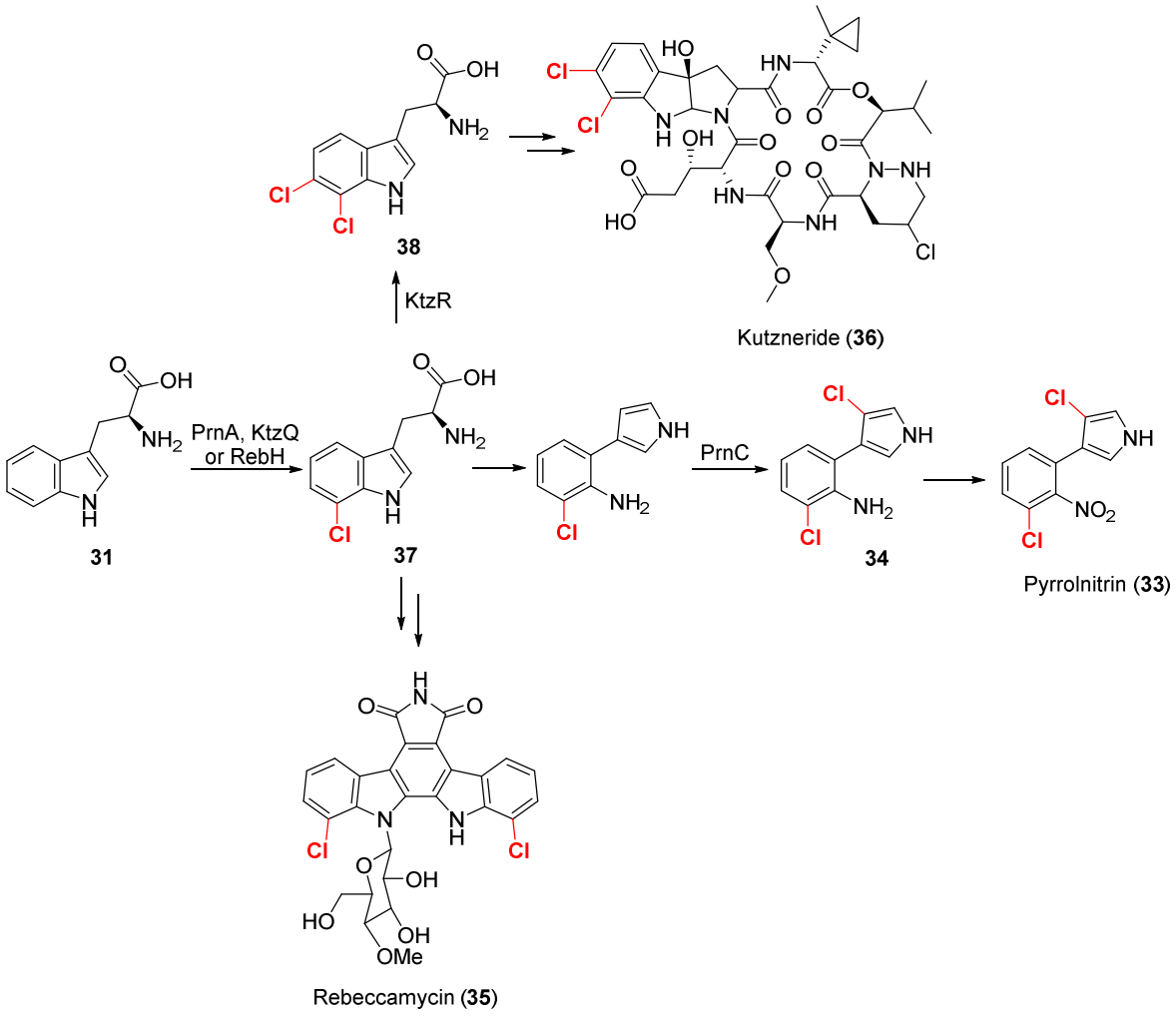


Figure 5: Crystal structures of PrnA (**A**) and PyrH (**B**) showing positioning of the proposed catalytic lysine and general base (blue) relative to the C-H position of substrate which is halogenated. Selected active site residues thought to be involved in the orientation of tryptophan relative to this lysine are also shown. (PDB: 2AQJ and 2WET respectively).

The FlHals can be further subdivided based upon the nature of substrates which they halogenate.[74-76,78] The tryptophan (**31**) FlHals are the most extensively studied and characterised FlHals with PrnA, the tryptophan 7-halogenase involved in biosynthesis of pyrrolnitrin (**33**) by *Pseudomonas fluorescens*, being the first.[131] After halogenation of the 7-position of tryptophan by PrnA, a second FlHal is responsible for halogenation of the pyrollic intermediate **34** (**Scheme 22**). RebH was subsequently identified from the rebeccamycin (**35**) biosynthetic pathway in *Lechevalieria aerocolonigenes*, and found to also be a tryptophan 7-halogenase.

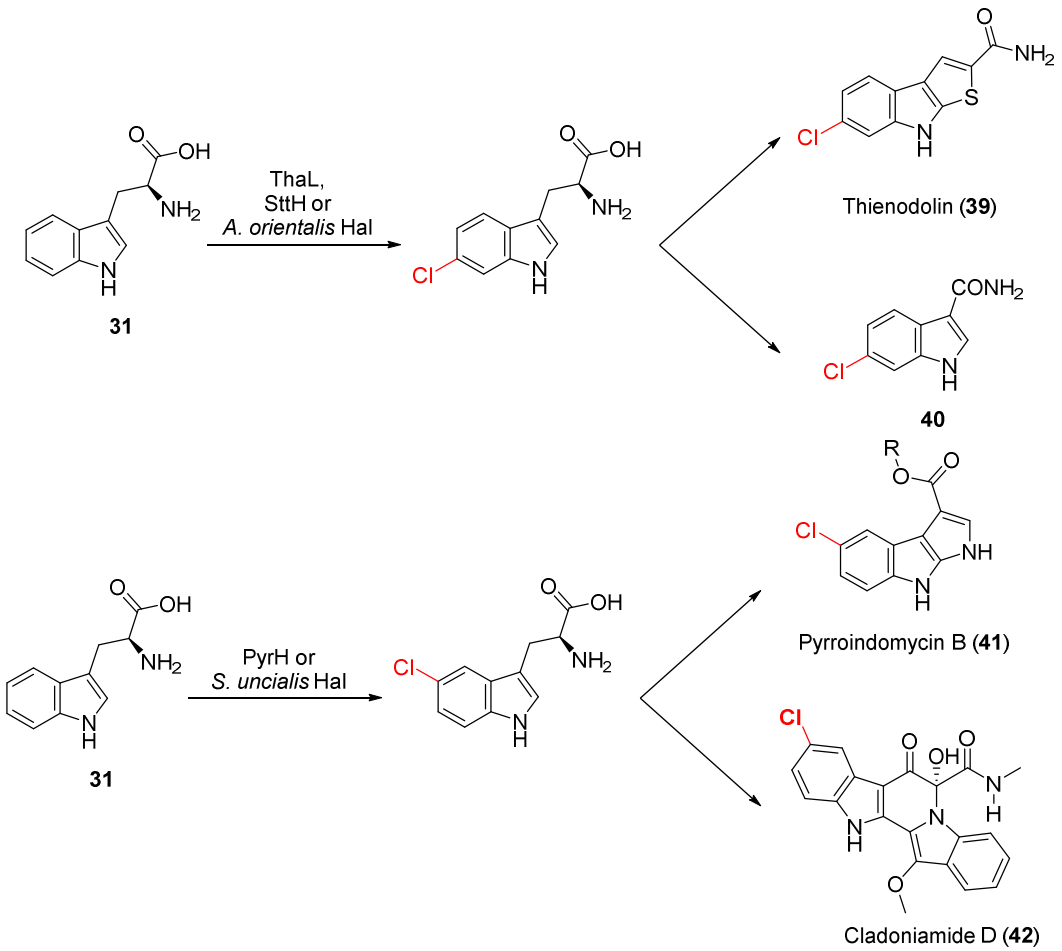
Similar to the biosynthesis of pyrrolnitrin, *Kutzneria sp.* 744 employs two different FlHals with complementary regioselectivity during the biosynthesis of kutzneride (**36**), in addition to a cryptic Fe^{II}/α-KG halogenase.[132] KtzQ is responsible for the 7-chlorination of tryptophan, whilst KtzR is responsible for the 6-chlorination of **37** to afford 6,7-dichloro tryptophan (**38**, **Scheme 22**).[132] Tryptophan 6-halogenases have also been identified which operate in isolation of other FlHals. ThaL for example, involved in the biosynthesis of thienodolin (**39**) by *Stretomyces albogriseolus*,[133]

and SttH, from a currently unidentified biosynthetic pathway in *Streptomyces toxytricini*,[134] have both been demonstrated to act as tryptophan 6-halogenases *in vitro* (**Scheme 23**). Additionally a tryptophan 6-halogenase was recently identified from *Amycolatopsis orientalis*, which is thought to be involved in the biosynthesis of **40**.[135] Interestingly a FlHal was recently identified from a marine sponge which was found to chlorinate the 6-position of 5-hydroxy tryptophan *in vitro*, as well as a number of other non-natural substrates.[136]



Scheme 22: Involvement of flavin-dependent tryptophan 7- and 6-halogenases in the biosynthesis of various natural products.

The tryptophan 5-halogenase PyrH, from the biosynthesis of pyrroindomycin (**41**) by *Streptomyces rugosporus*, is the only fully characterised tryptophan 5-halogenase to date which processes freely-diffusing tryptophan as a substrate (**Scheme 23**).[137] There is a tryptophan 5-halogenase known to be involved in the biosynthesis of cladoniamide (**42**) by *Stretomyces uncialis*, although full *in vitro* characterisation is still lacking.[138] A flavin-dependent halogenase has also been recently identified as being involved in the 5-chlorination of a tryptophan moiety in a peptide chain during biosynthesis of the lantibiotic NAI-107.[139]



Scheme 23: Involvement of tryptophan 6- and 5-halogenases during biosynthesis of various natural products.

With the exception of PrnC, the pyrrolic FlHals characterised to date halogenate peptidyl carrier protein (PCP) tethered pyrrol-2-carboxy thioester substrates, generated from the oxidation of proline-*S*-PCP intermediates.[74,76] HalB and Clz5, the flavin-dependent halogenases responsible for chlorination of PCP-bound pyrrole thioester substrates during the biosynthesis of pentachloropseudilin (**43**) and chlorizidine A (**44**) respectively (**Figure 6A**), appear to have relaxed regioselectivity compared to the tryptophan halogenases, which is likely due to the highly activated nature of the pyrrole ring.[140,141] A particularly interesting flavin-dependent pyrrole halogenase, Bmp2, was recently identified from the biosynthesis of brominated marine natural products (**Figure 6B**).[142,143] This particular pyrrole FlHal appears to be capable of bromination and iodination but not chlorination.[143] This is surprising since only one other FlHal has been found capable of iodination, and tryptophan FlHals have been found to be inhibited by the presence of I^-.[128,130] Understanding the structural reasons for the halide preference of Bmp2 may allow engineering of other FlHals from chlorinases and brominases into iodinases.

The substrate selectivity for PCP-bound substrates means that substrates must have a moiety which allows conjugation, thereby limiting substrate scope, and would mean that additional biocatalysts and biotransformations are required for the synthesis of such substrates when

compared to halogenases which process freely-diffusing substrates. Most of the currently known pyrrole FlHals are therefore of limited direct application to biocatalysis.

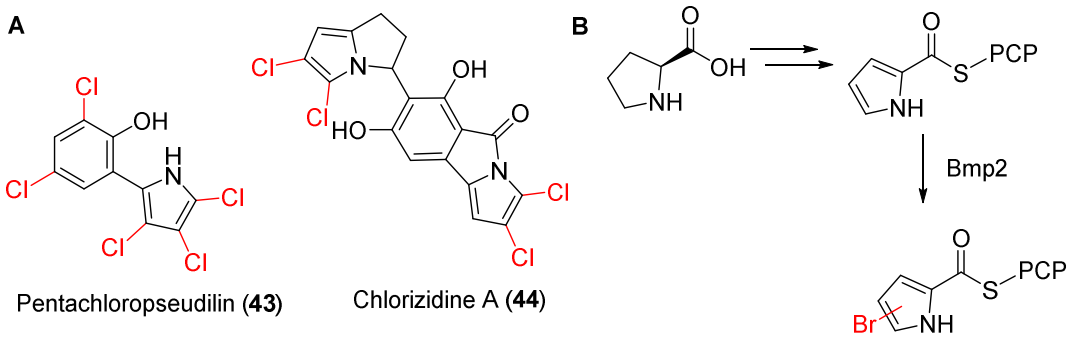


Figure 6: (**A**) Chlorinated natural products generated by the chlorination of pyrrole thioester-PCP intermediates by pyrrole FlHals. (**B**) Halogenation of PCP-*S*-pyrrole by Bmp2.

FlHals are also known which are capable of halogenating phenolic moieties.[76,78] Like the pyrrole FlHals, many of these require substrates which are PCP-bound. Given the limited potential of these enzymes as biocatalysts, only the phenolic FlHals which halogenate freely-diffusing substrates will be discussed. Two such enzymes, Rdc2 from *Pochonia chlamydosporia* and RadH from *Caetonium chiversii*, are both involved in the biosynthesis of radicicol (**45**) – a potent Hsp90 inhibitor.[144-147] These two enzymes act as post-PKS tailoring enzymes, and therefore can halogenate free-standing substrates. Several other phenolic FlHals which act on free-standing substrates have also been identified from the biosynthesis of diverse fungal natural products (**Figure 7**) such as griseofulvin (**46**), aspirochlorine (**47**), chaetovirdin (**48**), ochratoxin (**49**) and dihydrogeodin (**50**).[148-152] Each of these enzymes, like RadH and Rdc2, chlorinate *ortho*- to a hydroxyl group, except for the halogenase from ochratoxin (**49**) biosynthesis which appears to chlorinate *para*-.

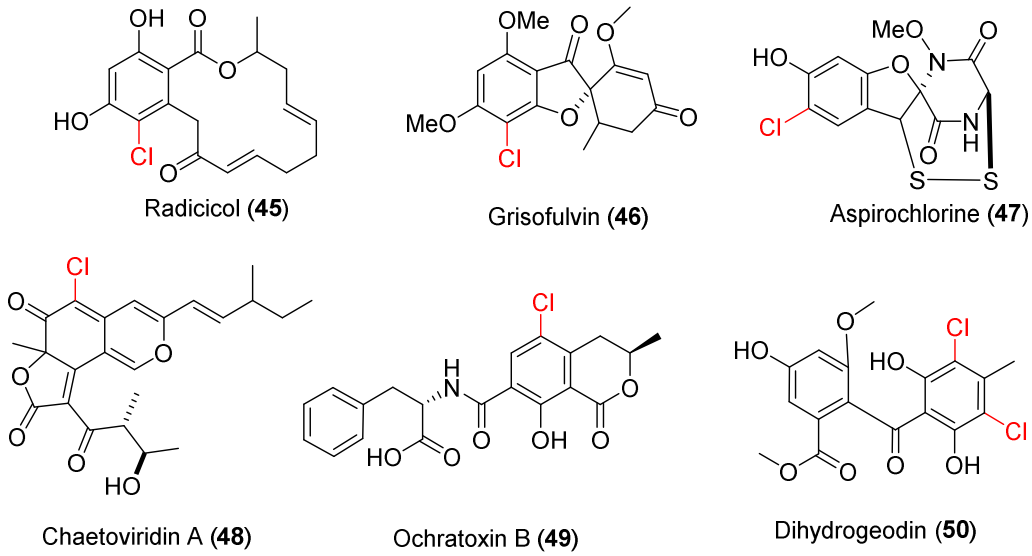
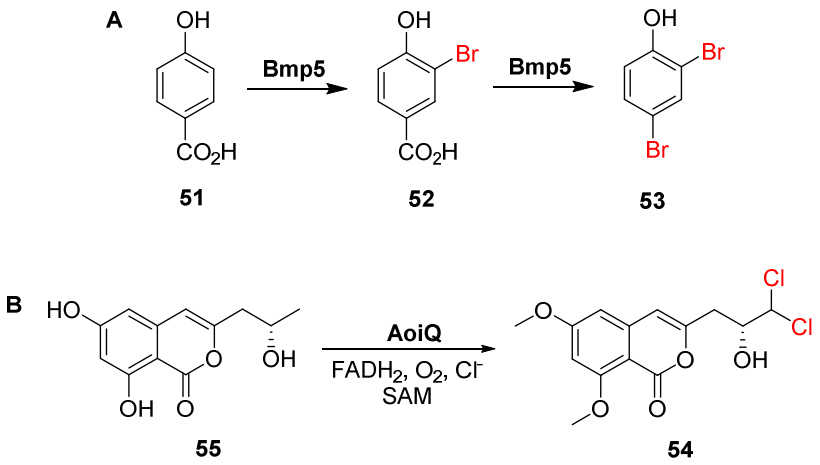


Figure 7: Chlorinated natural products biosynthesised by the chlorination of free-standing phenolic substrates by phenolic FlHals.

These phenolic FlHals are thought to operate via the same mechanism as the tryptophan halogenases, with the observed preference for chlorination *ortho*- to OH groups arising from deprotonation of the hydroxyl group or stabilisation of a benzoquinone intermediate. A phenolic FlHal was recently identified however, which appears to operate via a modification of the canonical electrophilic aromatic substitution mechanism.[143] Bmp5, from *Pseudoalteromonas luteoviolacea*, was demonstrated to not require an external flavin reductase enzyme for *in vitro* halogenation activity – instead directly utilising bromide, FAD and NADH.[143] This observation suggests that Bmp5 is similar to other single-component monooxygenases, meaning that $FADH_2$ generation and substrate oxygenation are catalysed by separate domains of the same enzyme.[121] Bmp5 was found to halogenate 4-hydroxybenzoic acid (**51**) *ortho*- to the hydroxy group. After the first bromination, it was then observed that **52** is converted to the di-bromo phenol **53** with concurrent decarboxylation (**Scheme 24A**) .[143] Such a decarboxylative halogenation has no precedent with other FlHals. Additionally, like Bmp2, Bmp5 is capable of utilising bromide or iodide but not chloride. The full mechanistic and structural elucidation of Bmp5 is therefore of great interest for the development of biocatalytic halogenation methods.



Scheme 24: Activity of the non-canonical FlHals (**A**) Bmp5 and (**B**) AoiQ.

Surprising FlHal activity has also been found with ClmS, the halogenase from chloramphenicol biosynthesis, and AoiQ, from diaporthins (**54**) biosynthesis, which are responsible for the chlorination of aliphatic C-H bonds.[76,153,154] Such a reaction is not intuitive from the electrophilic substitution mechanism proposed for other FlHals. ClmS is thought to halogenate acyl-CoA intermediates, suggesting that halogenation could occur via attack of hypohalous acid by an enolate intermediate.[153] This mechanism does not fit with the halogenation of diaporthin (**55**) by AoiQ *in vitro* (**Scheme 24B**), however, as the generation of such an enolate intermediate would require additional oxidoreductase activity.[154] Understanding the mechanism of this halogenation is therefore of great academic interest.

2.1.4 Conclusions and Outlook

Given their ability to halogenate aromatic substrates in a regio-divergent and regio-selective manner, a transformation which is challenging by traditional means, the flavin-dependent tryptophan halogenases have attracted most interest as biocatalysts for halogenation.[74-76] Early work found a number of these enzymes to be capable of both chlorinating and brominating non-natural tryptophan derivatives, indoles, benzamides, benzoic acids and napthyl substrates.[155-160] Similarly, the phenolic FlHals have also been explored for the halogenation of phenolic substrates.[144-146,161] Other work has focussed on using structure-guided and random mutagenesis to alter the substrate scope and regioselectivity of these enzymes.[155,156,158,162] The utility of these biotransformations is limited, however, by the low activity and poor stability of the FlHals.[163,164] Since the FlHals operate via an electrophilic aromatic substitution reaction, an energetically demanding process, k_{cat} values are typically in the region of 1 min^{-1}. Coupled with the poor stability of FlHals, this has meant that application of FlHals as biocatalysts is currently limited to laboratory-scale investigations. To address this, a number of groups have reported using directed evolution to improve the thermostability and catalyst lifetime of these enzymes to give variants more appropriate for biocatalytic applications.[164,165] To facilitate the expedient screening of the mutant libraries generated during directed evolution, a number of high-throughput screening methods for halogenase activity have also been reported.[162,165,166]

An alternative approach to enzymes with improved stability and productivity, which are usually better biocatalysts, is to identify enzymes which catalyse analogous transformations in thermophilic organisms. Such variants have been found to possess improved thermostability and catalyst lifetime, in addition to improved tolerance to organic solvents.[164,165,167] Given our aims to integrate halogenase enzymes with chemocatalytic reactions, which typically have markedly different operating conditions to biocatalysts, we identified a putative tryptophan halogenase gene from a thermophilic *Stretomyces* strain and probed its use as a biocatalyst. In addition to increased tolerance of the chemocatalytic conditions, it was hoped that this thermophile would provide a generally more robust halogenase biocatalyst for production of intermediate aryl halides.

2.2 Results and Discussion

2.2.1 Identification of a Thermophilic Tryptophan Halogenase (Th_Hal)

There are two motifs which are considered key for the identification of putative flavin-dependent halogenase genes; WxWxIP and GxGxxG. The former is thought to be responsible for preventing direct monooxygenase activity by creating spatial separation of C4a-hydroperoxy flavin (**28**) and substrate, whilst the latter is a conserved nucleoside-binding motif.[76] With this in mind Binuraj Menon, a co-worker on this project, conducted a BLAST analysis of published genome sequences, which revealed a putative gene from *Streptomyces violaceusniger* SPC6, a thermophilic and halotolerant strain, that contained both the FlHal conserved motifs.[168] Other groups had characterised a chitinase from this strain and found it to be of increased stability compared to mesophilic chitinases,[169] suggesting that the putatative halogenase may also possess increased stability. Subsequent sequence alignment of the translated putative FlHal gene with those of other

FlHals revealed high sequence similarity and identity with several (**Table 1**). Highest sequence similarity (85.9 %) was found to the tryptophan 6-halogenase SttH, suggesting that the putative halogenase may also function as a tryptophan 6-halogenase.

To allow the *in vitro* characterisation of this enzyme (termed Th_Hal hereafter), the relevant codon-optimised synthetic gene was acquired and cloned into pET28b(+) by Binuraj Menon for overexpression in *E. coli* BL21 (DE3). Gratifyingly, subsequent purification of the hexahistidine tagged protein by immobilised ion affinity chromatography (IMAC) allowed soluble Th_Hal to be afforded with reasonable purity (**Figure 8A**).

Table 1: Sequence identity and similarity of Th_Hal to several other FlHals. Generated using the optimal global alignment tool of BioEdit and the BLOSUM62 similarity matrix.*KtzR is proposed to chlorinate the 6-position of 7-chlorotryptophan, rather than tryptophan.[132]

FlHal	Regioselectivity	Identity (%)	Similarity (%)
PyrH	5-	53.8	71.0
SttH	6-	75.6	85.9
ThaL	6-	38.7	57.4
KtzR*	6-	71.3	81.5
PrnA	7-	36.9	55.2
RebH	7-	37.7	54.9

As almost all FlHals are two-component monooxygenases, a partner flavin reductase is required for concurrent regeneration of $FADH_2$.[76,121] As the flavin reductase previously used by our group (Fre from *E. coli*) is obtained from a mesophilic organism,[155] it was reasoned that a flavin reductase from a thermophilic organism may be more appropriate for our purposes. A thermophilic flavin reductase would likely be of increased stability, and therefore allow better determination of Th_Hal activity at higher temperatures - as well as a generally more robust biocatalytic halogenation system. Accordingly, a previously characterised thermophilic flavin reductase (Th_Fre) from *Bacillus subtilis* WU-S2B was identified for this purpose.[170] As with Th_Hal, a codon-optimised synthetic gene encoding this protein was acquired and cloned into pACYCDUET-1 by Binuraj Menon for expression in *E. coli* BL21 (DE3). Again, purification by IMAC was sufficient to obtain Th_Fre of the necessary purity (**Figure 8B**).

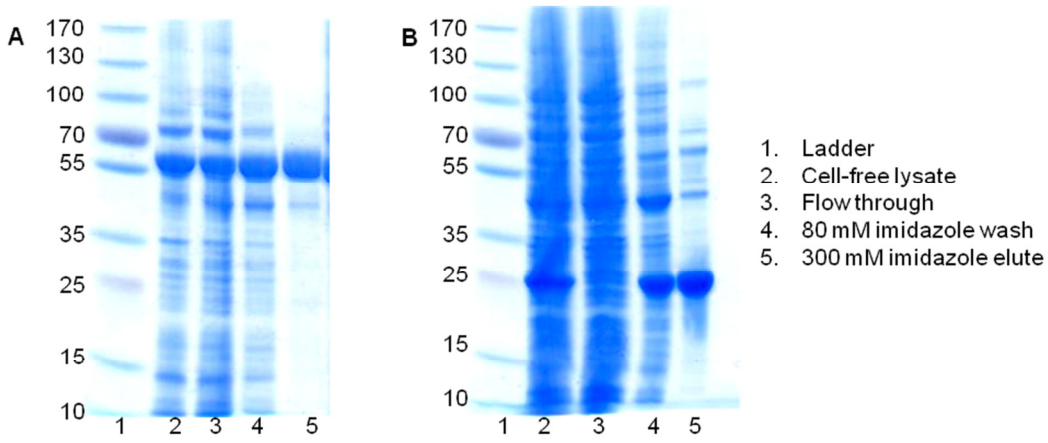
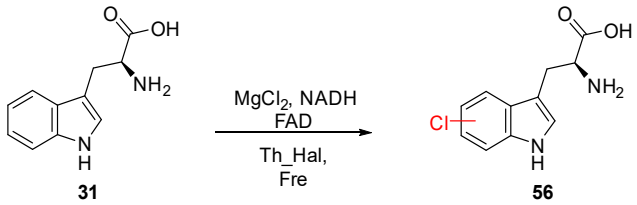


Figure 8: SDS-PAGE of fractions from the purification of (**A**) Th_Hal and (**B**) Th_Fre. In both cases the ladder used was PageRuler^{TM} Prestained Protein Ladder from ThermoScientific^{TM}. Protein mwt: 57.9 kDa (Th_Hal); 20.7 kDa (Th_Fre).

2.2.2 Substrate Scope and Regioselectivity of Th_Hal

With purified Th_Hal in hand, its proposed activity was verified *in vitro* using analytical scale biotransformations. Upon incubation of tryptophan (**31**) with purified Th_Hal, Fre, FAD, NADH and $MgCl_2$, conversion of substrate to a new product was observed by high-pressure liquid chromatography (HPLC, **Figure 9A**). This new peak was not observed in control reactions omitting Th_Hal. Liquid chromatography mass spectrometry (LC-MS) suggested the product to be chlorinated tryptophan (**56**), with the distinctive 3:1 ratio of peaks separated by 2 *m/z* units attributed to the natural abundance of ^{35}Cl and ^{37}Cl (**Figure 9B**)



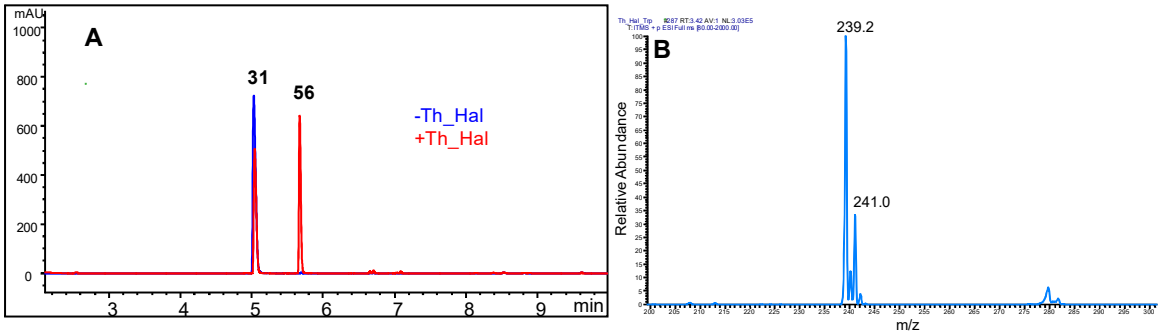


Figure 9: (**A**) HPLC UV chromatogram recorded at 280 nm of the incubation of tryptophan with and without Th_Hal. (**B**) LRMS of the product peak observed upon incubation of tryptophan with Th_Hal and $MgCl_2$.

Subsequent upscale of the reaction allowed isolation of the chlorinated tryptophan product, and NMR analysis showed this to be 6-chloro tryptophan (**57**) based on comparison of the ^{1}H NMR

splitting pattern to that reported in the literature (**Figure 10**).[156] Given the high sequence identity of Th_Hal to the tryptophan 6-halogenase SttH, it is unsurprising that Th_Hal shows tryptophan 6-halogenase activity.

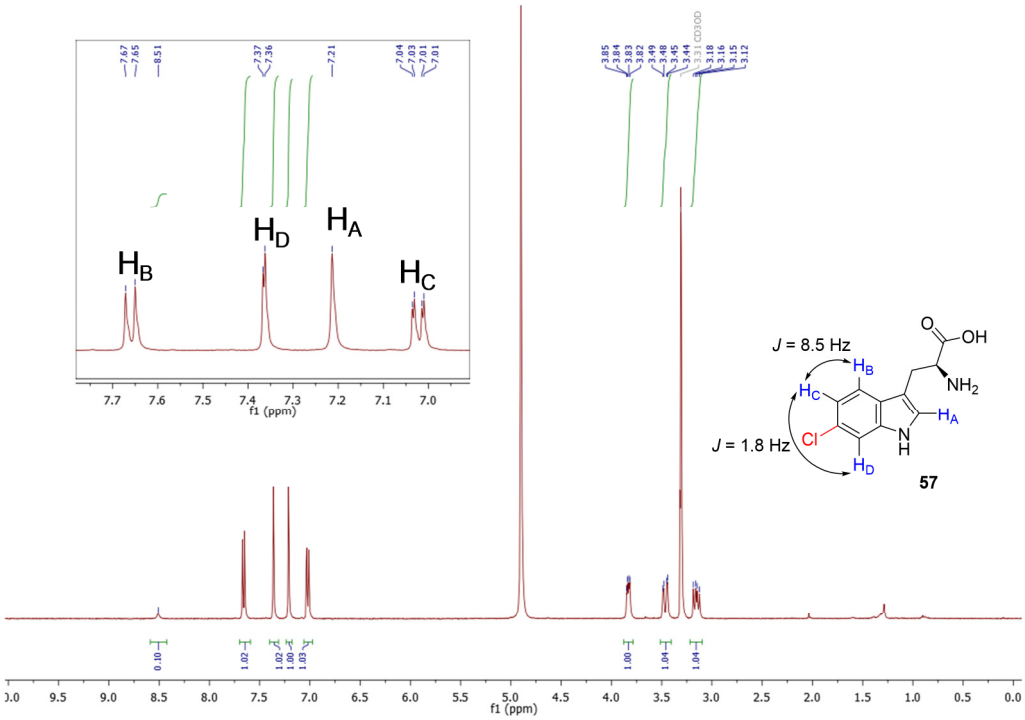


Figure 10: ^{1}H NMR of 6-chloro tryptophan (**57**) (MeOD, 400 MHz) with insert showing expansion between δ 6.9 and 7.8 ppm.

To assess the potential of Th_Hal as a biocatalyst for aromatic halogenation, a panel of additional substrates were tested for halogenation activity by Th_Hal (**Figure 11**). Given the observation that Th_Hal appears to be a tryptophan 6-halogenase, these were selected based upon our group's previous work on the related 5-, 6- and 7-tryptophan halogenases.[155,156] HPLC and LC-MS analysis of analytical scale biotransformations as above suggested that *N*-methyl tryptophan (**58**), 5-hydroxy tryptophan (**59**), kynurenine (**60**), anthranilic acid (**61**) and anthranilamide (**62**) were all halogenated by Th_Hal to give singly chlorinated products (**Table 2**). N-phenyl anthranilic acid (**63**) and 2-amino-*N*-ethyl benzamide (**64**), did not appear to be substrates however.

Chlorination of **58** by Th_Hal appeared to be less efficient than the analogous reaction with tryptophan, suggesting that the loss of a hydrogen-bonding contact between the indole NH and active site has a deleterious effect upon the binding efficiency of **58** to Th_Hal. Indeed, the indole NH has been proposed to form a key hydrogen bonding contact between tryptophan and the active site of other FlHals.[124,130] 5-hydroxy tryptophan (**59**) gave even lower conversion to chlorinated product under identical conditions. It is possible that increased steric bulk around the C-H position which is halogenated reduces the efficiency of halide transfer from the catalytic lysine to substrate. Alternatively, the introduction of such a coordinating moiety could force substrate to adopt a non-optimal conformation within the Th_Hal active site.

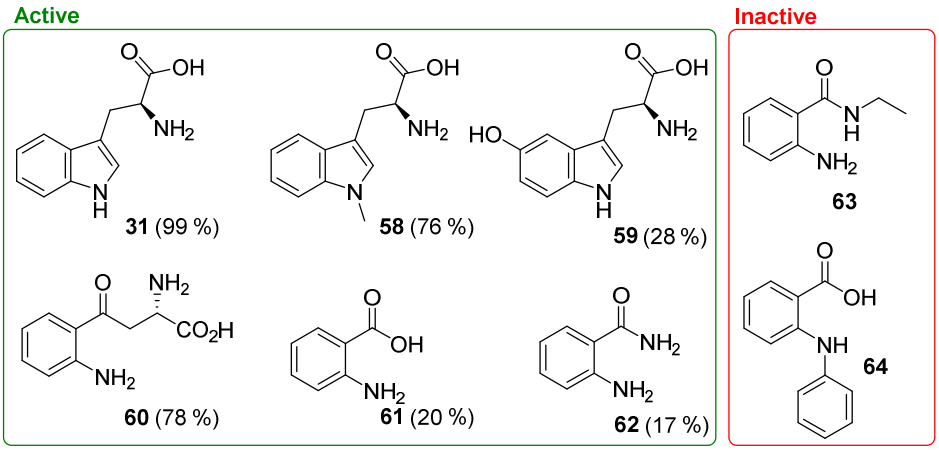


Figure 11: Compounds tested with Th_Hal for chlorination activity. For those compounds where chlorination was confirmed by LC-MS, conversions are shown in brackets. Conversions were determined by analytical HPLC based upon consumption of substrate. Conditions: substrate (0.5 mM), Th_Hal (2.5 μM), Th_Fre (1.0 μM), FAD (1.0 μM), NADH (2.5 mM), $MgCl_2$ (10 mM), 30 °C, 3 hrs.

Of the non-indolic substrates tested, kynurenine (**60**) afforded highest conversion to chlorinated product. This has been observed by our group with a number of tryptophan halogenases, and is thought to be because the relatively large substrate is capable of maintaining several hydrogen bonding contacts with the halogenase active site. The smaller anthranilic acid (**61**) and anthranilamide (**62**) were very poor substrates, with minimal product detected under identical conditions. **61** and **62** are likely to have limited contact to the Th_Hal active site residues due to their small size.

Table 2: Low resolution mass spectrometry (LRMS) peaks observed in the product peak of analytical scale biotransformations of substrates **58** to **62** with Th_Hal. *Approximate.

Substrate	Product *m/z* peaks (relative intensity*, %)	Identity
58	254.9 (25) 253.1 (75) 238.1 (5) 236.0 (20)	$[^{37}M+H]^+$ $[^{35}M+H]^+$ $[^{37}M-NH_2]^+$ $[^{35}M-NH_2]^+$
59	255.9 (25) 253.9 (75)	$[^{37}M+H]^+$ $[^{35}M+H]^+$
60	244.9 (25) 242.9 (75)	$[^{37}M+H]^+$ $[^{35}M+H]^+$
61	171.9 (25) 169.9 (75) 127.9 (20) 125.9 (80)	$[^{37}M-H]^-$ $[^{35}M-H]^-$ $[^{37}M-CO_2]^-$ $[^{35}M-CO_2]^-$
62	172.8 (25) 170.8 (75)	$[^{37}M+H]^+$ $[^{35}M+H]^+$

The regioselectivity of these Th_Hal catalysed chlorinations was subsequently determined using ^{1}H NMR in the same manner as prior. In each case, the regioselectivity was found to follow what would be expected from our group's work on the related tryptophan 6-halogenase SttH (**Figure 12**).[156] Unsurprisingly, *N*-methyl tryptophan was also halogenated exclusively at the 6-position to give **65**. The non-indolic substrates **60** to **62** were halogenated *para*- to the NH_2, a highly activated position of the aniline moiety, to give **66** to **68**.

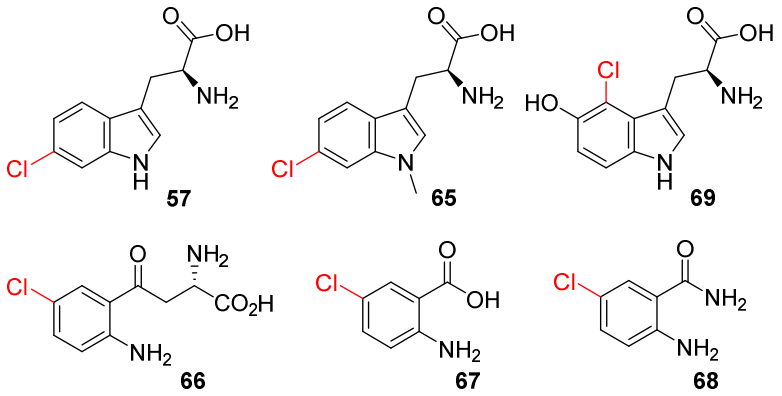


Figure 12: Regioselectivity of chlorinated products obtained from Th_Hal catalysed biotransformations.

For these small compounds, 5-substitution is revealed by the fine doublet splitting of H_A to H_B, which in turn has a larger splitting to H_C (**Figure 13**). Identification of this splitting pattern in the ^{1}H NMR spectra of each of the chlorinated products from **66** to **68**, and comparison of these spectra to literature, supports their characterisation as 5-chlorinated products. Chlorination at the 3-position is excluded by this splitting pattern, which would give rise to an apparent triplet due to the coupling of both H_A and H_C to H_B. Chlorination at the 4-position, which would give rise to a similar splitting pattern to 5-chlorination, was excluded on the basis of the *meta*- substitution proceeding via an unfavourable Wheland intermediate and comparison of the ^{1}H NMR chemical shifts to literature values for these compounds.

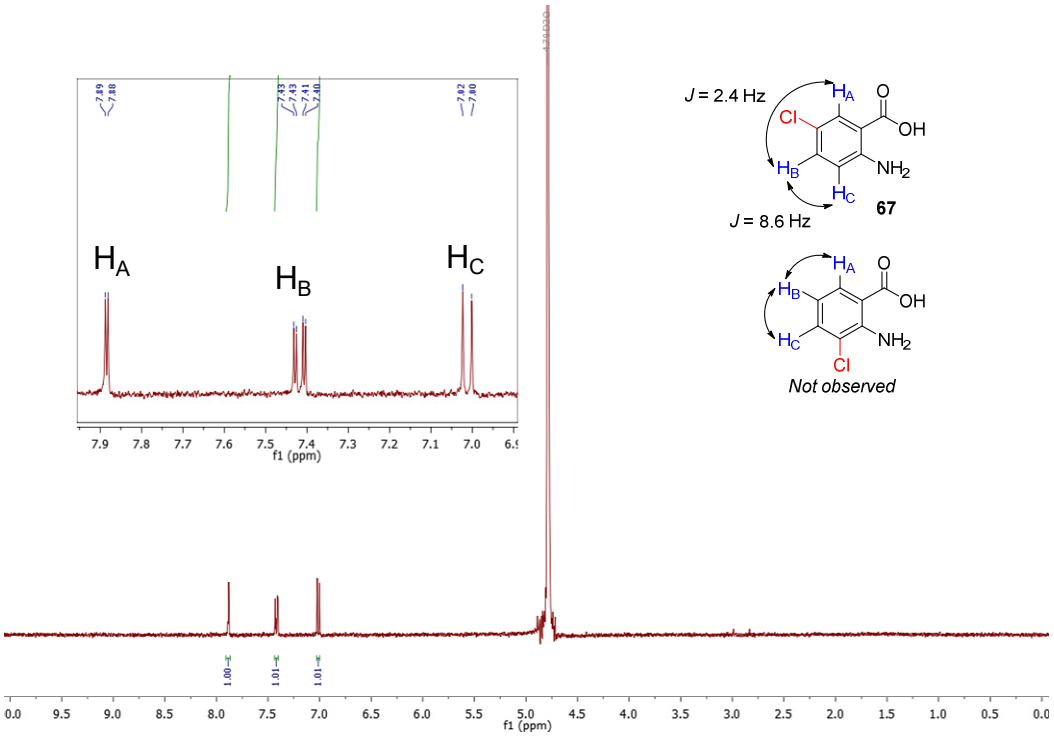


Figure 13: ^{1}H NMR of 5-chloro anthranilic acid (**67**) (D_2O, 400 MHz) with insert showing expansion between δ 6.9 and 8.0 ppm. The proposed coupling pattern of 3-chloro anthranilic acid is also displayed.

Surprisingly 5-hydroxy tryptophan (**59**) was chlorinated at the 4-position. This could be due to the increased stabilisation of a benzoquinone intermediate, or deprotonation of the hydroxyl group leading to increased electron density at the position *ortho*- to the OH group. Such halogenation patterns have been observed most frequently with phenolic FlHals.[76] Given that Th_Hal afforded **69** in a manner which would be expected by a phenolic FlHal, a number of other phenolic substrates were tested (**Figure 14**). These were selected based on our group's work on RadH, the phenolic halogenase responsible for *ortho*-chlorination *en route* to radicicol, which had found RadH to be particularly effective at halogenating **70** to **73**.[171] HPLC and LC-MS analysis of these reactions did not suggest that halogenation of any of these phenolic substrates occurred when incubated with Th_Hal as above, however. This suggests that the substrate binding modes of Th_Hal are much more similar to those of the tryptophan halogenases rather than the phenolic halogenases. Indeed sequence alignment of of Th_Hal with RadH and Rdc2 revealed only 34.5 % and 36.5 % sequence similarity respectively.

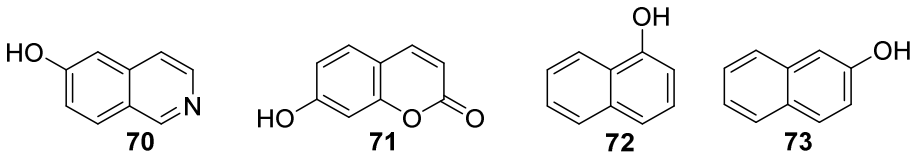


Figure 14: Phenolic substrates found not to be chlorinated by Th_Hal by HPLC or LC-MS analysis of analytical-scale reactions using the same conditions as prior.

2.2.3 Temperature Stability of Th_Hal

The activity of Th_Hal at higher temperatures was subsequently determined, to assess the feasibility of a high temperature biocatalytic halogenation. Assays using Th_Hal and Th_Fre for the chlorination of tryptophan at between 15 °C and 70 °C showed chlorination activity of Th_Hal to be optimal between 35 °C and 45 °C (**Figure 15A**). Over the same temperature range, the mesophilic FlHal PyrH was found to display optimal chlorination activity at around 30 °C however. Comparision of Th_Hal to the other mesophilic FlHals PrnA, RebH and SttH also showed Th_Hal to posess the highest conversion to chlorotryptophan at 45 °C (**Figure 15B**). The demonstration that Th_Hal activity at elevated temperatures was greater than that of the mesophilic FlHals encouraged the idea that Th_Hal may indeed be a stabilised FlHal variant.

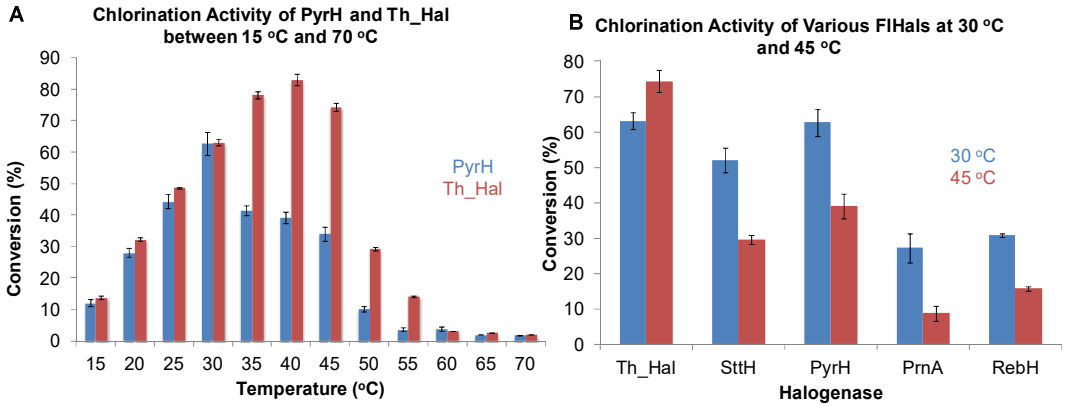


Figure 15: (**A**) Temperature-activity profiles of PyrH and Th_Hal. Conditions as prior except reactions were quenched after 30 min. (**B**) Chlorination activity of various FlHals at 30 °C and 45 °C. Conditions as prior, using the chlorination of tryptophan (**31**) as test reaction.

Timecourse reactions, plotting the conversion of tryptophan to chlorotryptophan by Th_Hal and PyrH over time, demonstrated that Th_Hal also possesses increased catalyst lifetime (**Figure 16**). At 30 °C, conversion of tryptophan by PyrH started to plateaux after around 300 min, whilst Th_Hal continued to convert substrate until the timecourse endpoint at 18 hrs. At 45 °C, this effect was more pronounced with conversion by PyrH slowing significantly after approximately 120 min.

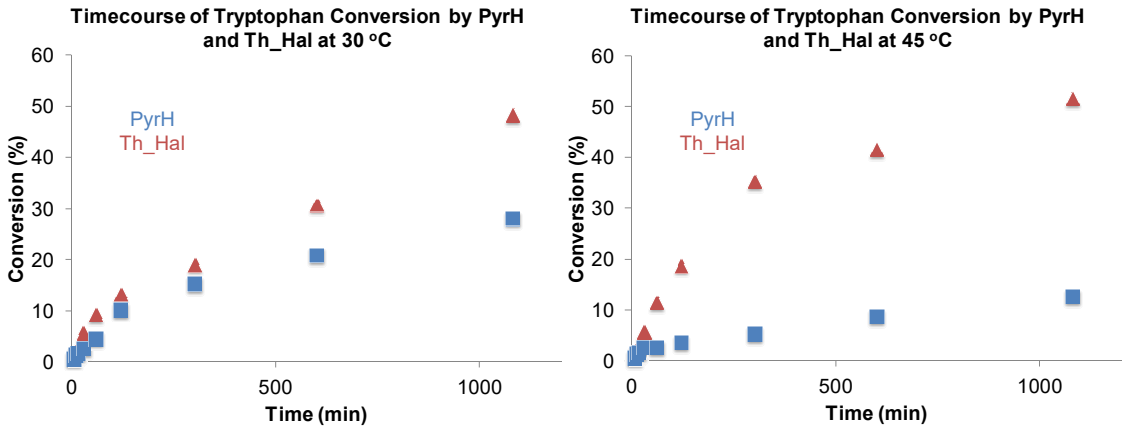


Figure 16: Timecourse of tryptophan chlorination by PyrH and Th_Hal at 30 °C and 45 °C. Conditions as prior except a substrate concentration of 1.0 mM was used to allow longer reaction courses to be measured.

In order to quantify the extent of Th_Hal stabilisation, Binuraj Menon determined the melting temperatures of Th_Hal and the other mesophilic tryptophan halogenases (**Figure 17C**). This technique involves using circular dichromism (CD) spectroscopy to determine the extent of protein unfolding in solution. In our case, the CD spectrum of each halogenase was measured between 20 °C and 80 °C at 2 °C intervals (**Figure 17A**). Plotting the mdeg of the CD spectra at a specified wavelength against temperature then gives an indication of at which temperature the protein begins to unfold and adopt a linear, disorganised, state. After fitting this data to a sigmoidal curve (**Figure 17B**), the temperature at which the protein is between the highly organised, folded, state and highly disorganised, denatured, state can then be determined – termed the melting temperature (T_m) of the enzyme. By using this process, it was found that the T_m of Th_Hal is approximately 10 °C above that of the other tryptophan halogenases tested (**Figure 17C**). The thermophilic flavin reductase Th_Fre was similarly found to have a T_m 9 °C higher than that of the mesophilic Fre from *E. coli*.

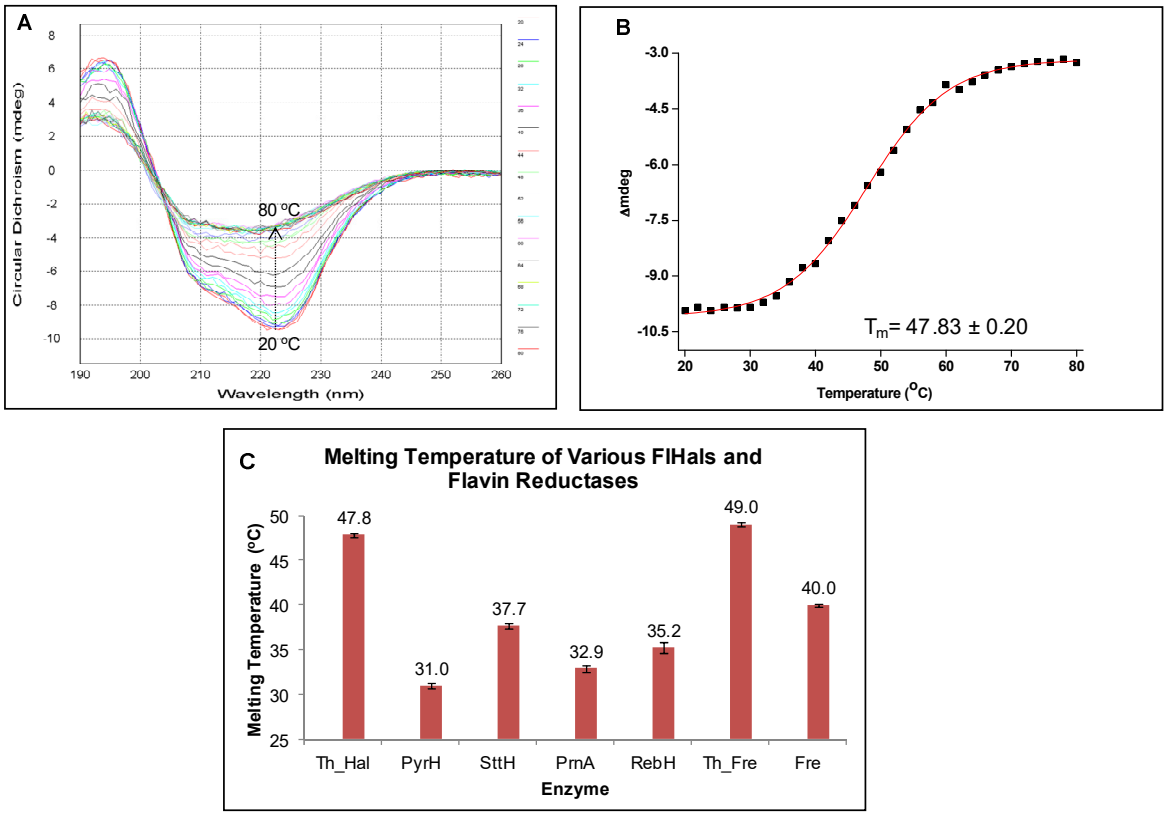


Figure 17: (**A**) Representative CD spectra of Th_Hal recorded between 20 °C and 80 °C at 2 °C intervals. (**B**) Δmdeg at 222 nm of Th_Hal plotted versus temperature and fitted to a sigmoidal curve. (**C**) Melting temperature of various Fl_Hals and flavin reductases. Data provided by Binuraj Menon.

2.2.4 Structure of Th_Hal

In addition to the academic interest of understanding the structural features of Th_Hal which allow stabilisation, such knowledge could guide the engineering of other FlHals into enzymes more amenable to application in biocatalysis. Comparison of the amino acid sequences of Th_Hal and SttH, which have 75.6 % sequence identity, revealed that the differences between the two are spread through the entire length of the sequence (**Figure 18**). The understanding of how these mutations influence stability based purley on sequence would therefore be limited. An apo crystal structure of Th_Hal was therefore obtained by Binuraj Menon, using PrnA (PDB 2AQJ) as the molecular replacement model, to allow such rationalisation.

```
SttH        NTRNPDKVVI VGGGTAGWMT ASYLKKAFGE RVSVTLVESG TIGTVGVGEA
THal        -L-NN--VVI VGGGTAGWMT ASYLKAAFGD RIDITLVESG HIGAVGVGEA

SttH        TFSDIRHFFE FLDLREEEWM PACNATYKLA VRFQDWQRPG HHFYHPFEQM
THal        TFSDIRHFFE FLGLKEKDWM PACNATYKLA VRFENWREKG HYFYHPFEQM

SttH        RSVDGFPLTD WWLQNGPTDR FDRDCFVMAS LCDAGRSPRY LNGSLLQQEF
THal        RSVNGFPLTD WWLKQGPTDR FDKDCFVMAS VIDAGLSPRH QDGTLIDQPF

SttH        DERAEEPAGL TMSEHQGKTQ FPYAYHFEAA LLAEFLSGYS KDRGVKHVVD
THal        DEGADEMQGL TMSEHQGKTQ FPYAYQFEAA LLAKYLTKYS VERGVKHIVD

SttH        EVLEVKLDDR GWISHVVTKE HGDIGGDLFV DCTGFRGVLL NQALGVPFVS
THal        DVREVSLDDR GWITGVRTGE HGDLTGDLFI DCTGFRGLLL NQALEEPFIS

SttH        YQDTLPNDSA VALQVPLDME ARGIPPYTRA TAKEAGWIWT IPLIGRIGTG
THal        YQDTLPNDSA VALQVPMDME RRGILPCTTA TAQDAGWIWT IPLTGRVGTG

SttH        YVYAKDYCSP EEAERTLREF VGPEAADVEA NHIRMRIGRS EQSWKNNCVA
THal        YVYAKDYLSP EEAERTLREF VGPAAADVEA NHIRMRIGRS RNSWVKNCVA

SttH        IGLSSGFVEP LESTGIFFIH HAIEQLVKHF PAGDWHPQLR AGYNSAVANV
THal        IGLSSGFVEP LESTGIFFIH HAIEQLVKNF PAADWNSMHR DLYNSAVSHV

SttH        MDGVREFLVL HYLGAARNDT RYWKDTKTRA VPDALAERIE RWKVQLPDSE
THal        MDGVREFLVL HYVAAKRNDT QYWRDTKTRK IPDSLAERIE KWKVQLPDSE

SttH        NVFPYYHGLP PYSYMAILLG TGAIGLRPSP ALALADPAAA EKEFTAIRDR
THal        TVYPYYHGLP PYSYMCILLG MGGIELKPSP ALALADGGAA QREFEQIRNK

SttH        ARFLVDTLPS QYEYFAAMGQ RV
THal        TQRLTEVLPK AYDYF-T--- Q-
```

Figure 18: Sequence alignment of SttH and Th_Hal. Conserved residues are highlighted in Red and non-conserved residues in black. Generated using the optimal global alignment tool of BioEdit.

The overall structure of Th_Hal revealed that it forms a dimer, as observed with other FlHals, in addition to distinct “box” and “pyramid” domains which form the flavin- and substrate-binding domains of other FlHals respectively (**Figure 19**). Given the structural similarity between Th_Hal and other FlHals, overlay of the Th_Hal secondary structure with the substrate-bound PrnA structure allowed identification of a potential substrate binding pocket of Th_Hal containing lysine and glutamate residues at positions similar to the catalytic residues of PrnA (**Figure 19**).

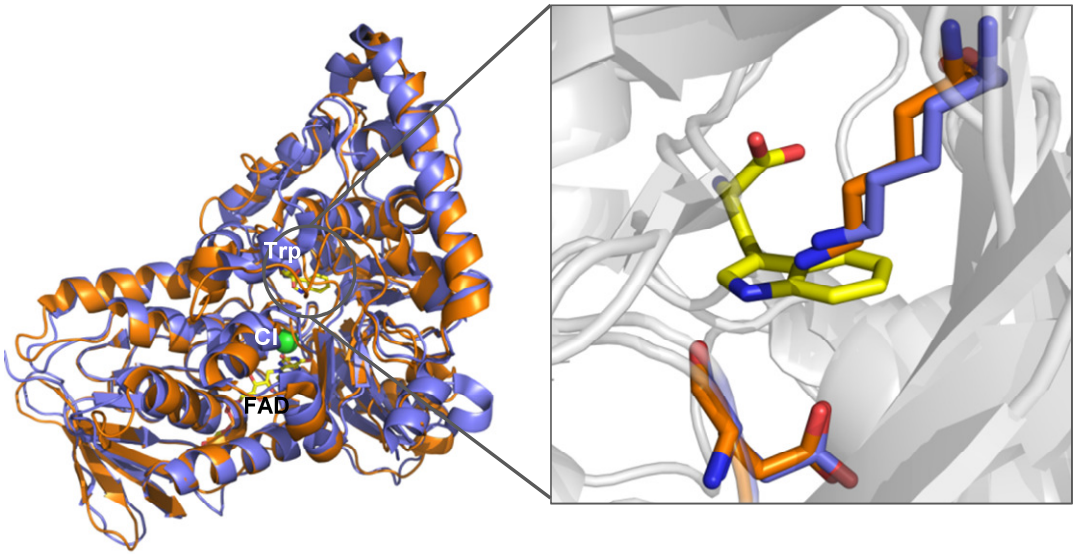


Figure 19: Overall secondary structure of Th_Hal in orange (PDB 5LV9) overlaid with PrnA in blue (PDB 2AQJ). Insert shows the tryptophan binding site of PrnA, with the proposed catalytic residues of Th_Hal and PrnA shown as sticks in orange and blue respectively.

With this region identified as the potential substrate-binding portion, AutoDock was used to dock tryptophan into this binding site in order to give an indication of the binding modes which contribute to the positioning of substrate relative to the proposed catalytic lysine (K75, **Figure 20**).[172] The lowest energy conformation of substrate returned by AutoDock (-3.63 kcal/mol) revealed a number of binding modes similar to those observed in the substrate-bound crystal structure of PyrH (**Figure 5B**).[124] Most notably the proposed catalytic residues of Th_Hal (K75 and E359) are both positioned close to the C-6 position of tryptophan, which would account for the observed regioselectivity. Additionally, the π-system of tryptophan appears to be sandwiched between H92 and F94, stabilising binding through π-stacking and shielding most of the indole C-H positions from halogenation. A number of directional hydrogen bonds are also feasible based on this conformation, with the indole NH within range of a hydrogen-bonding interaction to the backbone carbonyl of P93. The carboxylate and amino residues of the tryptophan side chain may also be able to hydrogen bond to S50 and Y459 respectively. Using the same docking process to rationalise the binding of anthranilic acid (**61**) to the active site did not afford any binding poses consistent with interaction between **61** and K75, suggesting a significant loss of binding energy with smaller substrates capable of forming fewer hydrogen bonds.

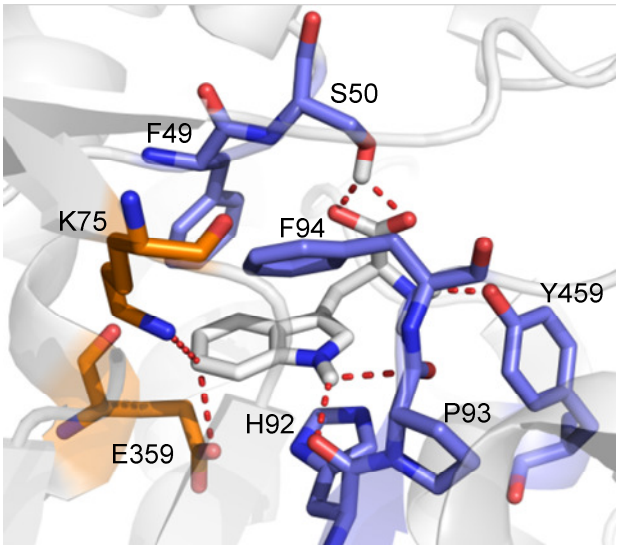


Figure 20: Docked model of tryptophan in the proposed Th_Hal active site generated using AutoDock 4.2.[172] The proposed catalytic residues are highlighted in orange and those involved in binding in blue.

To determine the structural features which directly contribute the increased stability of Th_Hal compared to SttH, apo- crystal structures of the two were overlaid (**Figure 24**, PDB 5LV9 and 5YH5 respectivley). This revealed that most of the residues mutated in Th_Hal are present on the surface of the protein, with the majority of the inner catalytic "core" conserved between the two. None of the mutations were found to reside on the interface of the FAD and substrate binding domains. The localisation of mutations to the surface of the protein may have come about in nature because the mutation of surface residues, rather than those involved in catalysis of inter-domain interactions, is more likely to retain catalytic activity during evolution towards increased thermostability.

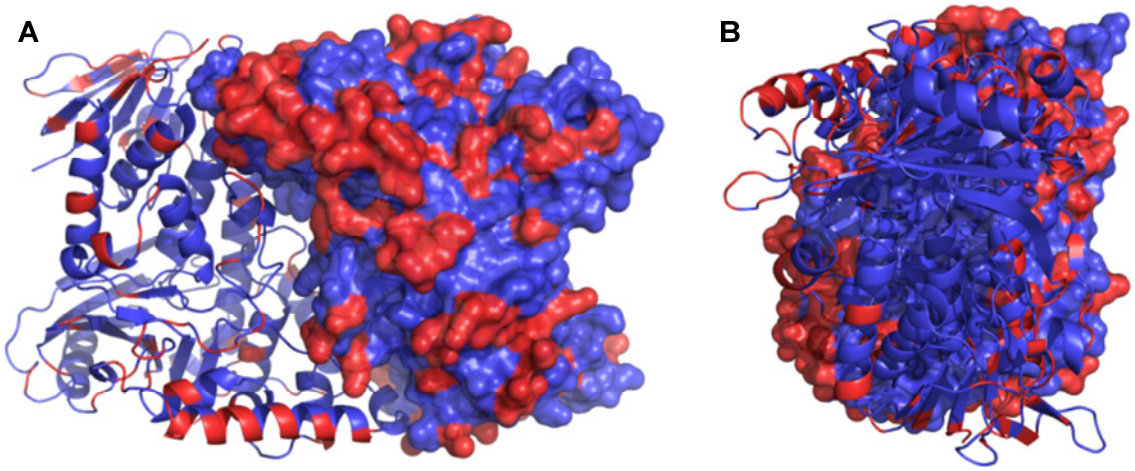


Figure 21: Two views of the apo crystal structure of SttH (PDB H5Y5) with residues mutated in Th_Hal highlighted in Red, showing localisation of the mutations to the surface of Th_Hal (**A**), rather than the inter-domain region (**B**). One monomer of the crystal structure is shown as a cartoon, and the other as a molecular suface for clarity.

Further analysis of the surface residues mutated in Th_Hal showed that most are to residues of increased polarity or charge. Increasing surface charge in this manner has been observed to lead to increased stability with FlHals, in addition to a number of other enzymes.[164,173] For example an engineered RebH variant was found to possess an arginine residue on the surface in place of a glutamine.[164] It is postulated that increased surface charge deters aggregation through electrostatic repulsion and could lead to the formation of increased hydrogen bonding interactions to solvent – thereby stabilising the enzyme in solution.[173]

2.3 Conclusions and Outlook

There have been several attempts to generate FlHal variants which possess increased thermostability and catalyst lifetime, in order to create enzymes more amenable to application as biocatalysts for aromatic halogenation. Previous work adopted the approach of directed evolution using mesophilic halogenases as the starting point and temperature as the selection pressure.[164,165] Using this approach, a RebH variant with T_m 18 ^oC higher than wild-type and a ThaL variant with T_m 10 ^oC higher than wild-type was obtained. These are both impressive improvements, and demonstrate the utility of directed evolution for biocatalyst engineering.

This chapter describes an alternative approach, however. A putative halogenase was identified from a thermophilic *Streptomyces* strain, which was then demonstrated *in vitro* to act as a tryptophan 6-halogenase. The substrate scope was also briefly explored, and found to follow what would be expected based on its high sequence similarity to SttH. Although the thermophilic halogenase (Th_Hal) was found to be more active at higher temperatures, and for longer reaction periods, than the previously reported mesophilic FlHals tested, chlorination activity with non-natural substrates was markedly lower than has been reported for other FlHals. Additionally a number of substrates demonstrated to have some activity with the closest homologue SttH were not chlorinated at all by Th_Hal.[156] Similar effects have been noted with other thermophilic FlHal variants, whereby increased thermostability does not correlate well to increased activity with some substrates. This is proposed to be because the increase in stability is concurrent with a reduction of enzyme flexibility. Such loss of flexibility could therefore reduce the ability of Th_Hal to modify its conformation to allow binding of non-natural substrates. This limited substrate scope is the main factor limiting the application of Th_Hal as a biocatalyst. Given the progress which has been made with modifying the substrate scope of FlHals using both structure-guided and random mutagenesis however, this enzyme may prove a useful starting point for other engineering efforts.[155,158,162]

The melting temperature of Th_Hal was also determined to allow comparison to other FlHals, and the associated engineered variants. This found Th_Hal to have a melting temperature approximately 10 ^oC above the other FlHals tested. This increase in T_m is comparable to the increase obtained from a single round of directed evolution using ThaL and slightly more than half of the increase obtained from three rounds of directed evolution upon RebH.[164,165] Although greater increases in melting temperature could be obtained from directed evolution approaches, it is worth noting that these methods require significant time and materials to reach this point. Using a thermophilic variant as a starting point for such efforts could therefore afford variants of similar

stability with fewer rounds of screening required. In the two examples using directed evolution, the stabilised variants contained two and eight mutations (ThaL and RebH respectively), whilst Th_Hal has only 75.6 % sequence identity to its closest homologue SttH. Crystallisation of the stabilised RebH variant revealed that some of the mutations result in an increase of surface charge,[164] a factor which is thought to contribute to the stabilisation of Th_Hal based on crystal structure analysis.

The fact that both methods of identifying thermostable variants afforded similar results suggest that the two approaches can be used complementarily, with these outcomes being used to direct the further engineering of FlHals. A recent report used a combinatorial codon mutagenesis approach to target random mutations to the FAD-binding portion of RebH, with limited improvement of catalytic activity.[174] Use of this approach to target mutagenesis to the surface portions of mesophilic FlHals which are mutated in Th_Hal may afford enzymes more appropriate for biocatalysis. Alternatively other thermophilic halogenase enzymes could be identified in the same way and then used as the starting point for targeted or random mutagenesis. Indeed several potential candidate genes from organisms with increased optimum growth temperature have been identified by Binuraj Menon (**Table 3**).

Table 3: Thermophilic organisms which may which possess genes with reasonable sequence identity to Th_Hal.*Based upon comparison with Th_Hal amino acid sequence using the optimum global alignment tool of BioEdit.

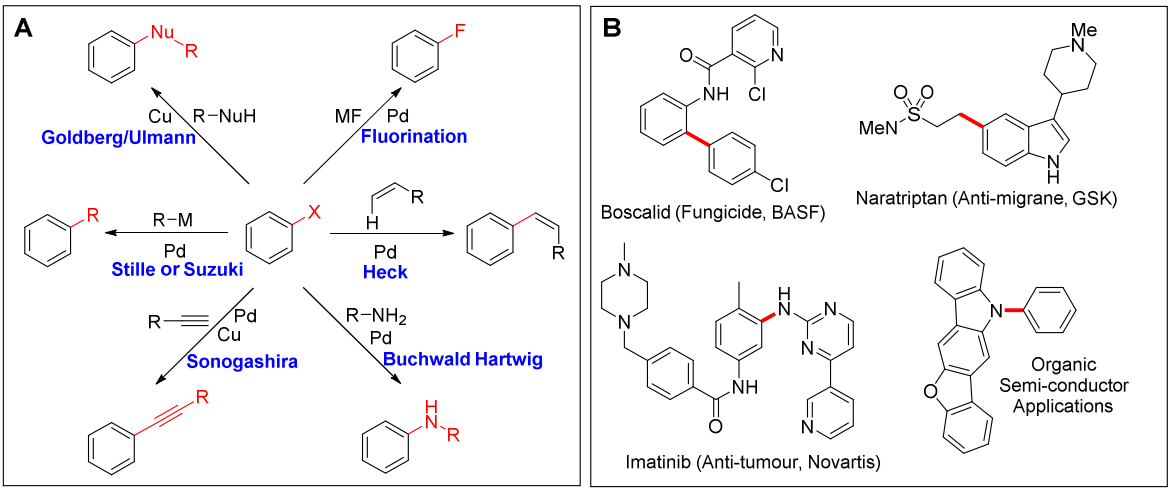
Organism	Identity (%)*	Optimum Growth Temperature (^{o}C)	Accession Number
Streptomyces thermoauthtrophicus	61	65	KWX04309.1
Microbulbifer thermotolerans	36	43 – 49	AMX03406.1
Rubellimicrobium thermophilum	36	45 – 54	WP_021098953.1
Sulfobacillus thermosulfidooxidans	21	50 – 55	WP_020375004.1

The increased stability of Th_Hal, rather than its activity at elevated temperatures, is arguably more useful in biocatalysis. As the FlHals require both NADH and FAD as cofactors, both of which are relatively unstable, the productivity of a high temperature biocatalytic halogenation is likely to ultimately be limited by these factors. Additionally, kinetic studies of Th_Hal found that the K_m of tryptophan (**31**) binding to Th_Hal was higher at 45 ^{o}C than at 30 ^{o}C.[175] This is logical, since binding of substrate to Th_Hal has an entropic penalty which will be more pronounced at higher temperatures. Since a thermophilic NADH recycling enzyme variant could not be obtained in the timescale of this study, preparative scale halogenations using Th_Hal required the use of stoichiometric amounts of NADH – an expensive cofactor – which therefore limits the utility of this process.

3. Halogenase-Controlled C-C Bond Formation

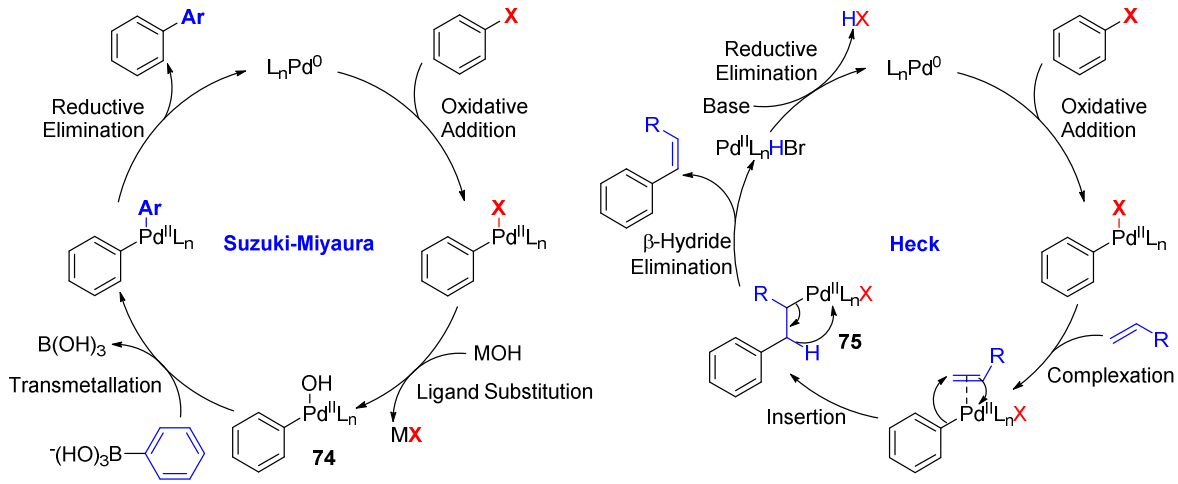
3.1 Introduction

Palladium-catalysed cross-coupling reactions play a key role in modern organic synthesis. The plethora of C-C and C-heteroatom couplings which are possible mean that they have become indispensible for the assembly of complex molecular scaffolds.[46,47] As such, reactions of this type are now found in synthetic routes towards compounds of relevance to all sectors of the chemical industry.[47,176,177] **Scheme 25** shows the versatility of these couplings, as well as a number of compounds of commercial relevance which are made on scale using Pd-catalysed C-C or C-N cross couplings.



Scheme 25: (**A**) Various transition metal-catalysed reactions utilising organohalogen compounds as substrates. All are shown as aryl halides for clarity, although many also accept other organohalogens. X = Cl, Br, I, OTf. (**B**) Compounds of industrial relevance synthesised using palladium-catalysed cross-coupling chemistry, with the bond formed by this chemistry highlighted in red.

Key to many of these reactions is the oxidative addition of Pd^0 across a C-halide bond (**Scheme 26**), meaning that organohalogen compounds have become indispensible intermediates in synthesis. Addition of Pd^0 across the C-halide bond generates a Pd^{II} intermediate which is generally believed to then undergo ligand substitution with a base to form intermediate **74**. In Stille, Suzuki-Miyaura and Sonogashira couplings, this intermediate then undergoes transmetallation with a second stoichiometric organometallic reagent. The resultant organopalladium species, bearing the organic portion of both coupling partners, then undergoes reductive elimination to regenerate Pd^0 and form the C-C bond.[46] The Heck reaction is thought to proceed in a similar manner, except that the Pd^{II} species formed from oxidative addition is then thought to insert across the C=C of substrate alkene to form a Pd-σ intermediate (**75**) .[46] The C=C is then restored through β-hydride elimination of this intermediate, followed by regeneration of Pd^0 through reductive elimination of halide (**Scheme 26**).

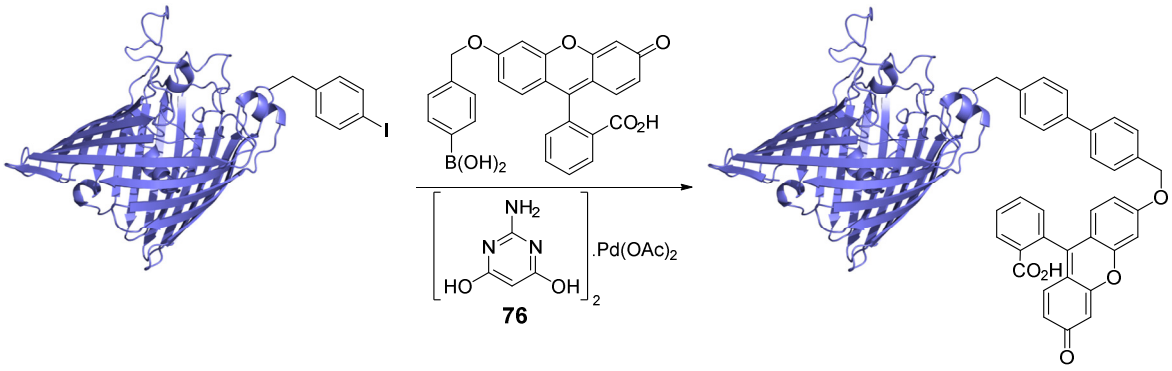


Scheme 26: Proposed catalytic cycles of the Suzuki-Miyaura and Heck reactions.[46] Note that the Stille, Sonogashira and Buchwald-Hartwig couplings operate via a mechanism analogous to the Suzuki-Miyaura reaction, except the nature of the transmetallation step varies.

3.1.1 Palladium Catalysis in Bio-orthogonal and Bio-compatible Conditions

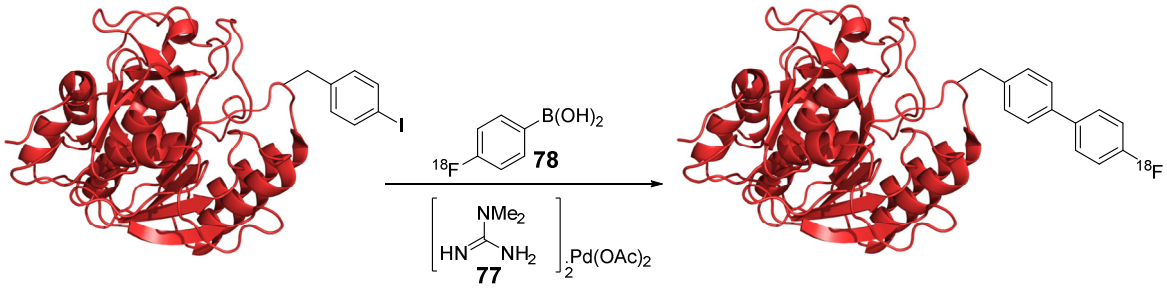
Since the initial discovery of these palladium-catalysed reactions, there has been a significant amount of work focussed on developing methods for conducting them under benign reaction conditions. The proliferation of palladium catalysis throughout industrial synthesis, due to the versatility and power of these transformations, means that methodologies with minimal environmental impact are required for sustainable chemical manufacture. These efforts have led to conditions for Pd-catalysed couplings which are mild enough to allow the modification of sensitive biological molecules, such as proteins or nucleic acids, or to allow the cross-coupling to occur within biological systems.[45]

One particular seam of work in the field of palladium-mediated biological molecule modification was based upon the identification of the pyrimidine $Pd(OAc)_2$ complex **76** as an effective catalyst for aqueous Suzuki-Miyaura couplings (SMCs),[178] which was based upon the application of this complex for Sonogashira couplings in organic solvents.[179] The conditions required for this cross coupling were sufficiently mild to allow the SMC of iodoinated aromatic amino acids, introduced into a model protein through derivatisation of surface cystine residues, to occur at 37 ^oC in pH 8.0 phosphate buffer. Efficient modification of the model protein required the addition of a vast excess (500 eq) of boronic acid and the palladium catalyst (50 eq).[178] This is likely due to the background deborylation of boronic acids in water, or the non-selective conjugation of electrophilic boronic acids with nucleophilic surface residues. Such surface residues could also coordinate to palladium, thereby preventing efficient catalysis.[178] Despite these issues, the ability to chemically modify a protein surface through orthogonal palladium catalysis was sufficient to encourage the application of this method to the modification of aryl iodide amino acids, introduced into a surface protein of *E. coli* cells through amber codon suppression, with a fluorescent boronic acid (**Scheme 27**).[180] In this example, similar incompatibility of the palladium catalyst and boronic acid with unidentified elements of the biological system meant that a vast excess of boronic acid was similarly required - necessitating the thorough washing of cells prior to visualisation of cell surface fluorescence.



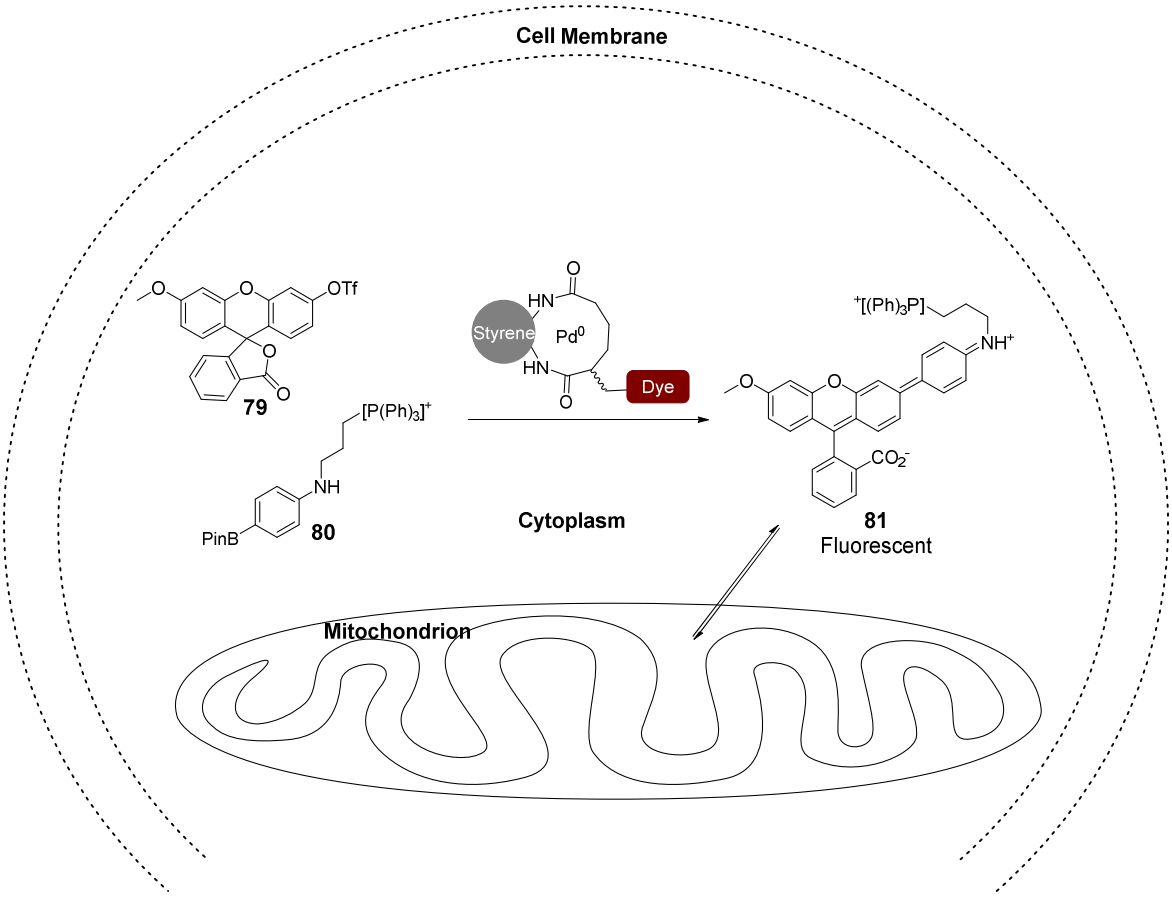
Scheme 27: Suzuki-Miyaura cross-coupling of non-natural iodophenyl alanine surface residues of OmpC (PDB 2J1N) at 37 °C using a pyrimidine-based palladium catalyst.[180]

An improvement of this approach was subsequently reported, which allowed the use of lower loadings of Pd and fewer equivalents of boronic acid, through use of the guanidine-based catalyst **77**. Because of the increased efficiency of this coupling, it was feasible to use the ^{18}F-labelled boronic acid **78** for the modification of a surface iodo phenylalanine residue of a target protein (**Scheme 28**), as well as a number of other peptides.[181] When labelling the protein however, a palladium loading of 10 equivalents was employed. This suggests some manner of deactivation or coordination of palladium to the protein, which limits the utility of such an approach for medical applications because of the likelihood of palladium contamination in the biological therapeutic.



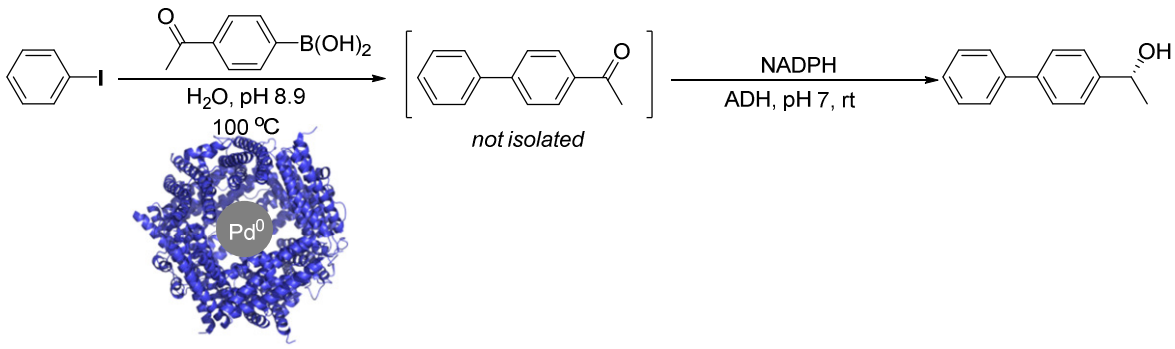
Scheme 28: Suzuki-Miyaura cross-coupling of non-natural iodophenyl alanine surface residues for ^{18}F labelling of SBL (PDB 1NDQ) at 37 °C using a guanidine-based palladium catalyst.[181]

The incompatibility of palladium catalysts to certain biological molecules was overcome in one example through the immobilisation of Pd^0 nanoparticles within styrene beads.[45] The tolerance of this particular catalyst allowed it to be used to catalyse non-natural reactions occurring within HeLa cells. In this example, the SMC product of the two non-fluorescent coupling partners **79** and **80** affords a product with fluorescent properties sufficiently distinct to allow the visualisation of product **81** within the biological system (**Scheme 29**). Conjugation of a fluorescent moiety with an excitation maxima distinct from **81** onto the styrene carrier of the Pd catalyst, allowed the localisation of Pd^0 catalyst within the cells to be visualised also. This revealed the Pd^0 catalyst and the fluorescent SMC adduct **81** to be visualised in the cytosol and mitochondria of the HeLa cells respectively.



Scheme 29: Intracellular fluorogenic palladium-catalysed Suzki-Miyaura Coupling.[45]

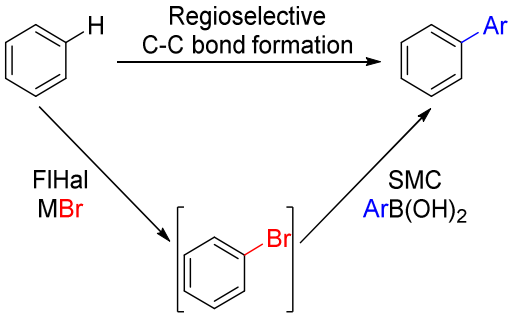
An alternate approach to circumvent the incompatibility of palladium catalysts and enzymes has involved the immobilisation of Pd^0 nanoparticles within the binding site of an enzyme.[182] In this example (**Scheme 30**) Pd^0 nanoparticles were immobilised within a DNA binding protein from a thermophilic organism. This preparation was found to be effective at catalysing the SMC of aryl iodides in water at elevated temperature. By coordinating the metal catalyst specifically within a protein "core", it is effectively prevented from interacting with other enzymes in the system – as with the examples of iridium catalysis described in Chapter 1. This allowed the subsequent addition of an ADH enzyme and NADH for enantioselective reduction of a ketone moiety (**Scheme 30**), without the need to remove palladium to prevent deactivation of ADH - which was reported in earlier attempts at this transformation.[26]



Scheme 30: Use of protein-immobilised Pd^0 nanoparticles in an integrated Suzuki-Miyaura Alcohol Dehydrogenase (ADH) reaction.[182]

3.1.2 Conclusions

Transition-metal catalysed cross-coupling reactions have become a cornerstone of organic synthesis because of their generality and utility. Many of these reactions require an aryl halide as one coupling partner, because of the ability of Pd^0 to undergo oxidative addition across the C-halide bond. Since the generation of these aryl halide intermediates is often fraught with issues of poor regioselectivity and the need for forcing conditions or toxic reagents, it was envisioned that the combination of the FlHal enzymes discussed herein with palladium-catalysed cross-coupling chemistry could provide a method for the direct, regioselective, formation of C-C bonds from C-H (**Scheme 31**). Given the relative stability and solubility of boronic acids with respect to water, and the precedent for aqueous Suzuki-Miyaura cross couplings (SMCs), the SMC was selected for the first demonstration of this combination.

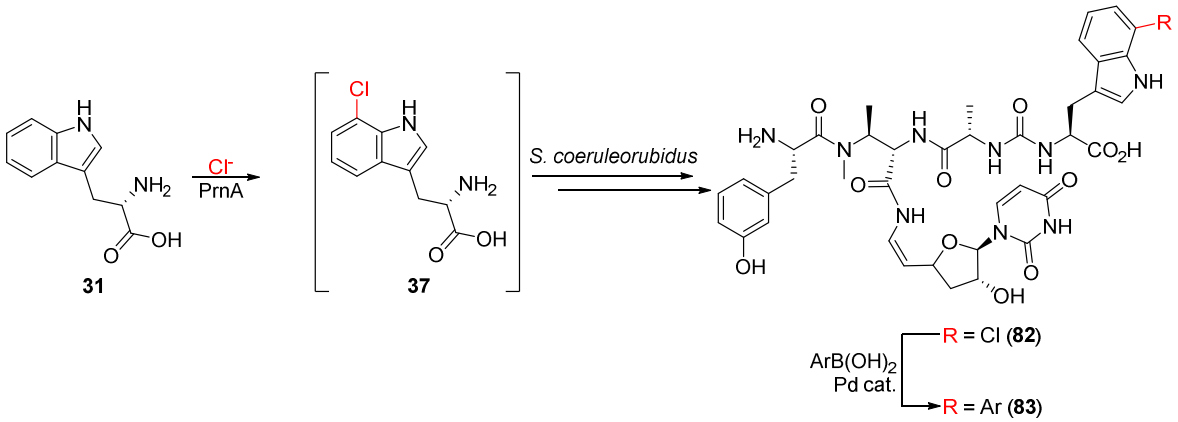


Scheme 31: Proposed combination of flavin-dependent halogenases (FlHal) with Suzuki-Miyaura cross-couplings (SMCs) to allow regioselective C-C bond formation.

Prior to this study, other groups had attempted such combinations for the generation of natural product analogues. In this work, the gene encoding the tryptophan 7-halogenase PrnA was introduced into the pacidamycin-producing organism *Streptomyces coeruleorubidus*.[183] This NRPS assembly line had previously been found to accept 7-chlorotryptophan (**37**) in place of tryptophan through substrate-directed biosynthesis,[184] and therefore the *in situ* chlorination of tryptophan by heterologously-produced PrnA within the pacidamycin producer resulted in the formation of chloropacidamycin (**82**, **Scheme 32**).[183] This approach to producing chlorinated non-natural products using FlHals has been exploited by a number of other groups also, and offers the

advantage of allowing the direct production of natural product analogues by fermentation without the need for additional synthetic steps or laborious total synthesis.[185-187]

In one example, the chloro substituent obtained from such biosynthetic engineering efforts was then utilised as a reactive handle for further derivatisation through SMC chemistry, generating a number of arylated pacidamycin analogues (**83**, **Scheme 32**).[183] A similar rationale has since been used to modify halogenated natural products generated through substrate-directed biosynthesis.[188] This work demonstrates the potential utility of combining FIHals and palladium catalysis, because of the ability to regioselectivley incorporate the reactive moiety for metallation and subsequent functionalisation. The complex mixtures of products and media components obtained from such fermentations however necessitated the extraction and partial purification of chloropacidamycin prior to the SMC reaction.[183] These intermediary processes reduce the efficiency of such a process, and ultimately make it less attractive from the green perspective.[3]



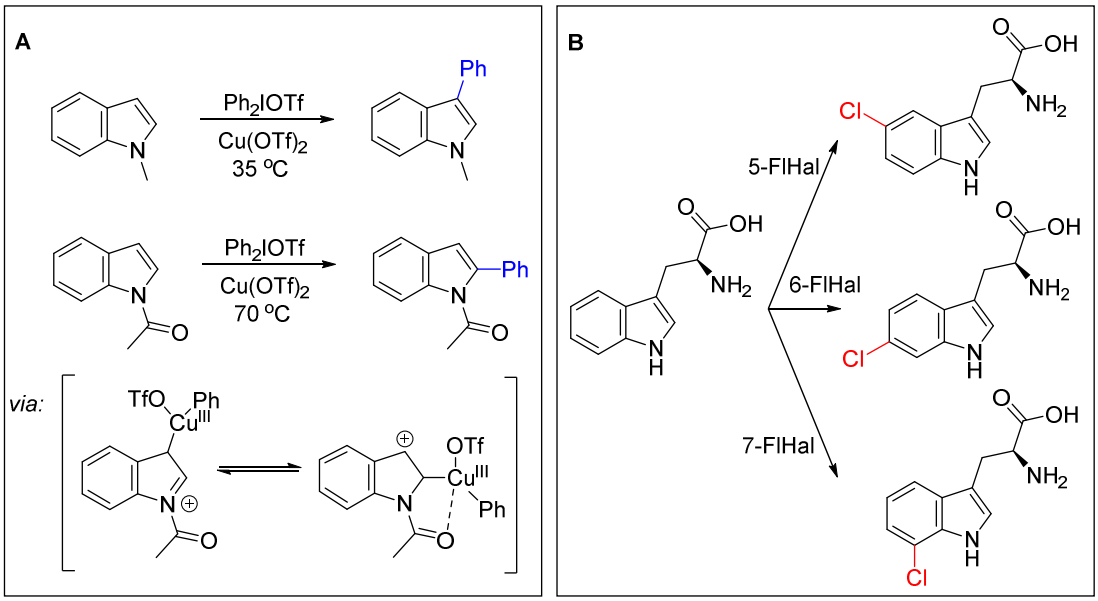
Scheme 32: Incorporation of the tryptophan 7-halogenase PrnA into the Pacidamycin-producing organism *S. coeruleorubidus* to produce chloropacidamycin (**82**) directly by fermentation. Subsequent arylation of the non-natural product (**82**) with a Suzuki-Miyaura coupling allowed further diversity to be generated.[183]

With this in mind, it was hoped to develop an integrated catalysis approach for the conversion of aromatic C-H to C-C bonds using FIHals and palladium-catalysed cross-coupling reactions in single pot transformations. It was envisioned that the regioselective and regio-divergent biocatalytic brominations offered by FIHals could allow similar control of C-C bond formations. Such a transformation (**Scheme 31**) could be viewed as a C-H activation reaction since, formally, inert C-H bonds are selectively activated and converted to C-C. These approaches are receiving increased attention in organic chemistry, since the ability to directly activate C-H bonds can lead to more step-economical synthetic routes without the need for preceding functionalisation reactions.

The indole moiety has been of particular interest in the field of C-H activation, because the number of similar C-H positions means that the selective control of these individual positions is a particular challenge.[67] Seminal examples of the selective arylation of indole C-H positions relied on approaches classed substrate control, whereby electronic or coordinating effects of the substrate control the position of metallation. For example the innate nucleophilicity of the indole C3 was

found sufficient to promote selective metallation of this position with electron-deficient d^8 Cu catalysts – thereby allowing arylation at this position of N-methylated indoles (**Scheme 33A**).[189]

Through introduction of a coordinating acetyl group onto the indole nitrogen however, this metallated intermediate was postulated to undergo rearrangement upon heating – thereby activating the C2 indole position for arylation (**Scheme 33A**).[189] The remaining C-H positions of indole were beyond the scope of C-H activation at the time, given their lack of innate reactivity and positioning relative to the other positions which could be activated. Tryptophan FlHals on the other hand can directly and selectively halogenate the C5, C6 and C7 positions of indole (**Scheme 33B**), through catalyst rather than substrate control, and therefore appeared a promising strategy to extend indole C-H activation beyond the known arylation patterns. This chapter will describe the attempts made to integrate the FlHals with palladium-catalysed cross-coupling reactions in single pot transformations, including studying the tolerance of each type of catalyst and the methods employed to overcome deactivation.



Scheme 33: (**A**) Cu-catalysed C-H activation of the indole C3 and C2 positions facilitated by substrate control, showing the rearrangement process thought to allow C2 activation.[189] (**B**) C-H activation using flavin-dependent halogenase (FlHal) enzymes to confer catalyst control. FlHals which exhibit different regioselectivity are discussed in detail in the introduction to Chapter 2.

3.2 Results and Discussion

3.2.1 Cross-Coupling Reaction Screening

To allow the development of effective FlHal-SMC cascades, FlHal biotransformations which allowed the efficient generation of intermediate aryl halides were required. Although the bromination of tryptophan by FlHals with different regioselectivity has been demonstrated by a number of groups to be robust *in vitro*, low concentrations of brominated product are often afforded.[155,157,159] This is thought to be due to a combination of poor enzyme stability and substrate or product inhibition. Additionally, many protocols for the SMC of amino acids require protection of the free amine moiety – presumably to prevent coordination of this group to Pd which could inhibit

productive SMC cycles.[190] It therefore seemed appropriate to identify non-natural halogenase substrates for the initial screening of SMC conditions. Our group had previously found PyrH, a tryptophan 5-halogenase, to be effective at the regioselective 5-chlorination of anthranilamide (**62**) at relatively high substrate concentrations.[155] A glucose dehydrogenase (GDH) was also employed for the concurrent recycling of NADH using glucose as cosubstrate, allowing catalytic quantities of NADH to be used. Application of these conditions, utilising NaBr as halide source in place of $MgCl_2$, fortunately afforded high conversion of anthranilamide (**62**) to a brominated product - identified by liquid chromatography high-resolution mass spectrometry (LC-HRMS) (**Figure 22**). Organobromine compounds are readily identified using mass spectrometry (MS) by the presence of peaks separated by 2 *m/z* units at an almost 1:1 ratio, due to the natural abundance of ^{79}Br and ^{81}Br. Subsequent upscale of this reaction allowed isolation of the brominated product, which was identified as 5-bromo anthranilamide (**84**) by ^{1}H NMR.

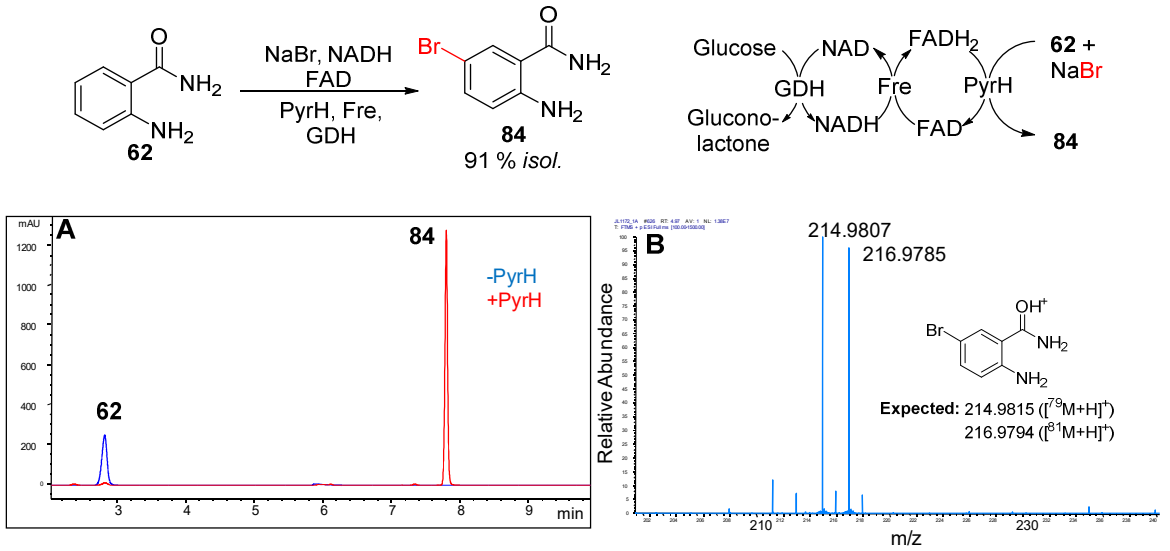
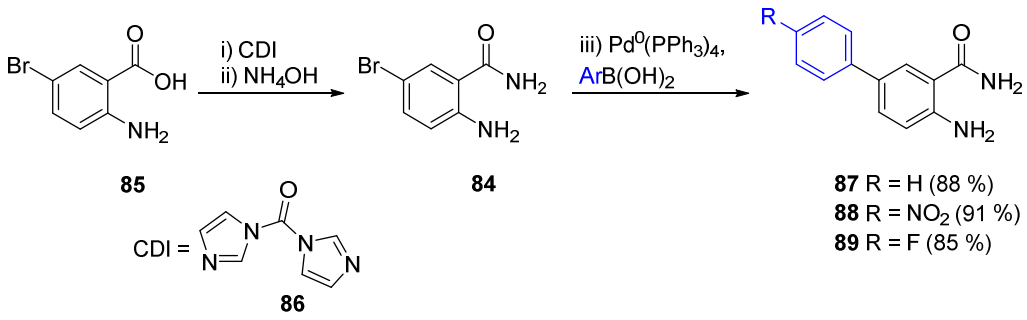


Figure 22: (**A**) HPLC chromatogram monitored at 280 nm resulting from the incubation of anthranilamide (**62**) with and without PyrH. (**B**) HRMS of the product peak resulting from incubation of **62** with PyrH showing the distinctive bromine isotope pattern. Conditions: anthranilamide (2.0 mM), NaBr (100 mM), NADH (100 µM), FAD (1 µM), PyrH (20 µM), Fre (2 µM), GDH (12 µM), rt, overnight.

With the knowledge that PyrH could be utilised to afford high concentrations of 5-bromo anthranilamide, the SMC of this product with phenyl boronic acid was selected as the test reaction for screening SMC conditions (**Figure 23C**). To allow the conversion of such reactions to be accurately determined on an analytical scale, standards of 5-bromo anthranilamide and a number of 5-aryl anthranilamide derivatives were prepared using literature methods (**Scheme 34**).[191] The commercially available 2-amino-5-bromobenzoic acid (**85**) was converted to **84** though activation with 1,1'-carbonyldiimidazole (**86**) followed by treatment with saturated ammonium hydroxide solution. The intermediate bromide was then arylated using conditions reported for the SMC of anilines.[192] This sequence afforded **87** to **89** in good yields (**Scheme 34**), which were used to construct calibration curves for analytical HPLC.



Scheme 34: Synthetic approach to standards of 5-bromo anthranilamide (**84**) and 5-arylated anthranilamides (**87** to **89**). Isolated yields are shown in brackets. Conditions: i) CDI (1.1 eq), THF, N_2, rt, 6 hrs, 74 % yield; ii) saturated NH_4OH, rt, overnight; iii) $Pd^0(PPh_3)_4$ (10 mol %), $ArB(OH)_2$ (2.6 eq), K_2CO_3 (9.0 eq), PhMe, H_2O, EtOH, N_2, reflux, overnight.

Based on previous reports of aqueous SMC reactions, the pyrimidine and guanidine ligands **L1** to **L4** in combination with $Pd(OAc)_2$ were selected for an initial catalyst screen (**Figure 23B**). The dimethylamino primidine **L1** was readily prepared from the guanidine **L2** and diethyl malonate (**Figure 23A**). **L1** to **L4** were then pre-complexed with $Pd(OAc)_2$ according to literature protocols.[178,181]

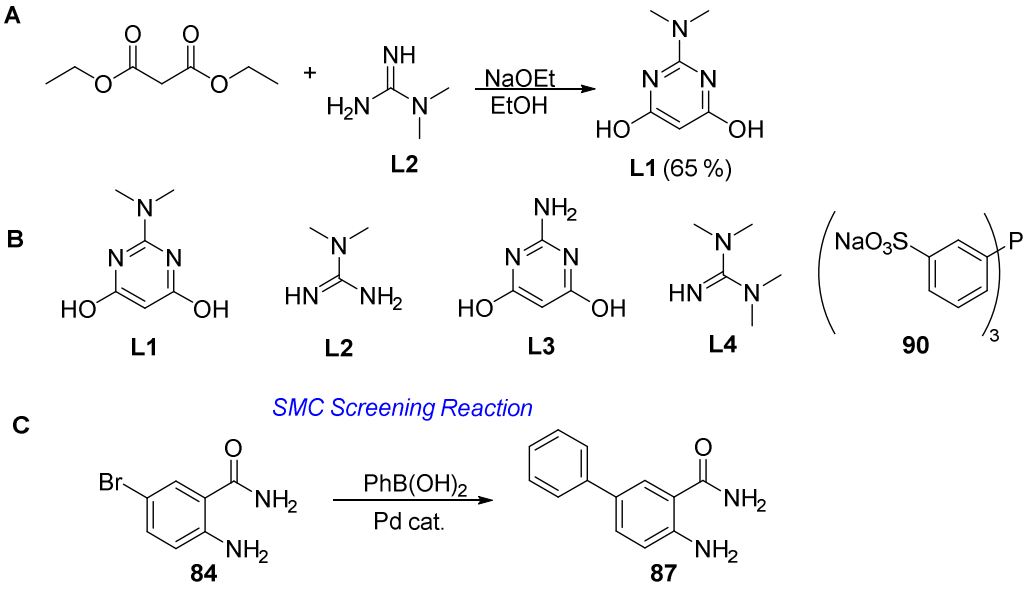


Figure 23: (**A**) Synthesis of **L1**. Conditions: **L2** (1.0 eq), sodium ethoxide (3.0 eq), diethyl malonate (1.0 eq), ethanol, reflux, 5 hrs. (**B**) Ligands screened in the test SMC of **84** and ligand identified by Humera Sharif (tppts). (**C**) Test SMC reaction of **84** with phenyl boronic acid.

Table 4: Screening of the pyrimidine and guandidine ligands **L1** to **L4** in the test SMC reaction (**Figure 23C**). Conditions: **84** (2.0 mM), $Pd(OAc)_2.L_2$ (10 mol %), CsF (5.0 eq), $PhB(OH)_2$ (5.0 eq), KPi (pH 7.2), overnight. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs. † Isolated yield = 81 %.

Ligand	Temperature (^oC)	Conversion (87, %)*
L1	30	6
L2	30	<1
L3	30	<1
L4	30	5
L1	80	86†
L2	80	56
L3	80	32
L4	80	32

Incubation of these complexes with 5-bromo anthranilamide (**84**) and phenyl boronic acid in water with CsF at 30 ^oC afforded only trace amounts of product after incubation overnight. Considering that oxidative addition may be a limiting factor, the same screen was subsequently repeated at 80 ^oC and found reactions using ligand **L1** to afford 86 % conversion to product **87** (**Table 4**). Upscale of these conditions allowed the product to be isolated in a promising 81 % yield and identified as **87** through comparison of spectral data to the standard prepared above. A parallel study by Humera Sharif subsequently found that the combination of Na_2PdCl_4 with tppts (**90**, triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt) was also effective at this coupling under similar conditions.[193]

3.2.2 Suzuki-Miyaura Integration with Halogenation through MWCO Filtration

3.2.2.1 Optimisation

With two catalyst systems for the test SMC in hand, their tolerance to the FlHal biotransformation conditions was assessed. Both the cofactors NADH and FAD, required for generation of $FADH_2$ and HOCl respectively, are rich in heteroatoms. It was therefore considered that they may strongly coordinate to Pd species and prevent their interaction and desired reactivity with substrates. Additionally, given the redox activity of both co-factors, it is possible that they could cause the formation of catalytically inactive Pd species through redox reactions. The deactivation of Pd catalysts by proteins has been reported by several other groups and is considered to be due to the coordination of soft donor atoms on the surface of proteins to Pd catalysts.[178] Indeed, some have exploited the coordinating effect of metal catalysts to surface cystine residues to template reactions on the surface of proteins.[194] To assess the extent of SMC deactivation by each of these factors, the test SMC reaction was conducted in the presence of each component using both catalyst systems (**Table 5**).

Table 5: Effect of FlHal biotransformation additives on the efficiency of the test SMC (**Figure 23C**). Conditions (A); **84** (2.0 mM), $Pd(OAc)_2$.**L1**$_2$ (10 mol %), CsF (5.0 eq), $PhB(OH)_2$ (5.0 eq), KPi (pH 7.2), 1 hr. (B) **84** (2.0 mM), Na_2PdCl_4 (2.5 mol %), tppts (5.0 mol %), K_3PO_4 (2.4 mM), $PhB(OH)_2$ (2.0 eq), 1 hr. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs.

Entry	Conditions	Additive	Conversion (87, %)*
1	A	None	30
2	A	NADH (0.1 mM)	15
3	A	FAD (1 μM)	19
4	A	Fre (1 μM)	25
5	A	GDH (6 μM)	<1
6	A	Glucose (40 mM)	10
7	B	None	21
8	B	NADH (0.1 mM)	9
9	B	FAD (1 μM)	13
10	B	Fre (1 μM)	10
11	B	GDH (6 μM)	2
12	B	Glucose (40 mM)	17

This screen revealed that the co-factors NADH and FAD had minimal impact on SMC conversion, whilst the presence of protein caused a significant reduction in conversion to **87**. The observation that the SMC reactions gave lowest conversion in the presence of GDH, which is employed at higher concentration than the other protein Fre, suggests that the source of deactivation could be through coordination of Pd to either surface residues or the hexahistidine tags used for protein purification. An increased loading of Pd was therefore employed to essentially saturate Pd-binding portions of the protein – ideally leaving some in solution to catalyse SMC reactions. Whilst this allowed reasonable conversions to be obtained in the presence of NADH and FAD, no such improvement was observed in reactions containing protein (**Table 6**).

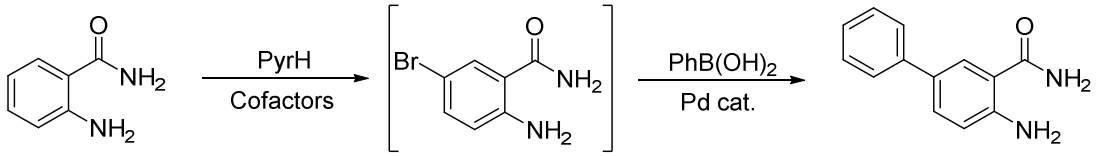
Table 6: Effect of increasing Pd loading on the tolerance of the test SMC (**Figure 23C**). Conditions as prior except both are left overnight. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs.

Conditions	Additive (Concentration)	Conversion (87, %)*			
		2.5	5	10	20 mol %
A	None	3	62	93	78
A	NADH (0.1 mM)	3	14	17	14
A	FAD (1 μM)	12	19	54	90
A	Fre (1 μM)	4.7	11	4	<1
A	GDH (6 μM)	<1	2	8	2
A	Glucose (40 mM)	10	55	93	52
B	None	21	40	62	9
B	NADH (0.1 mM)	9	26	30	3
B	FAD (1 μM)	13	34	38	21
B	Fre (1 μM)	10	39	39	15
B	GDH (6 μM)	2	4	3	2
B	Glucose (40 mM)	17	35	61	63

Given that unsatisfactory conversion to **87** was observed even when an increased Pd loading was employed, it was decided that the proteins required for halogenation must be somehow removed prior to addition of SMC components, or compartmentalised separately from them. To determine the effectiveness of this approach, the test reaction (**Scheme 35**) was run under a number of different circumstances:

1. All components of the FlHal biotransformation and SMC reaction were added simultaneously, with heating once the halogenation reaction had completed
2. The halogenation was allowed to proceed at 30 °C until completion, before addition of the SMC components and heating at 80 °C
3. Once the halogenation reaction had completed, protein was denatured and removed by centrifugation before addition of the SMC components

Test FlHal-SMC Reaction



62 **84** **87**

In situ

Scheme 35: Test FlHal-SMC reaction used for screening. FlHal conditions: antrhanilamide (**62**, 2.0 mM), NADH (100 μM), FAD (1 μM), NaBr (100 mM), glucose (20 mM), PyrH (20 μM), Fre (2 μM), GDH (12 μM), 20 °C, overnight. SMC conditions as stated in the text.

Table 7: Effect of altering the order of SMC components addition and protein removal by denaturation (see text) upon conversion of the test FlHal-SMC reaction. (**Scheme 35**). Conditions A and B as prior using 20 mol % Pd catalyst. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs.

Conditions	Reaction Mode	Conversion (87, %)*
A	1	<1
A	2	<1
A	3	<1
B	1	N.C
B	2	<1
B	3	1

Each of these conditions afforded trace to nil conversion to arylated product however (**Table 7**), suggesting that either protein removal was inefficient or that a non-protein component was responsible for inhibition of the SMC. It was also considered that the poor conversion could be due to a cumulative effect from the presence of all the inhibitory cofactors at once. Prolonged reaction times and further increased loadings of palladium catalyst afforded a slight increase in conversion to arylated product (**Table 8**, entries 3 and 7). Above 30 mol % of either palladium catalyst seemed to cause a decrease in conversion to arylated product however (**Table 8**, entries 4,5, 9 and 10), with a concurrent increase in the amount of reduced product (i.e. anthranilamide) formed – suggesting that transmetallation may have become the yield-determining step.

Table 8: Effect of increased Pd catalyst loading upon conversion of the integrated FlHal-SMC test reaction (**Scheme 35**) when protein was removed by denaturation (mode 3). Conditions A and B as prior. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs.

Entry	Conditions	Pd Loading (mol %)	Conversion (87, %)*
1	A	10	5
2	A	20	23
3	A	30	26
4	A	40	18
5	A	50	11
6	B	10	1
7	B	20	14
8	B	30	13
9	B	40	12
10	B	50	13

Simply increasing the number of boronic acid equivalents did not afford any meaningful increase in conversion, however. In order to eliminate the possibility that poor conversion was due to inefficient protein removal by denaturation, thereby allowing optimisation to continue in more productive directions, FlHal biotransformations were passed through 10 kDa molecular-weight cut-off (MWCO) filters prior to addition of the SMC components.

Reactions conducted in this manner were found to give significantly increased conversion to arylated product compared to those where denatured protein was removed by centrifugation (**Table 9**, entries 7 and 10). Further optimisation of palladium and boronic acid loading afforded a satisfactory conversion of 83 % when using Na_2PdCl_4 and tppts as the palladium source and ligand respectively (**Table 9**, entry 12). Similar increases were not observed when the pyrimidine-based catalyst was used, suggesting incompatibility with a component other than protein. Use of the optimised conditions involving Na_2PdCl_4 and tppts, however, allowed the regioselective arylation of anthranilamide (**62**) to 5-phenyl anthranilamide (**87**) in a 79 % yield without isolation of the intermediate aryl halide. Spectral data of the proposed arylated product was consistent with that obtained for the standard of **87** obtained herein and that reported in the literature.

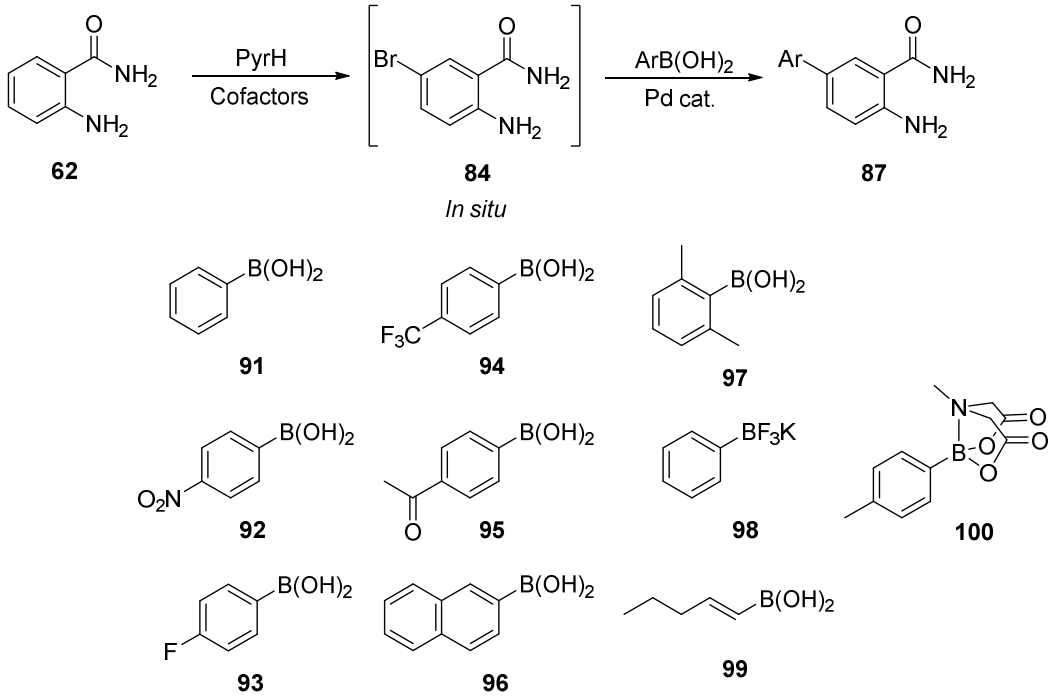
Table 9: Conversion of the test integrated SMC reaction (**Scheme 35**) with higher loading of Pd and increased equivalents of boronic acid. Conditions A and B as prior. Removal method 1 refers to the precipitation of protein after the PyrH-catalysed biotransformation. Removal method 2 refers to the filtration of PyrH-catalysed biotransformations through 10 kD MWCO filters. *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs. † Isolated yield = 79 %.

Entry	Conditions	Pd loading (mol %)	Equivalents $PhB(OH)_2$	Removal Method	Conversion (87, %)*
1	A	40	20	1	13
2	A	40	30	1	11
3	A	50	30	1	14
4	A	40	20	2	18
5	A	40	30	2	12
6	A	50	30	2	6
7	B	40	20	1	30
8	B	40	30	1	24
9	B	50	30	1	15
10	B	40	20	2	60
11	B	40	30	2	32
12	B	50	30	2	83†

3.2.2.2 Substrate Scope

With an initial set of conditions for the regioselective arylation of anthranilamide (**62**) in hand, the scope of this process with respect to boronic acid coupling partner was determined on an analytical scale. A range of aromatic boronic acids including electron-withdrawing (**92** to **94**) and electron-donating substituents (**95** to **97**) were screened with the optimised conditions using the MWCO membrane filtrate of a test reaction using anthranilamide as substrate (**Scheme 35**). 1-Penten-1-yl boronic acid (**99**) was also tested, in addition to a trifluoroborate (**98**) and MIDA (*N*-methyliminodiacetic acid) boronate ester (**100**) coupling partners. Of these, only the aromatic boronic acids bearing electron-withdrawing substituents (**Figure 24**) gave reasonable conversion to arylated products, as did the vinyl example. Since electron-deficient aryl boronic acids deborylate faster in basic water,[195] the higher conversions in these examples suggest that transmetallation is

sufficiently facile in these cases to compensate. The electron-rich and sterically hindered coupling partners **95** to **97** gave very poor conversion to arylated product, whilst use of the potassium trifluoroborate salt (**98**) did not afford any arylated anthranilamide (**87**). Since transmetallation of **95** to **97** would be expected to be slower, it is likely that deborylation becomes competitive in these cases.[195]



Coupling Partner	Conversion (%)*
91	83
92	98
93	72
94	45
95	20
96	7
97	<1
98	N.C
99	35
100	30

Figure 24: Screening of the integrated FIHal-SMC reaction using MWCO filtration with a number of organoboron coupling partners. Conditions B as prior (**Table 9**). *Determined by analytical HPLC and based on the consumption of **62** and **84** using calibration curves of these compounds. Average of two runs. N.C. = no conversion.

Such a tight substrate spectrum suggests that the poor conversion and requirement for large excesses of boronic acid, even in the optimised example, may be due to a competitive reaction of the organoboron donor with a FIHal reaction component which is not removed by the MWCO filtration. The reaction of boronic acids with the diol moieties of sugars, for example, is a well

known.[196] Since glucose is utilised in excess in the FlHal reaction for the recycling of NADH (**Figure 22**), there is a possibility that the organoboron coupling partner could be sequestered as glucose esters (**Figure 25**). Both of the redox co-factors NADH and FAD both contain diol moieties in the sugar portion also, which could form esters with the organoboron coupling partner (**Figure 25**). Since the formation of these esters is a pH-dependent equilibrium in water, and because boronate esters are known to be reactive under SMC conditions,[196] it is possible these factors are inconsequential in this case and therefore poor conversions may stem from the interaction of these components with the Pd catalysts.

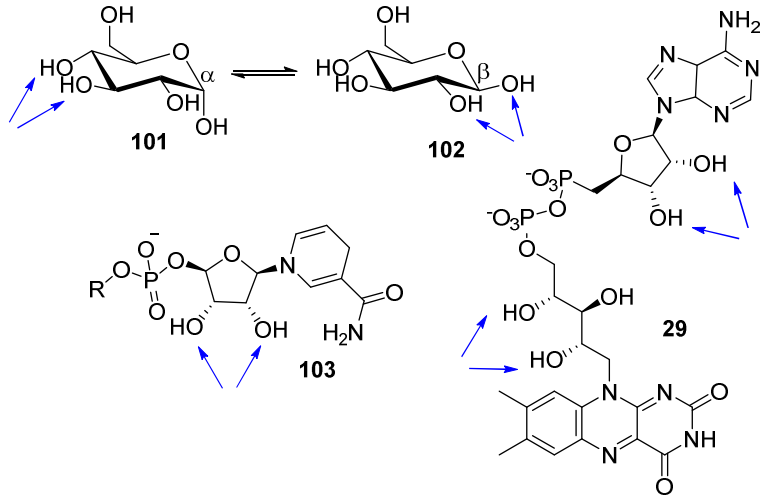
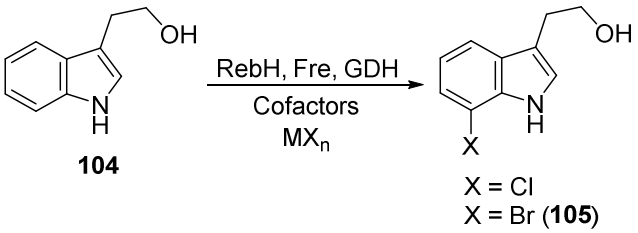


Figure 25: Structure of the α- (**101**) and β-D-pyranose (**102**) forms of glucose, NADH (**103**) and FAD (**29**) with diol moieties potentially capable of forming boronate esters with boronic acids highlighted with blue arrows.

With the scope of organoboron coupling partners accepted by this process seemingly very limited, focus was turned to determining the substrate scope of the process with respect to the types of aromatic scaffold which could be regioselectivley arylated. With the goal of regioselective indole arylation in mind, the tryptophan 7-halogenase RebH was selected based on previous reports of its effective halogenation of a number of indoles.[159] Of these, tryptophol (**104**) was selected as the most suitable for SMCs because of the lack of free amine groups. Expression and purification of RebH as previously described by our group allowed the 7-bromination of this substrate to be optimised on a small scale. This found an increased concentration of FAD to be required for effective halogenation compared to PyrH (**Table 26**) - consistent with previous observations about chlorination using RebH, suggesting that the binding affinity of RebH for $FADH_2$ is lower than the other FlHals. Upscale of this reaction allowed the isolation of 7-bromo tryptophol (**105**) in good yields however, using the GDH recycling system for NADH.

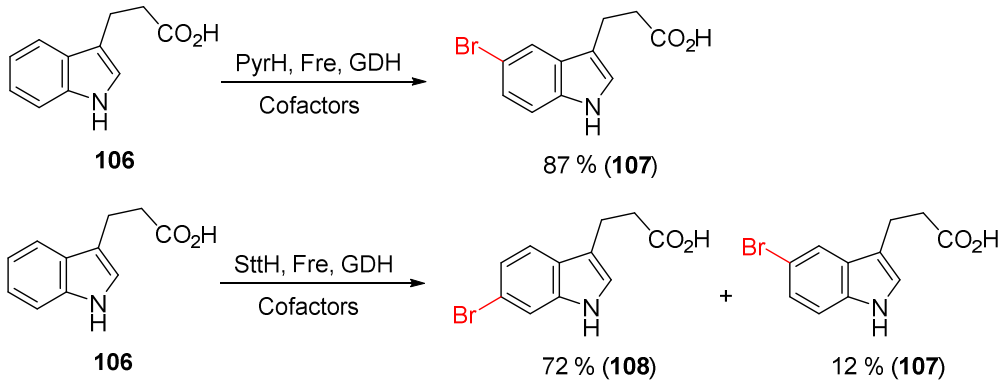


104

Halide Source	[FAD] (μM)	[MX_n] (mM)	[RebH] (μM)	Conversion (%)*
$MgCl_2$	100	10	25	51 (Cl)
NaBr	100	10	25	<5 (Br)
NaBr	100	100	25	12 (Br)
NaBr	100	1000	25	N.C. (Br)
NaBr	10	100	25	N.C. (Br)
$MgCl_2$	100	10	50	98 (Cl)
NaBr	100	100	50	99† (Br)

Figure 26: Optimisation of the RebH-catalysed bromination of tryptophol (**104**). *Determined by analytical HPLC based upon consumption of starting material. N.C. = no conversion. Analytical scale conditions: tryptophol (**104**, 0.5 mM), NADH (100 μM), Fre (2.5 μM), GDH (5 μM), glucose (20 mM), KPi (pH 7.2), propan-2-ol (4 % v/v), 20 °C, overnight. †Isolated yield = 92 %. Preparative scale conditions: tryptophol (**104**, 2.0 mM), NADH (100 μM), FAD (100 μM), NaBr (100 mM), glucose (20 mM), RebH (50 μM), Fre (5.0 μM), GDH (10 μM), KPi (pH 7.2), propan-2-ol (4 % v/v), rt, overnight.

To access the 5- and 6- positions of the indole ring, PyrH and the tryptophan 6-halogenase SttH were selected. Analytical scale reactions, however, found neither of these halogenases to effectively chlorinate or brominate tryptophol (**104**) under standard conditions. The larger indolic substrate indole-3-propionate (**106**) had however been found to be chlorinated with good conversion by both PyrH and SttH to give the 5- and 6-chloro products respectively.[156] PyrH and SttH fortunately also catalysed the regioselective bromination of this substrate, allowing the 5- and 6-bromo indole-3-propionate (**107** and **108**) products to be obtained in good yield (**Scheme 36**). In the case of the SttH-catalysed bromination, an approximately 9:1 ratio of 6:5 brominated products is obtained.



106

87 % (**107**)

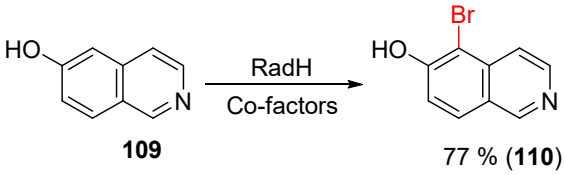
106

72 % (**108**)

12 % (**107**)

Scheme 36: Isolated yields of the regioselective 5- and 6-bromination of indole-3-propionate (**106**) using PyrH and SttH respectively. PyrH and SttH biotransformation conditions were as reported for PyrH earlier, except a FAD concentration of 7.5 μM was used with SttH.

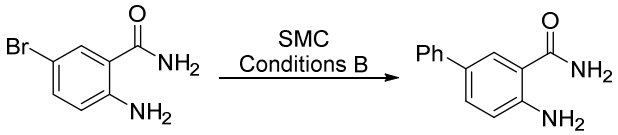
To extend the scope of this process beyond indoles and benzamides, the phenolic FlHal RadH was also investigated. Our group had previously found RadH to regioselectivley chlorinate *ortho*- to the OH group of 6-hydroxy isoquinoline (**109**) with good conversion. Application of these conditions to the bromination of this substrate, using NaBr in place of $MgCl_2$, fortunately afforded good conversion to brominated product which after isolation was identified to be the selectively brominated 5-bromo-6-hydroxy isoquinoline product (**110**) by ^{1}H NMR (**Scheme 37**).



109 77 % (**110**)

Scheme 37: Isolated yield from the regioselective bromination of 6-hydroxy isoquinoline (**109**) catalysed by RadH. Conditions: substrate (0.5 mM), NaBr (10 mM), NADH (2.5 mM), FAD (1.0 µM), RadH (25 µM), Fre (4 µM), KPi, EtOH (1 % v/v), 30 °C, overnight.

Despite many attempts to include a co-factor recycling system into these RadH-catalysed biotransformations, neither GDH nor ADH was found to be appropriate by co-workers in our group.[171] NADH was therefore required in stoichiometric quantities for direct regeneration of $FADH_2$ by Fre. Given that both NADH and FAD had been found to inhibit the SMC of our test substrate in earlier experiments, the requirement for increased concentration of these co-factors for effective halogenation by RadH and RebH respectively was of concern. Fortunately, screening of the test SMC reaction in the presence of the cofactor concentrations required by RadH and RebH found good conversion to arylated product using the optimised SMC conditions (**Scheme 38**), suggesting that the integration of these biotransformations into the FlHal-SMC reaction may be effective.



84 **87**

Entry	[NADH] (µM)	[FAD] (µM)	Conversion (87, %)*
1	0.1	1	92
2	5	1	96
3	0.1	100	93
4	5	100	87

Scheme 38: Effect of increased NADH and FAD concentration, required for effective bromination by RadH and RebH respectively, upon conversion to arylated product **87** using the test SMC reaction. Conditions B as prior (**Table 9**). *Determined by analytical HPLC using a calibration curve of **87**. Average of two runs.

The integrated FlHal-SMC reaction was subsequently conducted using the optimised biotransformation and SMC conditions, with these newly-identified FlHal-catalysed brominations.

The boronic acid coupling partners found to afford highest conversion on an analytical scale were selected to maximise the diversity of products which could be obtained (**Figure 27**). Following initial demonstration of the substrate scope of this process, others in our group carried out a number of other reactions using this regime for further exemplification of this process. These examples are included below for completeness, and the contribution of others noted. Importantly, this method allowed the direct formation of C-C bonds at the indole C5, C6 and C7 positions (**Figure 27**) – transformations which were at the time unprecedented.

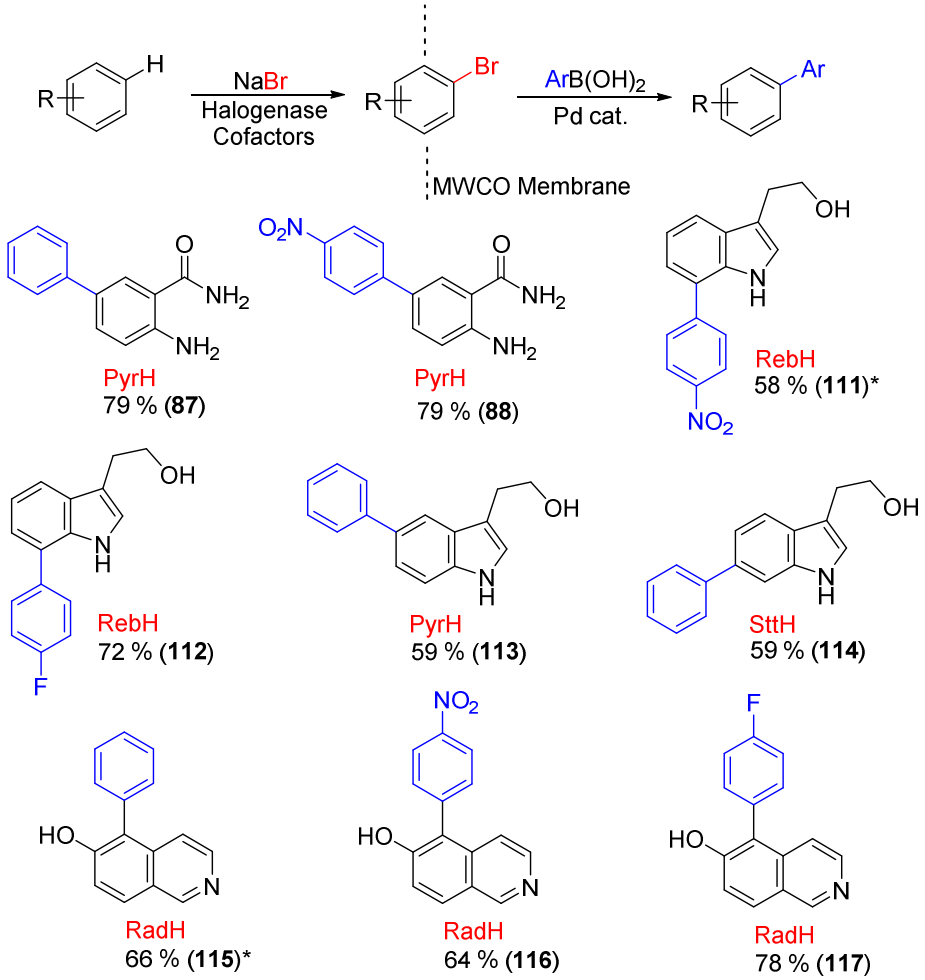


Figure 27: Isolated yields of arylated products obtained from the integration of various FlHal enzymes with SMC reactions using a molecular weight cut-off (MWCO) membrane to compartmentalise the bio- and chemo-catalysts. The enzyme used in each case is shown in red. *Obtained by Jean-Marc Henry. Biotransformation conditions are as described earlier in the text and in the experimental section. SMC conditions: Na_2PdCl_4 (50 mol %), tppts (1.0 eq), $PhB(OH)_2$ (30 eq), base (2.4 eq), KPi, 80 ^{o}C, overnight.

The FlHal-SMC process involving removal of proteins by MWCO filtration allowed a modest selection of arylated products to be obtained in good to moderate yields (**Figure 27**). The main factor limiting the generality of this approach is the narrow substrate scope with respect to boronic acid coupling partners, the reasons for which are discussed above, meaning that only the electron-withdrawing partners **92** and **93** could be used in addition to phenyl boronic acid (**91**). The use of

near stoichiometric quantities of palladium source and ligand also reduces the attractiveness of this process significantly, and suggests that factors other than simply the presence of biocatalyst are having a deleterious effect upon SMC conversion when integrated with FlHal brominations.

3.2.3 Suzuki-Miyaura Integration with Halogenation using Biocatalyst Immobilisation

It was considered that one reason for the requirement to employ such a high Pd loading in the FlHal-SMC reaction using MWCO filtration could be the dilute aryl bromide relative to cofactors and other components of the FlHal biotransformation. Indeed other groups have found dilute SMC reactions in aqueous media, in the presence of biological components, to require very high loadings of Pd catalyst.[178] It was therefore postulated that improving the productivity of FlHals could allow higher concentrations of intermediate aryl bromide to be afforded, thereby allowing lower loadings of Pd catalyst to be used for the SMC.

The poor productivity of FlHals has been reported by many groups, and is thought to be due to a combination of their poor stability and kinetic values (k_{cat} ca. 1 min^{-1}) as well as substrate or product inhibition.[155,164,197] The reason for such slow reactions is likely to be their dependence upon an electrophilic aromatic substitution reaction, which involves an energetically-demanding electrophilic attack as the rate-determining-step.[128] It appeared logical therefore that a biocatalyst preparation with increased catalyst lifetime and stability may allow the generation of higher concentrations of aryl bromide and hence allow efficient SMC reactions with lower Pd loadings. Such a preparation may also improve the scalability of the FlHal-SMC process beyond what is feasible using purified enzymes.

3.2.3.1 Cross-Linked Enzyme Aggregate (CLEA) Preparation

During the course of this study, the group of Sewald reported the RebH-catalysed halogenation of tryptophan on a gram scale.[197] This advance in halogenase productivity and scalability was made possible using a cross-linked enzyme aggregate (CLEA) of RebH with a flavin reductase and ADH. Preparation of a CLEA involves precipitation of the protein of interest from a crude cell lysate, followed by incubation with a short bifunctional molecule which acts as a linker – in this case glutaraldehyde (**118**, **Figure 28B**). The two aldehyde moieties of **118** form imine functionalities with lysine residues on the surface of the protein of interest – therefore forming covalent bonds between individual enzyme molecules and the individual subunits of the protein of interest (**Figure 28A**).[198] In this case, this process was carried out using the FlHal RebH, as well as the Fre and ADH enzymes required for cofactor recycling, to create a combined heterogeneous halogenation biocatalyst which utilises halide source, O_2 (from air) and propan-2-ol as the only stoichiometric reagents.[197]

The improved stability of CLEAs with respect to time is well reported in the literature, and has been used to make many practicable biocatalysts. The cross-linking in such preparations is believed to reduce the rate of enzyme denaturation by replacing weak non-covalent interactions between individual enzyme subunits with stronger covalent bonds.[199] The formation of heterogenous aggregates containing enzyme is also thought to improve productivity by effectively shielding some

enzyme molecules from experiencing high substrate or product concentrations in the bulk solution – essentially creating small microenvironments within the heterogeneous catalyst at lower local substrate and product concentration. Similar effects are also thought to contribute to the stability of these materials with respect to mechanical sheer and organic solvents.[198] The increased stability of the RebH-Fre-ADH CLEA prepared by Sewald's group was found to be sufficient to allow recycling over 40 times, and in one example was used to chlorinate 1 g of tryptophan continuously over a period of 7 days.[197] It was hoped that such improvements in biocatalyst efficiency could be achieved by application of this method to the FlHals utilised in the integrated FlHal-SMC process.

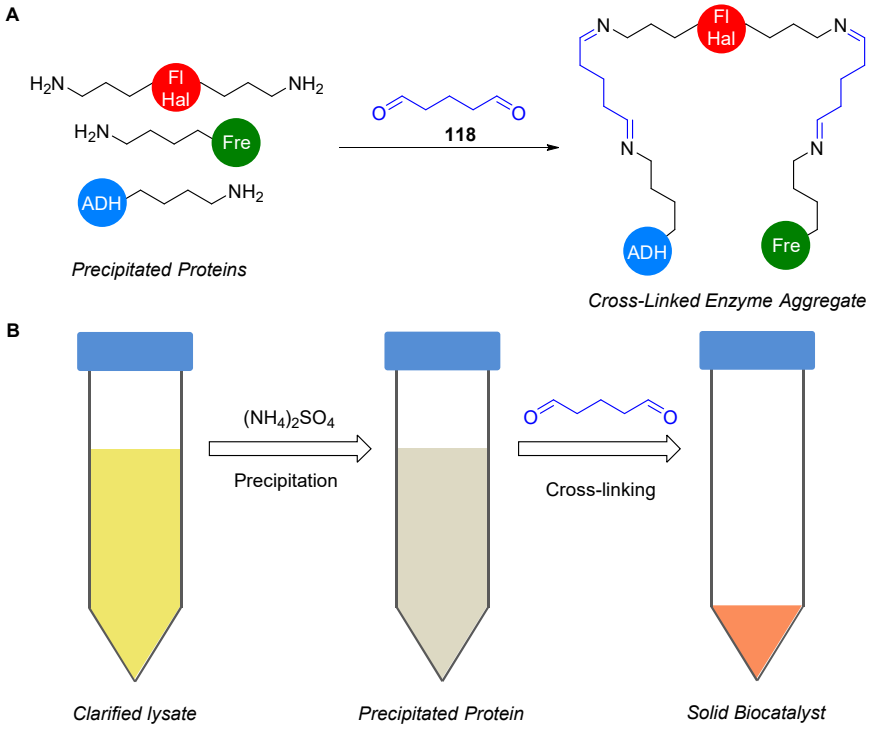


Figure 28: Schematic illustration of the cross-linking process involved in preparation of a cross-linked enzyme aggregate (CLEA) using glutaraldehyde as cross-linker showing (**A**) imine formation between the individual enzymes and (**B**) workflow of CLEA preparation.

As the 5-arylation of anthranilamide (**62**) using PyrH and a GDH-based NADH recycling system had become the standard test reaction for the FlHal-SMC cascade (**Scheme 35**), attempts were made to generate a PyrH-Fre-GDH CLEA following Sewald's reported protocol. Initial trials at this process did not generate an active aggregate, however. This was initially presumed to be due to a difference in lysine content of the two FlHals. Since surface lysine residues are involved in the cross-linking process through the formation of imines (**Figure 28A**), it is possible that too low a concentration of glutaraldehyde during the cross-linking process could prevent efficient cross-linking of the biocatalysts. Since both FlHals also possess an active site lysine however, it is also possible that too high a concentration of glutaraldehyde during the cross-linking process could cause deactivation by capping of this catalytically-essential residue. The concentration of glutaraldehyde utilised during the cross-linking process, as well as cross-linking time, was therefore varied (**Table 10**). A commercially-available ADH enzyme was also utilised in place of GDH to assess if the difference in lysine content of these two enzymes caused the formation of inactive

CLEAs. Neither of these approaches led to the formation of an active PyrH CLEA, however. None of these conditions afforded an active RebH-containing CLEA either, suggesting that the relative lysine content of the enzymes involved was not related in this case.

Table 10: Concentrations of glutaraldehyde and cross-linking times trailed for the formation of PyrH-containing CLEAs. Each combination attempted twice, with either GDH or ADH as the NADH recycling enzyme. The same set of conditions was also trialled for the formation of a RebH CLEA. Activity was assayed under standard conditions using the chlorination of tryptophan as a test reaction. *Optimum conditions reported by Sewald for the preparation of a RebH-Fre-ADH CLEA.[197]

[Glutaraldehyde] (wt/v %)	Cross-linking time (hr)
1.0	2
0.5	2*
0.25	2
0.1	2
1.0	1
0.5	1
0.25	1
0.1	1

Attention was then turned to the system used for expression of the FlHals. Many groups, including our own, have found FlHals to be difficult to recombinantly express in a soluble form using *E. coli*.[163] For this reason, the FlHals are often co-expressed with chaperone proteins to aid protein folding and therefore increase yields of soluble protein. Our laboratory utilise the cold-adapted chapteronins Cpn10 and Cpn60 (arctic express) whilst Selwad reported the use of GroES and GroEL (pGro7). Since CLEAs are prepared from crude cell lysates, the resulting aggregate contains proteins endogenous to the expression system in addition to the protein of interest. The expression level of PyrH relative to these chaperones, as well as other proteins endogenous to *E. coli*, could therefore be a limiting factor. The relative expression level of PyrH in both the arctic express and pGro7 systems, in addition to BL21 containing no chaperones, was therefore assessed by SDS-PAGE of cleared cell lysates. This suggested that the level of PyrH expression in BL21 containing the pGro7 chaperones was significantly higher than in arctic express, by comparison of the size of the band at approximately 60 kDa (**Figure 29**) . Using the same method, almost no PyrH was observed in the cleared cell lysate of BL21 without the use of chaperones.

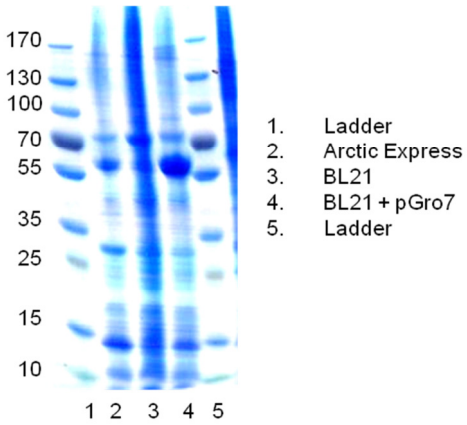
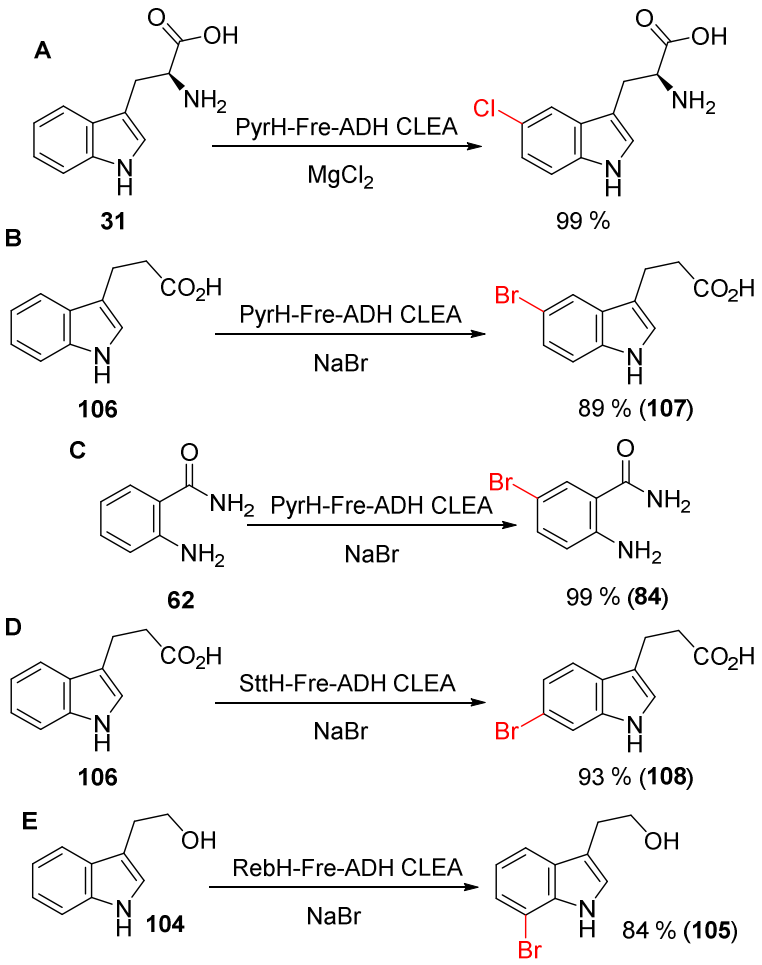


Figure 29: SDS-PAGE of cleared cell lysates of PyrH (mwt 58.1 kDa) expressed in *E. coli* with various chaperones. The ladder used is PageRuler^{TM} Prestained Protein Ladder from ThermoScientific^{TM}.

Following the observation that a larger amount of soluble PyrH was present in cell lysates resulting from the co-expression of PyrH with the pGro7 chaperones, this expression system was subsequently used to trial further CLEA formation conditions. An active PyrH-Fre-ADH CLEA was successfully prepared, capable of the quantitative conversion of tryptophan to 5-chloro tryptophan at a substrate concentration of 3.0 mM (**Scheme 39A**) – an activity improvement in agreement with Sewald's report for RebH.[197] The PyrH CLEA also catalysed the bromination of anthranilamide (**62**) and indole-3-propionate (**106**) with high conversion (**Scheme 39B** and **C**). The RebH CLEA prepared using this method was also found to halogenate the non-natural substrate tryptophol (**104**) with similar efficiency (**Scheme 39E**). With these promising results, CLEAs of SttH were prepared using the same method and found to catalyse the 6-bromination of indole-3-propionate (**Scheme 39D**). In each case ^{1}H NMR confirmed that the regioselectivity of each of these biotransformations was not altered by the use of a CLEA instead of purified enzyme.



Scheme 39: Conversions of the biocatalytic chlorination and bromination reactions found to be efficiently catalysed by CLEAs of various FlHals with cofactor recycling using Fre and ADH. Conversions are based on analytical HPLC. Conditions: substrate (3.0 mM), NADH (100 µM), FAD (10 µM), propan-2-ol (5 % v/v), $MgCl_2$ (30 mM) or NaBr (30 mM) and KPi (pH 7.2) to a total volume of 30 mL, using FlHal-Fre-ADH CLEAs prepared from a total FlHal culture volume of 1.5 L.

One advantage of the CLEA preparations reported by Sewald, amongst others, is that their increased stability allows them to be recycled through multiple rounds of biotransformation without a significant loss in activity. The RebH and SttH CLEAs prepared using this method were similarly found to retain most of their halogenation activity over three 24 hour halogenation cycles of tryptophol (**104**) and indole-3-propionic acid respectively (**106**, **Figure 30**). In these examples, the heterogeneous biocatalyst is easily removed from the reaction after a halogenation cycle by filtration or centrifugation. After washing with buffer to remove residual starting material and cofactors, the CLEA can then be resuspended in fresh reaction buffer containing substrate and cofactors.

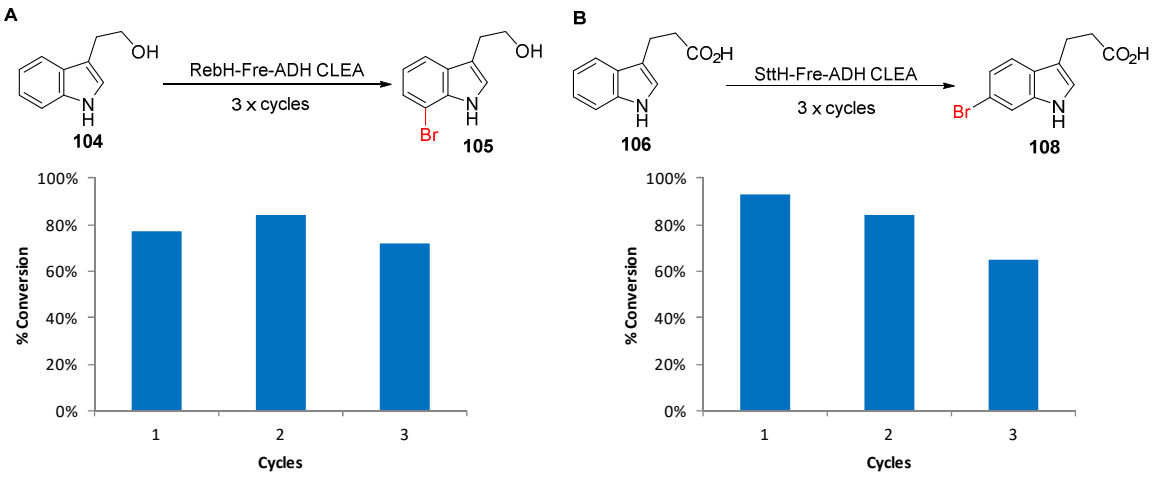


Figure 30: Conversion of bromination reactions catalysed by recycled RebH (**A**) and SttH (**B**) CLEAs over three biotransformations. Conversion determined by analytical HPLC. Conditions: substrate (3.0 mM), NADH (100 μM), FAD (10 μM), NaBr (30 mM), KPi (pH 7.2), propan-2-ol (5 % v/v), rt, overnight.

Since the MWCO filtration approach to compartmentalisation also leads to the collection of biocatalysts after the biotransformation is complete, recycling of the enzymes collected in this manner was also attempted for comparison. A timecourse of anthranilamide (**62**) bromination by PyrH suggested that after 3 hours, approximately 75 % conversion to brominated product was achieved (**Figure 31A**). In order to minimise deactivation of the biocatalyst, and hence maximise the efficiency of its recycling, this timepoint was chosen to collect the biocatalyst by MWCO filtration. Upon concentration of the reaction approximately 20-fold through the MWCO membrane, the harvested biocatalysts were resuspended in fresh reaction buffer containing substrate and cofactor. A small loss of activity was observed in the second cycle of halogenation, whilst the third cycle preceded to only ca. 5 % over the same time. Allowing the third cycle to proceed for a further 9 hours eventually allowed a 49 % conversion to be obtained (**Figure 31B**).

Although this method did allow the halogenation of approximately twice as much material as the same batch of enzyme would have done in a single cycle, the loss of activity between each cycle is significantly more pronounced than when using a CLEA. This rapid loss of activity, combined with the lower substrate concentration employed with recycling of pure protein, meant that an average aryl bromide concentration of 1.2 mM could be afforded over three cycles with the same batch of purified enzyme. When using the CLEA approach however, an average aryl bromide concentration of 2.3 mM was obtained over three cycles with the same CLEA. Despite the advantages of the CLEA, it is worth noting that they often lead to slower biotransformations because of the need for mass transport of substrate into the heterogeneous catalyst.

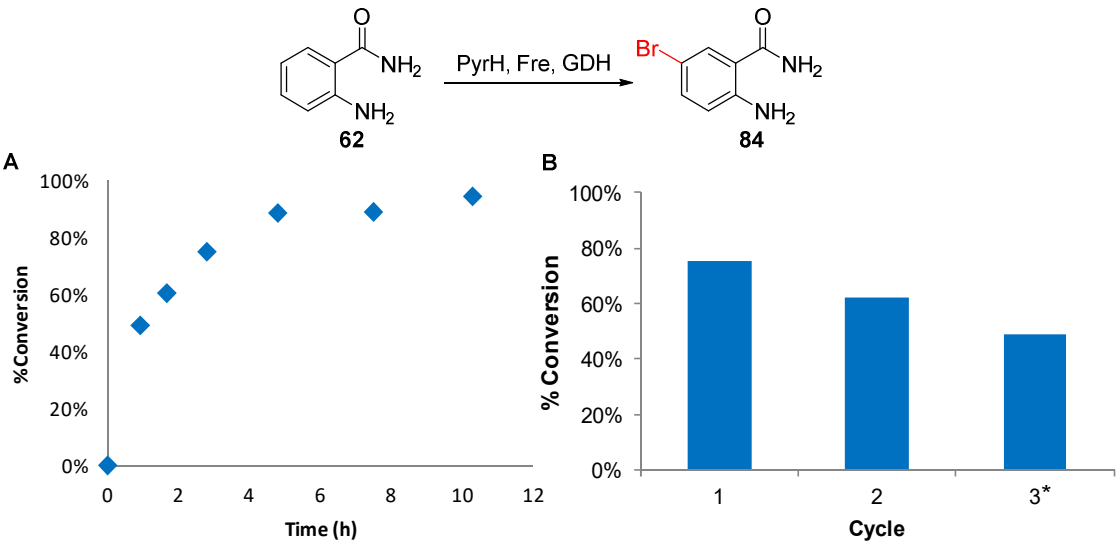


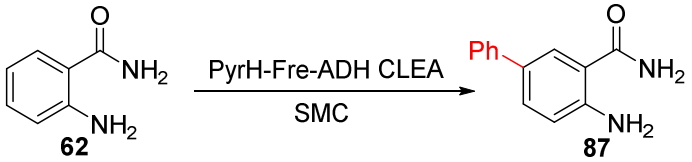
Figure 31: (**A**) Timecourse of the PyrH-catalysed bromination of anthranilamide. (**B**) Conversion of anthranilamide bromination using pure PyrH, Fre and GDH recycled over 3 3-hour biotransformations by MWCO filtration. Conversions are based on analytical HPLC using a calibration curve of **84**. PyrH biotransformation conditions as prior. *Allowed to proceed for an additional 9 hours.

3.2.3.2 CLEA Integration with Suzuki-Miyaura Couplings

Since the use of FlHal CLEAs had afforded a more robust method for the generation of aryl bromides at higher concentration than could be achieved using pure protein, the integration of this process with the SMC was investigated. It was thought that the increased concentration of aryl bromide relative to the cofactors inhibitory of the SMC may allow lower loadings of Pd catalyst to be employed. As formation of a CLEA leads to the capping of nucleophilic surface lysine residues on the biocatalysts, it was also considered that its removal prior to the SMC reaction may not be required. To assess this possibility, a number of analytical scale FlHal-SMC reactions were run using anthranilamide (**62**) and PyrH CLEAs under a number of different modes:

1. Both the FlHal biotransformation and SMC components were added at the start, with heating of the reaction after incubation at room temperature overnight.
2. After incubation of the FlHal biotransformation at room temperature overnight, the reaction was degassed prior to addition of the SMC components and heating overnight.
3. After incubation of the FlHal biotransformation at room temperature overnight, the CLEA was removed by centrifugation prior to degassing of the supernatant and heating overnight with SMC the components.

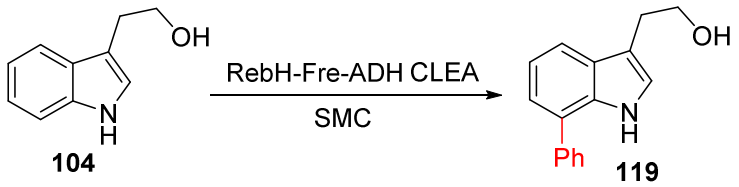
Under mode 1, the **L1**$_2$.Pd(OAc)$_2$ complex afforded trace conversion to arylated product (**Figure 32**, entries 1 and 2), whilst **87** could not be detected in reactions using the Na_2PdCl_4 and tppts catalyst system (**Figure 32**, entries 3 and 4). This is likely due to the air sensitivity of the phosphine ligand – indeed Humera Sharif had found reactions using this ligand under an atmosphere of air to afford no conversion to product.[193] In all cases using this mode (**Figure 37**, entries 1- 4) poor conversion to brominated intermediate was observed also, suggesting that the CLEA is not entirely efficient at preventing mutually-inhibitory protein-palladium interactions.



Entry	Catalyst System	Pd Loading (mol %)	Reaction Mode	Conversion (87, %)*
1	A	10	1	6
2	A	50	1	2
3	B	10	1	N.C.
4	B	50	1	N.C.
5	A	10	2	N.C.
6	A	50	2	24
7	B	10	2	12
8	B	50	2	48
9	A	10	3	51
10	A	50	3	40
11	B	10	3	97
12	B	50	3	94

Figure 32: Effect of the order of SMC components addition and presence of the FlHal CLEA upon conversion of anthranilamide to **87** using a number of reaction modes (see text). PyrH CLEA biotransformation conditions: substrate (3.0 mM), NADH (100 μM), FAD (10 μM), NaBr (30 mM), KPi (pH 7.2), propan-2-ol (5 % v/v), PyrH CLEA, rt, overnight. Catalyst system A: **L1**$_2$.Pd(OAc)$_2$ (x mol %), CsF (30 mM), PhB(OH)$_2$ (15 mM). Catalyst system B: Na_2PdCl_4 (x mol %), tppts (2x mol %), K_3PO_4 (2.4 mM), PhB(OH)$_2$ (15 mM). Both heated at 80 ^{o}C overnight. *Based on analytical HPLC using a calibration curve of **87**. Average of two runs.

Under mode 2, whereby the SMC catalysts were added after degassing of the completed biotransformation, higher conversions were observed in both catalyst systems at both 10 and 50 mol % palladium catalyst (**Figure 32**, entries 5-8). This confirms the idea that the presence of air during mode A could be a limiting factor. In all cases, conversions were highest when the CLEA was removed prior to addition of the SMC components (mode 3, **Figure 32**, entries 9-12) with a promising conversion of 97 % obtained when using 10 mol % of Na_2PdCl_4 in combination with tppts (**Figure 32**, entry 11) – a 5-fold reduction in the loading required for effective FlHal-SMC reactions using MWCO filtration. A similar set of screening reactions were also run using the arylation of tryptophol (**104**) by RebH to determine the reproducibility of these conditions (**Figure 33**), and were found to give conversions comparable to those in **Figure 32**.



Entry	Catalyst System	Pd Loading (mol %)	Reaction Mode	Conversion (119, %)*
1	A	10	2	N.C.
2	A	50	2	26
3	B	10	2	27
4	B	50	2	49
5	A	10	3	68
6	A	50	3	42
7	B	10	3	73
8	B	50	3	90

Figure 33: Effect of the order of SMC components addition and presence of the FlHal CLEA upon conversion of tryptophol to **104** using a number of reaction modes (see text). All conditions as prior (**Figure 32**). *Based upon analytical HPLC using a calibration curve of **119**. Average of two runs.

It is also worth noting that these conversions (**Figures 33** and **34**) were afforded with only 5 equivalents of phenyl boronic acid employed, relative to the 30 required during the arylation using MWCO filtration. This is possibly because the glucose cosubstrate for NADH recycling by GDH employed in the approach using pure protein was replaced by propan-2-ol in the method using CLEAs. Since the addition of glucose to analytical scale SMCs did not have a particularly pronounced effect upon conversion to arylated product in earlier experiments (**Table 5**), it is also possible that the increased efficiency of the integrated FlHal-SMC reaction using CLEAs is due to increased concentration of aryl bromide relative to the other cofactors required for FlHal bromination. The exact reason for this increase in efficiency is currently under investigation in our laboratory.

The observation that removing CLEA prior to addition of the SMC components affords highest conversions to arylated product suggests that the biocatalyst could be coordinating to Pd catalyst thereby rending it inactive. Since this effect is significantly less pronounced than when pure protein was present during SMC reactions (**Table 5**), it is possible that the capping of nucleophilic surface residues prevents this interaction somewhat. The residual reduction in SMC conversion when CLEAs are present could be because of the heterogeneous nature of the CLEA, which may reduce the efficiency of Pd catalyst mass transport throughout the reaction solution. Despite the need to remove the biocatalyst prior to addition of the SMC components in this instance, the heterogeneous nature of the CLEA makes this more facile than when pure protein was employed. Additionally, the ability to recycle the FlHal CLEAs through multiple cycles of halogenation means that their removal could ultimately be advantageous to the process overall.

3.2.3.3 Substrate Scope

With a seemingly more efficient halogenation-SMC cascade in hand, the scope of this process with respect to boronic acid coupling partners was explored. Given that this process required lower loadings of Pd catalyst and fewer equivalents of boronic acid than with approach using MWCO filtration, it was hoped that scope in this regard would be significantly broader when using CLEAs. Accordingly a test arylation of anthranilamide (**62**) was run using PyrH CLEAs, which were removed prior to addition of a number of organoboron coupling partners and Na_2PdCl_4 along with tppts. These analytical scale FIHal-SMC reactions revealed good conversion to arylated adducts by HPLC (**Figure 34**). Subsequent LC-MS analysis of these reactions found *m/z* peaks consistent with what was expected based on the structure of these adducts (**87** to **89** and **120** to **132**, **Appendix One**). Importantly, in addition to the electron-deficient boronic acids accepted by the MWCO filtration process (**Figure 24**), good conversions were observed with organoboron donors bearing electron-donating substituents in addition to heterocyclic and vinyl examples.

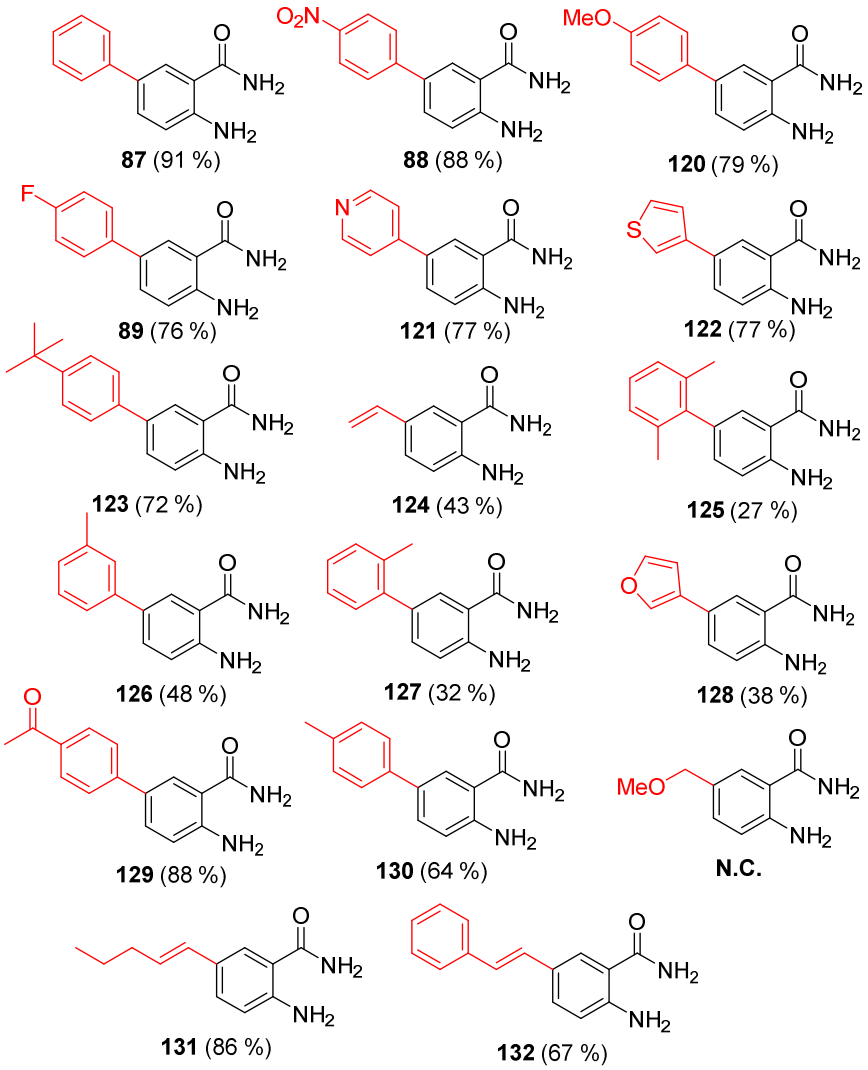


Figure 34: Adducts detected during the analytical scale screening of the integrated FIHal-SMC reaction using the arylation of anthranilamide (**62**) with PyrH CLEAs (**Figure 32**) as the test reaction using 10 mol % of Pd catalyst and 5 equivalents of organoboron donor. Conversions are based upon the consumption of anthranilamide (**62**) and 5-bromo anthranilamide (**84**) by analytical HPLC using calibration curves of these compounds.

The best coupling partners were then selected from this panel and used to perform the preparative FlHal-SMC cascade with both anthranilamide (**62**) and tryptophol (**104**). This allowed a larger and more diverse range of products to be prepared (**Figure 35**), in improved yields to the process involving MWCO filtration. Most importantly, the loading required in this process was significantly lower than previously employed indicating a generally more efficient system. Following initial demonstration of this transformation, the scope of this process was also explored by others in our group. These examples are included below for completeness, with the contribution of these members noted.

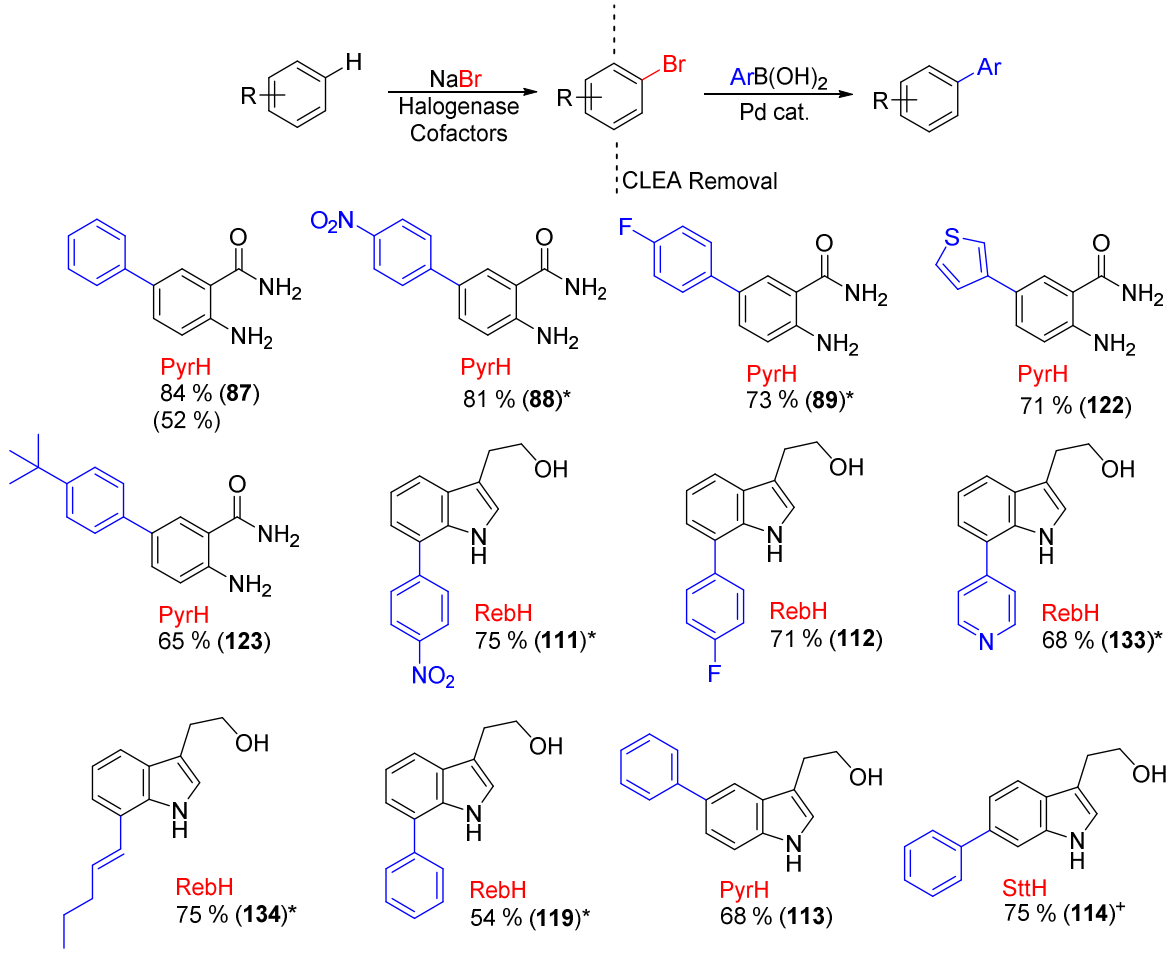


Figure 35: Isolated yields of arylated, vinylated and heteroarylated products obtained by using FlHal CLEAs and the optimised SMC reaction of these supernatants. The FlHal employed is highlighted in Red. Conditions are as described above, involving the removal of CLEA prior to addition of 10 mol % $Na_2PdCl_2.tppts_2$. *Performed by Jean-Marc Henry. +Isolated from three cycles of halogenation using the same SttH CLEA. The yield is brackets was obtained by performing the FlHal-SMC reaction on a 1.0 mmol scale.

By using CLEAs of the tryptophan 5- and 6-halogenases PyrH and SttH respectively, arylation of the 5- and 6-positions of the indole heterocycle was also accomplished – on an increased scale, using a lower Pd catalyst loading, and with higher yield than the MWCO approach. In the case of the 6-arylation of indole-3-propionic acid, the recyclability of CLEAs was exploited. In this example the combined supernatants from three SttH-catalysed brominations of **106** using the same CLEA were combined and arylated in a single reaction. This afforded more than 6 times the amount of

arylated material than could be obtained from a single arylation transformation using pure protein and MWCO filtration. The increased catalyst lifetime of CLEAs was also exploited for the arylation of anthranilamide. Here, CLEAs prepared from a total of 4.5 L of culture expressing PyrH were used for the continued halogenation of **62** over 6 days. After the biotransformation had reached 60 % conversion however, no further conversion to brominated product was detected. Arylation of this supernatant afforded a 52 % isolated yield of 5-phenyl anthranilamide (equivalent to approximately 0.5 mmol on this scale).

The use of CLEAs in the FlHal-SMC cascade significantly improved the efficiency of this coupling with respect to scalability, Pd catalyst loading and the number of equivalents of organoboron reagent which were required compared to the approach using MWCO filtration. The requirement to remove CLEA prior to the SMC to allow good conversion to arylated product to be obtained however, introduces an intermediary processing step between the two reactions which is undesirable from the perspective of space-time yield. This processing is minimal and facile due to the heterogenous nature of the biocatalyst and does not consisitute an isolation or purification process. The increased stability of this preparation also allows the efficient recycling of the biocatalyst between multiple cycles of halogenation, which somewhat compensates for the reduction in space-time yield from its removal prior to the SMC reaction.

3.2.4 Membrane Compartmentalisation for Halogenation Cross-coupling Cascades

In order to reduce the intermediary processing required between the FlHal and SMC reactions, methods for compartmentalising the bio- and chemo-catalysts within the same reaction vessel were explored. A biphasic reaction system was briefly investigated, which would allow the compartmentalisation of palladium catalysts bearing hydrophobic ligands in the organic phase, whilst the FlHal operated in a distinct aqueous phase. The FlHals were found to be stable with respect to a limited number of water-immiscible organic solvents, although the polar nature of the substrates employed meant that insufficient partition of aryl bromide into the organic phase was observed.

An alternative approach to biphasic reactions has been reported for the separation of incompatible reagents by the groups of Bowden and Gröger.[37-39] As described in Chapter 1 (**Scheme 7**), these approaches involve the use of partially-permeable membranes. Unlike the MWCO membranes utilised for earlier FlHal-SMC cascades, these membranes separate compounds based on their charge. The membranes are made from a hydrophobic polymer, poly-dimethylsiolxane (PDMS, **135**), in which only relatively non-polar components are soluble - meaning that highly polar or charged reagents are insoluble and cannot pass between the bulk solvent on either side of the membrane.[36,200] Since the biocatalysts used in this case are highly charged, it was envisioned that they are unlikely to dissolve sufficiently well in PDMS to pass through membranes made of PDMS. As the palladium catalysts employed also possess a net charge, it was expected that they also should be contained on a single side of the membrane – therefore allowing spatial separation of the bio- and chemo-catalysts without the need for intervention or removal (**Figure 36A**). The charged nature of the cofactors NADH and FAD, which also to have a deleterious effect on SMC efficiency,

were expected to be contained too - which may allow the use of further reduced loadings of Pd catalyst. By employing inorganic bases for the SMC reaction, it was expected that a basic pH could be achieved on the SMC side of the membrane whilst a pH closer to the optimum for FlHal activity could be maintained on the other.

3.2.4.1 Poly-dimethylsiloxane Thimbles

PDMS membranes are often utilised in the form of "thimbles" for the compartmentalisation of incompatible reagents (**Figure 36**).[36-38] Thimbles are small cylinders made of PDMS into which one set of reagents is added (**Figure 36B**). This small vessel can then be placed inside a larger reaction vessel containing the second set of reagents. Such thimbles are commonly manufactured on a laboratory scale using a glass sample vial as template, which is coated in the monomers **136** and **137** before polymerisation to afford a PDMS thimble of approximately the same size.[36]

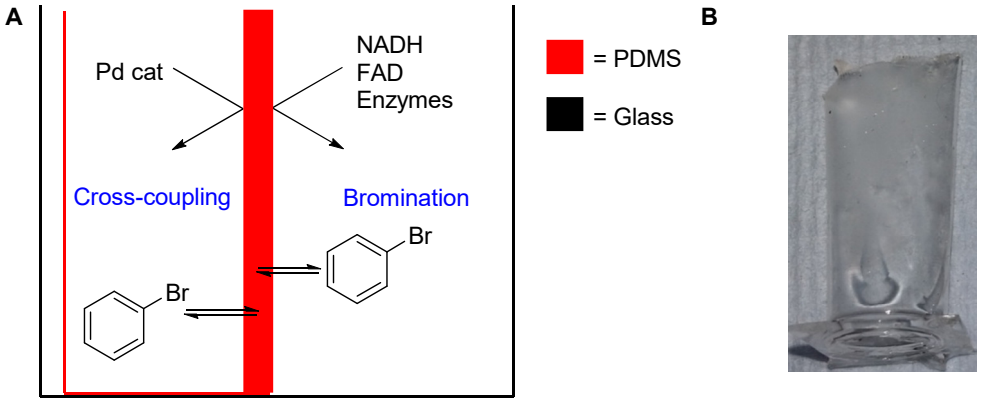
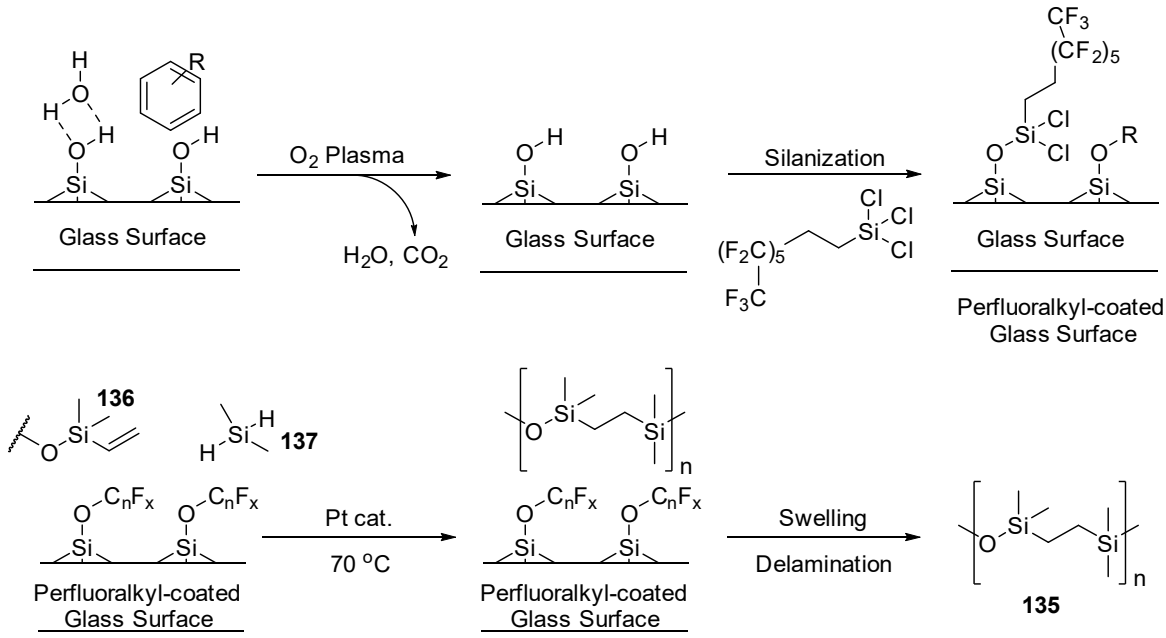


Figure 36: (**A**) Proposed use of a PDMS membrane for compartmentalisation of the FlHal and SMC reactions. One wall of the thimble is illustrated as being thicker to convey the idea that passage through the membrane requires partition of compound into PDMS. (**B**) Complete PDMS thimble fabricated using a glass vial with dimensions of approximately 18 mm x 50 mm as template.

In order to fabricate thimbles for use in the FlHal-SMC cascade, a modification of the process reported by Bowden was used (**Scheme 40**).[38] Firstly the glass vials are placed in an oxygen plasma cleaner for 15 minutes. This is used to ensure that all silanol functionalities on the glass surface are evenly exposed through the removal of foreign organic matter and water layers. The cleaned glass substrate is then treated with a perfluorinated chlorosilane by vapour deposition *in vacuo*. Reaction of the exposed silanol moieties on the glass substrate with the chlorosilane (silanization) then results in a template coated with fluoroalkyl chains, which aids in the removal of the PDMS layer after coating. The treated vial is then coated in a mixture containing the monomers **136** and **137** as well as a Pt catalyst (obtained as a Sylgard® 184 kit) and heated to aid polymerisation. The coated vial is then soaked in hexane, which leads to swelling of the PDMS layer because of the high solubility of hexanes in PDMS,[200] and a complete cylinder of PDMS is removed (delamination) – i.e. a thimble. The thimbles are then cleaned with solvents, to remove un-reacted monomers and Pt catalyst, then dried before use.



Scheme 40: Schematic representation of the process for fabrication of PDMS thimbles.[201] The nature of the platinum catalyst added to Sylgard® 184 is not specified by Dow.

In order to use these membranes for the integration of the FIHal and SMC reactions, it is a pre-requisite that intermediate aryl bromides dissolve sufficiently well in PDMS to flux from one side of the membrane to the other. To determine the feasibility of using this approach in FIHal-SMC cascades, the flux of 5-bromo anthranilamide (**84**) and 7-bromo tryptophol (**105**) from the outside to the inside of a PDMS thimble was determined. A PDMS thimble was placed into a solution containing each substrate and incubated at room temperature (**Figure 38A**). At specified timepoints, the UV spectrum of both sides of the membrane (i.e. inside and outside of the thimble) was measured and used to calculate the concentration of substrate on either side. Plotting the concentration of both aryl bromides over time found that, although flux is slow, the passage of substrate between the compartments is possible (**Figure 38A** and **38B**). Indole-3-propionic acid (**106**) was found move even more slowly though the membrane, reaching only 7 % flux after incubation overnight. The same experiment using the cofactors NADH and FAD in addition to the **L1**.$Pd(OAc)_2$ palladium catalyst found that even after overnight incubation, flux between the two compartments was negligible (**Figures 38C** to **38E**). This effect can even be seen with the naked eye using the highly coloured **L1**$_2$.$Pd(OAc)_2$ complex (**Figure 37**).

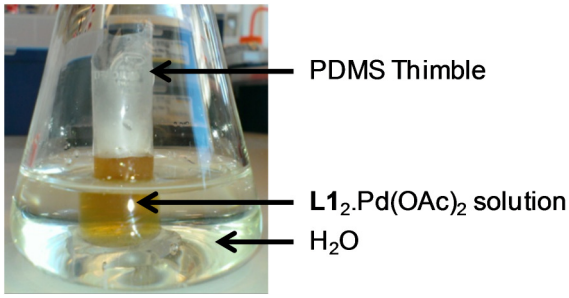


Figure 37: Photograph of **L1**$_2$.$Pd(OAc)_2$ contained within a PDMS thimble added to bulk H_2O.

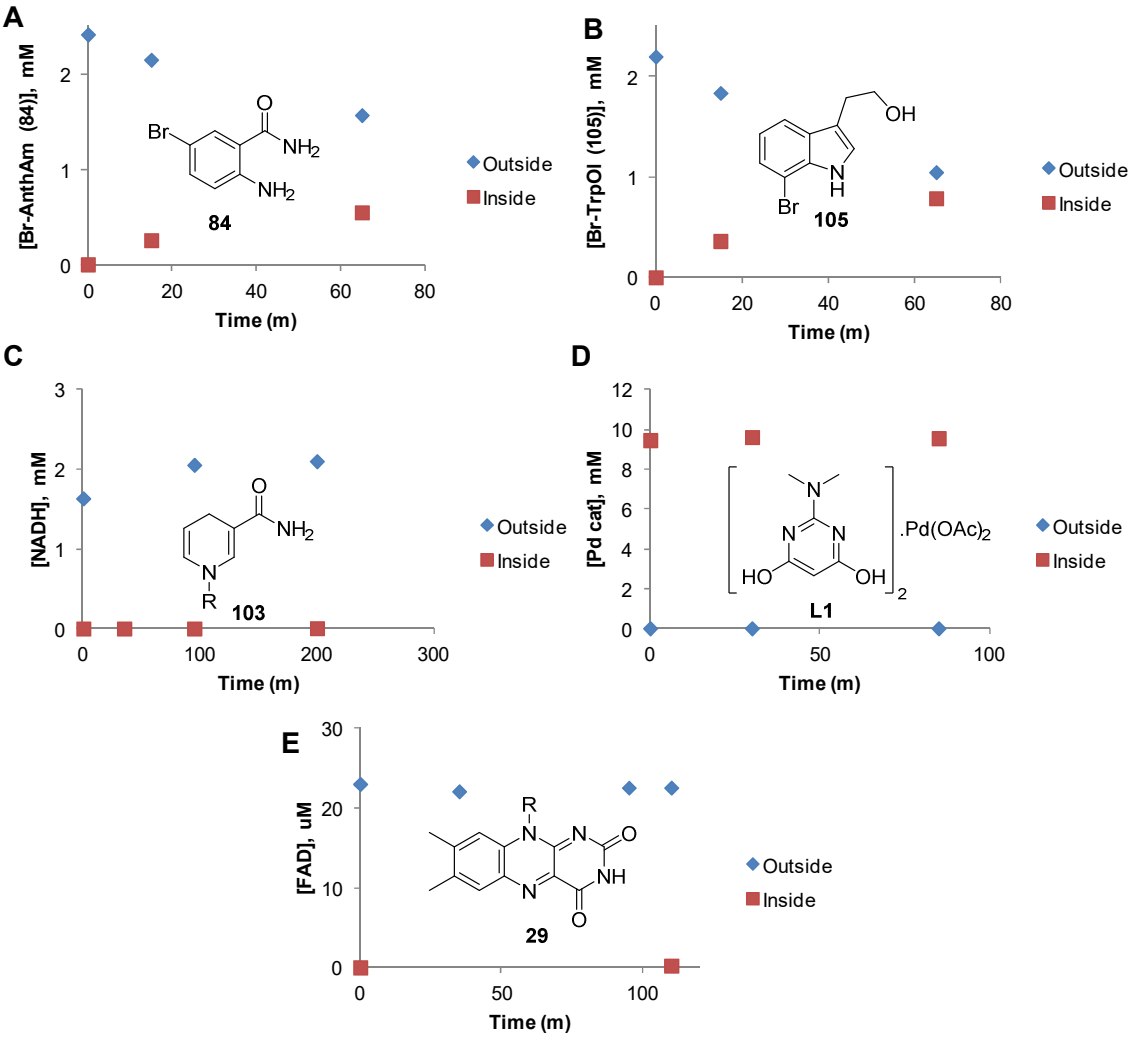


Figure 38: Flux of (**A**) 5-bromo anthranilamide (**84**), (**B**) 7-bromo tryptophol (**105**), (**C**) NADH (**103**), (**D**) **L1**$_2$.Pd(OAc)$_2$ and (**E**) FAD (**29**) through a PDMS thimble over time. Flux was determined by individually adding (**84**), (**105**), (**103**) or (**29**) to 10 mL of KPi followed by a PDMS thimble containing 1 mL of KPi. Concentration of each component was determined using UV-Visible spectroscopy at the timepoints stated. Flux in (**D**) was determined in the same way with **L1**.Pd(OAc)$_2$ added to the thimble. See section 7.3.9 for full experimental details.

3.2.4.2 Integration using Pure Protein and PDMS Compartmentalisation

With the knowledge that aryl bromides could move freely through a PDMS membrane, whilst **L1**$_2$.Pd(OAc)$_2$ could be effective contained, a test SMC reaction was conducted using a solution of pure 5-bromo anthranilamide (**84**) containing the cofactor concentrations utilised for a PyrH-catalyed bromination reaction. To this solution was added a PDMS thimble containing the previously described SMC components and the reaction heated under air at 80 ^{o}C overnight. Analytical HPLC showed this reaction to proceed to 78 % conversion, demonstrating that this method of compartmentalisation is suitable for use with the FlHal-SMC cascade. It was noted that the SMC-like reactivity observed could be due to residual Pt catalyst from the thimble fabrication process.[202] Control reactions did not show any evidence of arylated product **87** by analytical HPLC however when **L1** was used in the absence of any Pd source.

Based on the previous observation that enzymes are unable to move through PDMS membranes,[39] a test FlHal-SMC cascade was run using antranilamide (**62**) and purified PyrH. Addition of a PDMS thimble containing 10 mol % **L1**$_2$.Pd(OAc)$_2$, 5 equivalents of PhB(OH)$_2$ and 10 equivalents of CsF to the PyrH-catalyed biotransformation allowed isolation of **87** in 74 % yield (**Figure 39**). The earlier method for this transformation using purified PyrH and MWCO filtration necessitated a Pd loading of 50 mol % and a vast excess of PhB(OH)$_2$ to achieve similar yields. This reaction therefore demonstrates how the PDMS membrane can be utilised to overcome inhibitory interactions which reduce SMC efficiency. Given the compartmentalisation of the bio- and chemo-catalysts, the only intervention between the FlHal and SMC reactions was heating – no removal of reagents or degassing was required. To test the efficiency of this coupling, the particularly challenging 3-furanyl boronic acid coupling partner was selected. Attempts to use this coupling partner in the arylation of anthranilamide (**62**) using pure PyrH afforded disappointing isolated yields of the arylated product **128**, however (**Figure 39**).

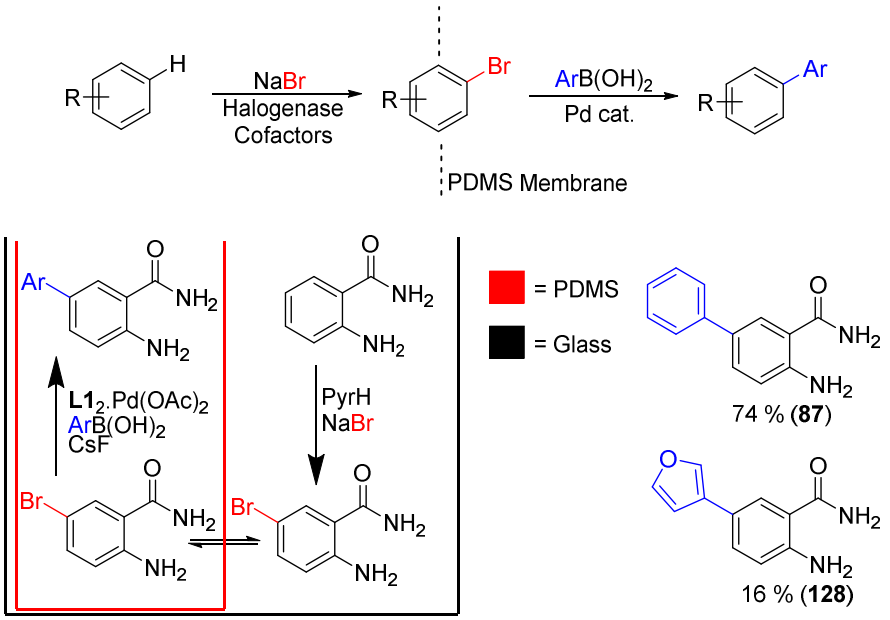


Figure 39: Isolated yields obtained from the integrated arylation of antranilamide (**62**) using purified PyrH and PDMS compartmentalisation. PyrH biotransformation conditions are as prior with 2.0 mM anthranilamide. SMC conditions: ArB(OH)$_2$ (5 eq), CsF (10 eq), **L1**$_2$.Pd(OAc)$_2$ (10 mol %), 80 °C, overnight.

It was envisioned that the compartmentalisation of the FlHal and SMC in this manner could allow the concurrent action of both the FlHal and the SMC – with both reactions forming product at the same time. This was not possible using any of the FlHal biotransformations used thus far because the SMC of these substrates required heating to 80 °C, presumably to aid oxidative addition of Pd^0 across the C-Br bond, at which temperature none of the FlHals were active. After finding Th_Hal to be efficient at catalysing chlorination reactions at increased temperature however,[175] it was hoped that this thermophilic halogenase variant may be used for this purpose. Use of Th_Hal in a PDMS compartmentalised arylation of anthranilamide (**62**) at 45 °C unfortunately did not afford any 5-bromo anthranilamide (**84**) intermediate or arylated product **87**, however. Although Th_Hal could

produce 5-chlorinated anthranilamide (**62**) and anthranilic acid (**61**) at 30 °C (**Figure 11**), conversion in both cases was low – suggesting that the binding of these small substrates to Th_Hal is poor. The effect of poor binding upon conversion in the halogenation reaction will be more pronounced at higher temperatures, since the entropic penalty for binding of substrate to active site is higher.

3.2.4.3 Integration using Immobilised Biocatalyst and PDMS Compartmentalisation

With the strategies involving the use of FlHal CLEAs and compartmentalisation of the SMC using PDMS both providing improvements to the FlHal-SMC cascade in their own right, it was hoped that the combination of both methods could allow the efficiency of the process to be increased further – both in terms of Pd catalyst loading and space-time yield. PyrH CLEAs from 3 L of culture were therefore used to generate a higher concentration of aryl bromide than achieved in reactions using pure protein. Addition of a PDMS thimble containing 10 mol % **L1**$_2$.Pd(OAc)$_2$, 3-furanyl boronic and CsF to the supernatant of this biotransformation prior to heating at 80 °C allowed the isolation of **128** in 64 % yield (**Figure 40**) – a significant increase in yield based upon the method using pure protein (**Figure 39**). Employing the same conditions using a RebH CLEA in place of PyrH allowed the 7-functionalisation of tryptophol (**104**) with both 1-pen-1-enyl boronic acid and 4-tertiarybutylphenyl boronic acid to afford **134** and **138** in similar yields (**Figure 40**). The increased aryl bromide concentration generated by using CLEAs allowed the Pd loading employed for synthesis of **87** to be reduced to 2 mol % from 10 mol % (**Figure 40**) - a significant improvement when compared to the 50 mol % loading of Pd required in the initial MWCO filtration approach (**Figure 27**).

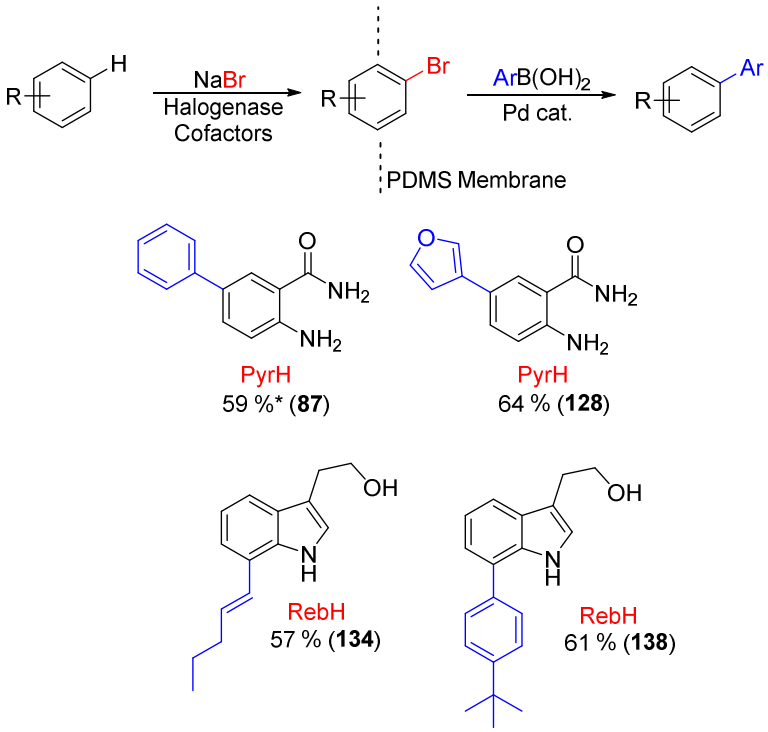


Figure 40: Isolated yields from the integrated arylation and vinyl-ation of anthranilamide (**62**) and tryptophol (**105**) using FlHal CLEAs and PDMS compartmentalisation. Obtained using the SMC conditions as prior (**Figure 39**), except a substrate concentration of 5.0 mM was used throughout and *used 2 mol % Pd rather than 10 mol %.

During the preparation of **134** and **138**, the RebH CLEA was left in the vessel whilst heating took place. After re-suspension into fresh reaction buffer once the SMC was complete however, these CLEAs were found to be completely inactive. Therefore although removing the CLEA prior to the SMC reaction does introduce an intermediary processing step, the inability to recycle this biocatalyst reduces the usefulness of doing so. Additionally the facile nature of the removal process limits its impact on the efficiency of the process as a whole.

In another attempt to develop a FlHal-SMC cascade which occurs in a truly concurrent manner, the ability of CLEAs to catalyse bromination reactions at high temperatures was also explored. Conversion of anthranilamide (**62**) to 5-bromo anthranilamide (**84**) by a PyrH CLEA at 45 °C however was very poor (**Figure 41**). Similar to the case involving Th_Hal, this could be due to an entropic effect. It is also possible however that the elevated temperature leads to degradation of either the cofactors required for FlHal activity or the imine linkages essential for stability of the CLEA.

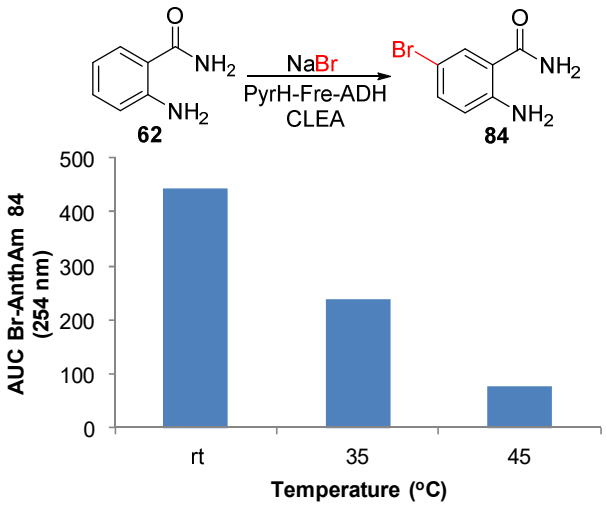
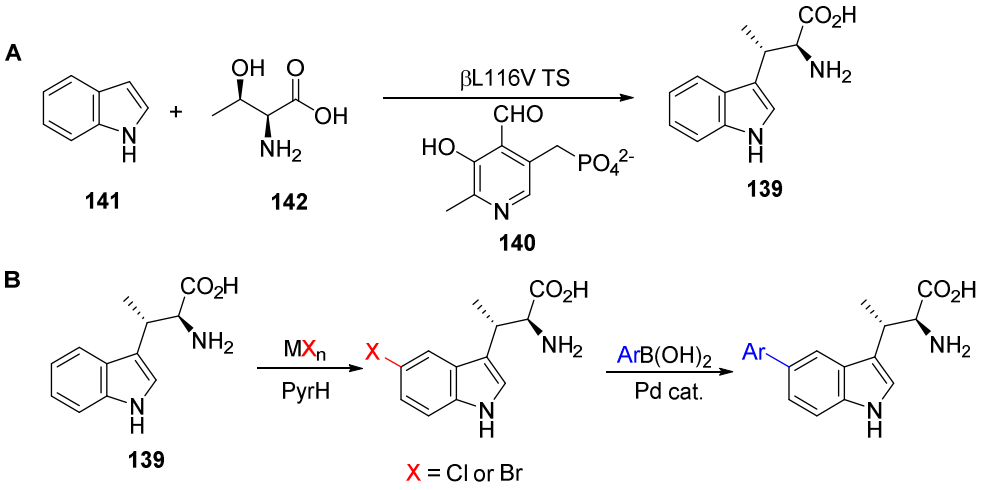


Figure 41: Effect of increasing temperature upon the conversion of anthranilamide (**62**) to 5-bromo anthranilamide (**84**) by PyrH CLEAs. Biotransformations were conducted as prior at the specified temperature overnight. AUC was determined by analytical HPLC. rt = room temperature.

3.2.5 Application to the Synthesis of Tryptophan Derivatives

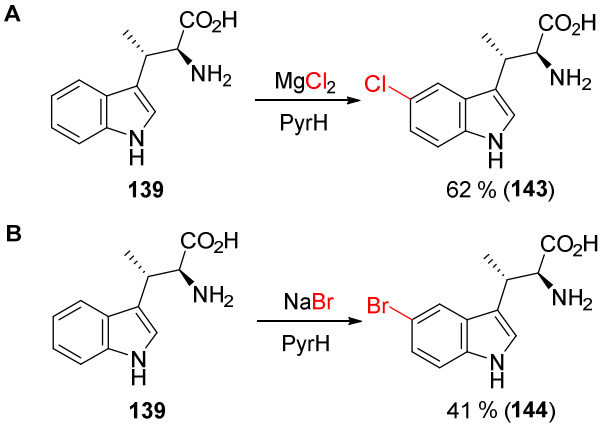
During the course of this study, others in our group had developed a biocatalytic method for the production of non-natural β-methylated tryptophan derivatives (**139**) using an engineered mutant of a tryptophan synthase enzyme.[203] In nature these enzymes are responsible for the synthesis of tryptophan (**31**) through the condensation of indole 3-glycerolphosphate with L-serine through a PLP (**140**)-mediated mechanism.[204,205] This well-characterised enzyme has been utilised to prepare many tryptophan derivatives using functionalised indoles and L-serine.[206] Daniel Francis found that mutation of a single active site residue of a tryptophan synthase enzyme allowed the condensation of indole (**141**) with L-threonine (**142**) instead however – thereby generating (2*S*,3*S*)- β-methyl tryptophan (**139**) stereoselectivley (**Scheme 41A**).[203] Whilst this engineered mutant could accept many indoles to generate the corresponding β-methyl tryptophan derivatives, yields with 5-fuctionalised indoles were poor. Given that the FlHal-SMC cascade developed above allowed the

direct 5-functionalisation of indoles, it was considered that this approach could be utilised to expand the scope of products prepared by this process beyond what is possible using the tryptophan synthase mutant alone (**Scheme 41B**).



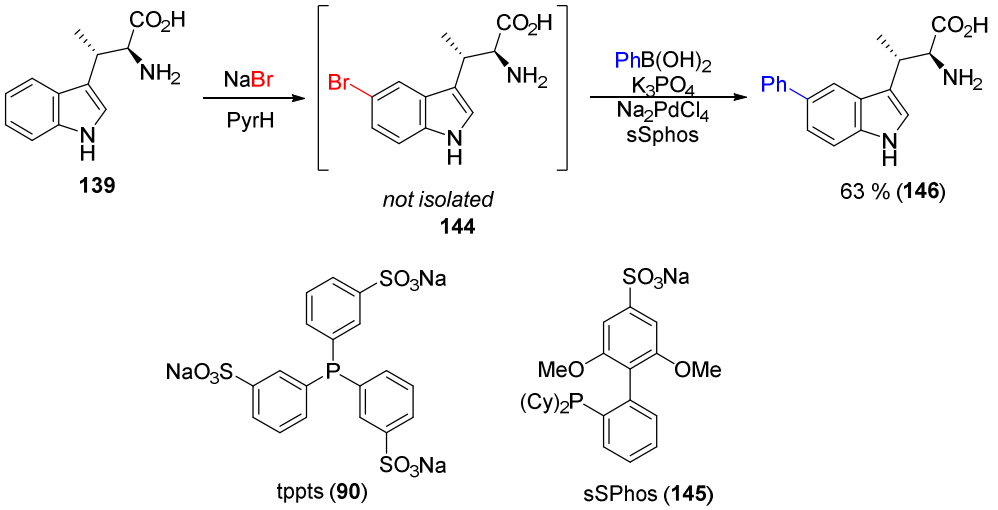
Scheme 41: (**A**) Preparation of (2*S*,3*S*)-β-methyl tryptophan (**139**) from indole (**141**) and L-threonine (**142**) using an engineered tryptophan synthase (TS) enzyme.[203] Conditions are described in the experimental section. (**B**) Proposed 5-derivatisation of **139** using PyrH in FlHal-SMC cascades.

To assess this possibility, (2*S*,3*S*)-β-methyl tryptophan (**139**) was prepared according to the optimised process (**Scheme 41A**), which allowed isolation of **139** in 51 % yield.[203] Analytical scale reactions of **139** with purified PyrH (a tryptophan 5-halogenase) under standard FlHal conditions showed PyrH to chlorinate **139** with 68 % conversion by analytical HPLC – relative to 98 % chlorination activity with tryptophan. After identification of *m/z* peaks consistent with **143** by LC-HRMS, the biotransformation was upscaled using a PyrH CLEA which allowed isolation of **143** in 62 % yield (**Scheme 42A**). Analysis by ^{1}H NMR confirmed the product to be 5-subsituted. Replacement of $MgCl_2$ with NaBr allowed the analogous bromination reaction to be carried out in the same way, with **144** isolated in 41 % yield (**Scheme 42B**).



Scheme 42: Isolated yields obtained from the (**A**) chlorination and (**B**) bromination of **139** using PyrH CLEAs. Conditions are as previously reported for PyrH CLEA biotransformations except CLEAs from 3 L of culture expressing PyrH were used.

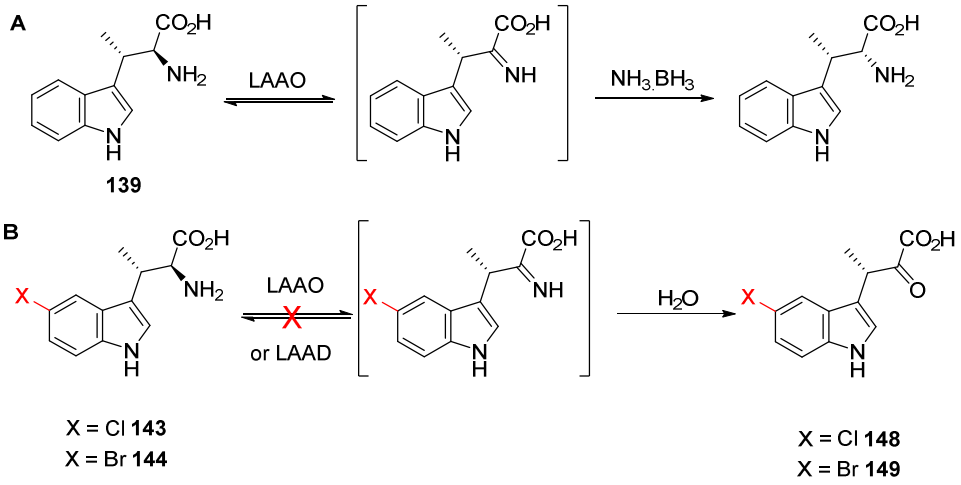
With a method for the regioselective bromination of **139** at the 5-position in hand, the use of this moiety for further functionalisation was explored. Integration with Pd catalysis using PDMS compartmentalisation was considered infeasible, since the existence of the zwitterionic amino acid moiety at the reaction pH would likely prevent efficient flux of intermediate aryl bromide through the membrane. The method involving removal of the CLEA prior to addition of SMC components was therefore explored. Initial attempts employing Na_2PdCl_4 and tppts (**90**) as catalyst afforded only a trace of product which could be identified as an aryl adduct of **139** by LC-HRMS. During attempts at optimising this reaction however, another group reported the efficient SMC of bromo tryptophan derivatives in aqueous media using sSphos (**145**) in place of tppts (**90**).[207] Application of these conditions to the SMC of biotransformation supernatants containing **144** allowed isolation of the arylated product **146** in 63 % yield from **139** (**Scheme 43**). Presumably the increased electron density of the P atom in sSPhos (**145**) compared to tppts (**90**) allows a tighter coordination of sSphos to Pd – thereby preventing inhibition of productive SMC cycles by coordination of the unprotected amine moiety.



Scheme 43: One-pot arylation of (2*S*,3*S*)-β-methyl tryptophan (**139**) using PyrH CLEAs. FlHal biotransformation conditions are as previously described except the bromination proceeded for two days. SMC conditions: $PhB(OH)_2$ (5 eq), K_3PO_4 (25 eq), sSPhos (15 mol %), Na_2PdCl_4 (5 mol %), 80 °C, overnight. Cy = cyclohexyl.

It was hoped that the L-amino acid oxidase (LAAO) catalysed inversion of **139**,[208] demonstrated previously (**Figure 44A**), could be used for the inversion of the halogenated derivatives **143** and **144** – thereby allowing access to additional diastereoisomers. Attempts at these inversions afforded only starting material, however. Further investigation by HPLC and LC-HRMS showed that whilst the keto-acid **147** could be detected upon incubation of LAAO with **139**, the analogous intermediates **148** and **149** could not be detected from the halogenated derivatives – suggesting **143** and **144** are not substrates for the LAAO due to their increased steric bulk. An analogous enzyme, L-amino acid deaminase (LAAD), developed by the Turner group was therefore explored, as this enzyme had been shown to be effective in the inversion of bulky non-natural phenylalanine substrates.[209] Unfortunately analogous reactions using this biocatalyst indicated only a trace of intermediate **148** by HPLC and LC-HRMS, suggesting that **144** is a poor substrate for LAAD. Some

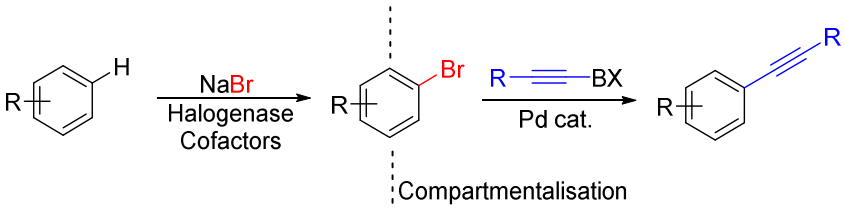
tryptophan halogenases have been demonstrated to accept both the L- and D- isomers of tryptophan however,[124,159] suggesting that access to additional diastereoisomers of 5-halo and 5-aryl **139** derivatives may still be possible through inversion of **139** prior to halogenation.



Scheme 44: (**A**) LAAO-catalysed stereo inversion of (2*S*,3*S*)- β-methyl tryptophan. (**B**) Attempted inversion of the halo (2*S*,3*S*)-β-methyl tryptophan derivatives.

3.2.6 Regioselective Alkynylation

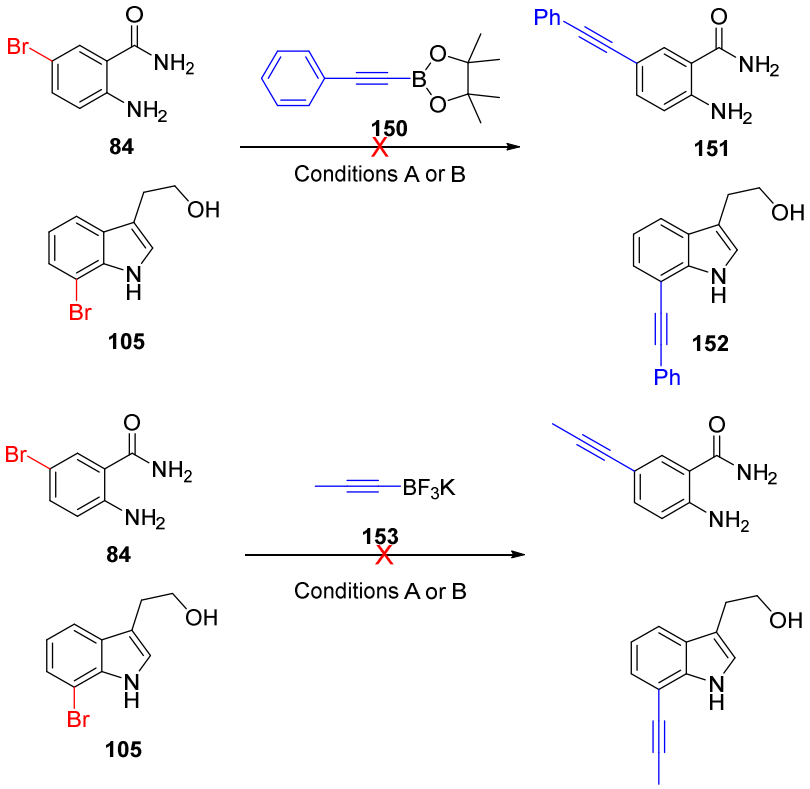
It was hoped to broaden the scope of products obtained from the integration of FlHal enzymes and Pd catalysis beyond the regioselective conjugation of aryl, vinyl and heteroaryl groups. Preliminary studies on the integration of Sonogashira couplings into these regimes by Humera Sharif however found very poor conversions with a number of different catalyst systems and conditions. With the FlHal-SMC cascade in hand however, it was postulated that the use of alkynyl boron species may allow a similarly regioselective alklynation reaction to be developed (**Scheme 45**).



Scheme 45: Proposed scheme of regioselective alkynylation using FlHals and SMCs.

Based on literature precedent for the SMC of this reagent with aryl bromides,[210] the phenylacetylene boronic acid ester **150** was selected for screening the feasibility of this reaction. The SMC of CLEA biotransformation supernatants containing either **84** or **105** using previously optimised conditions (**Scheme 46**) did not afford any conversion apparent by analytical HPLC, however. Further analysis of these reactions by LC-HRMS and crude ^{1}H NMR did not provide any evidence of successful coupling either. To rule out incompatibility of this particular reagent with the FlHal biotransformation components, the same reaction was conducted using purified **84** or **105** and the $Pd^{0}(PPh_3)_4$-catalysed conditions used to prepare standards **87** to **89** (**Scheme 34**). Similar analysis did not provide any evidence of the coupled products **151** or **152**. Considering that the

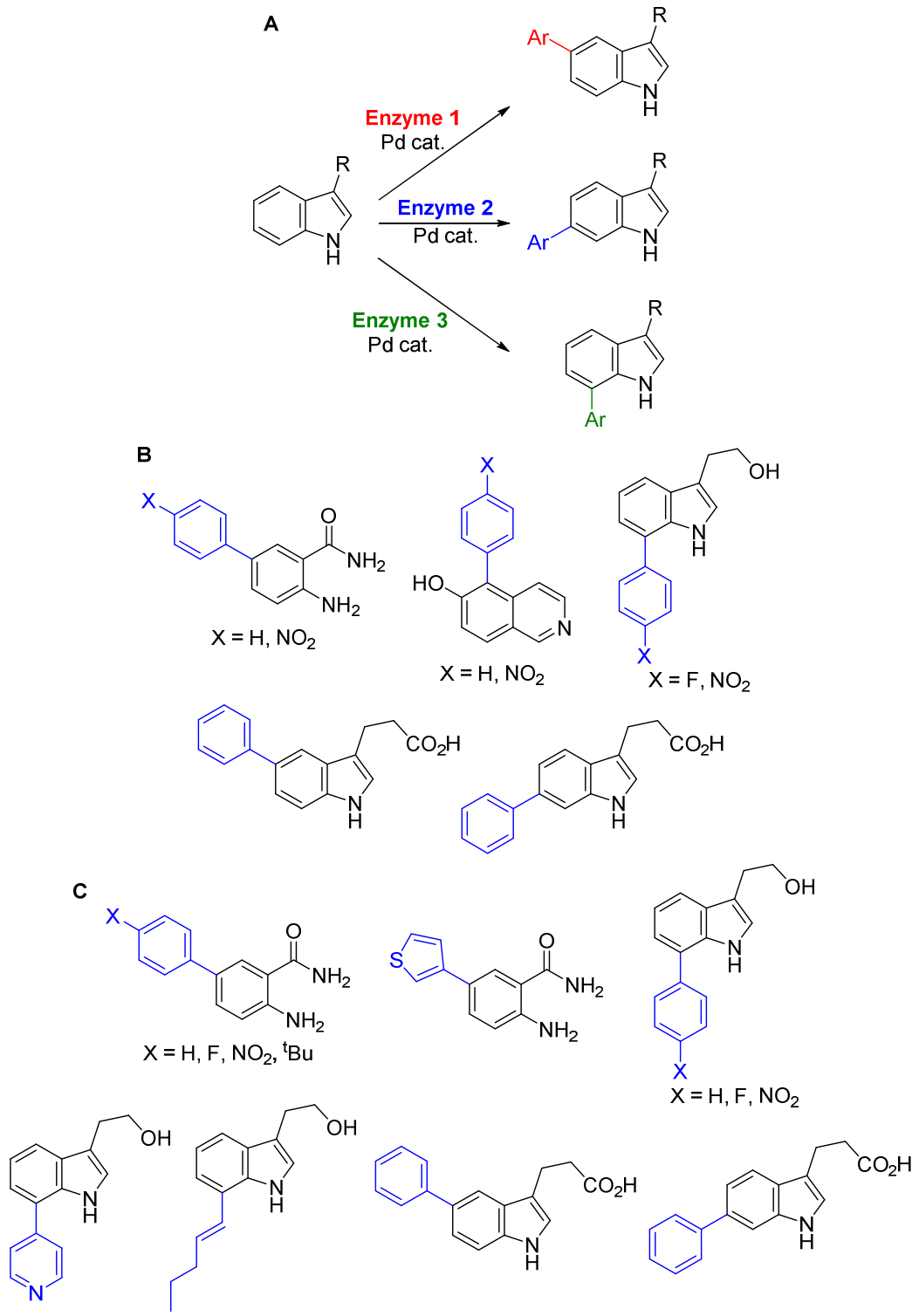
stability of reagent **150** could be a limiting factor, the alkynyl trifluoroborate salt **153** was explored as an alternative. Screening of **84** and **105**, both as purified compounds and present in FlHal CLEA biotransformation supernatants, with this reagent however did not provide any evidence of successful coupling by HPLC, LC-HRMS or ^{1}H NMR.



Scheme 46: Unsuccessful SMC reactions of FlHal bromination products attempted. Conditions A: To a supernatant from CLEA-catalysed bromination of **84** or **105**, at 3.0 mM substrate concentration was added organoboron reagent (5 eq), K_3PO_4 (10 eq), tppts (20 mol %), Na_2PdCl_4 (10 mol %), 80 °C, overnight. Conditions B: Purified bromide **84** or **105** (30 mM), organoboron reagent (5 eq), K_2CO_3 (9 eq), $Pd^0(PPh_3)_4$ (10 %), toluene, H_2O, EtOH, reflux, overnight.

3.3 Conclusions and Outlook

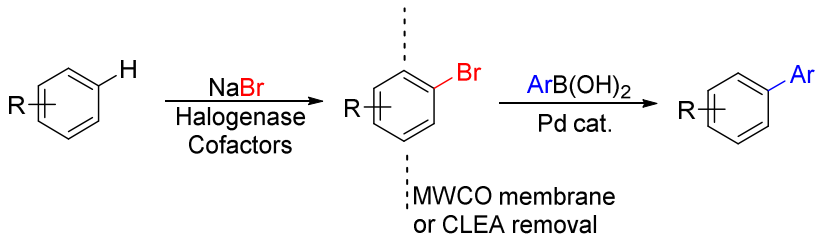
The selective manipulation of similar, unactivated, positions of the indole moiety has been a long-standing challenge in the field of C-H activation chemistry.[67] Early attempts at the direct arylation of indole C-H bonds employed substrate-control strategies – whereby electronic or coordinating effects of substrate to a transition metal are exploited to control regioselectivity. This chapter describes an alternative approach to this problem, whereby the regioselectivity of the flavin-dependent halogenase enzymes is used to control the position of halogenation. By integration with palladium-catalysed cross-coupling chemistry, the C-Br bond was transformed to a C-C bond (**Scheme 47A**).



Scheme 47: (**A**) Strategy employed herein, using FlHals of different regioselectivity to direct the halogenation and subsequent arylation of the C5, C6 and C7 indole positions through combination with palladium catalysis. (**B**) Scope of products obtained from integrated FlHal SMC reactions using pure protein and MWCO filtration. (**C**) Scope of products obtained from integrated FlHal SMC reactions using CLEA-immobilised FlHals.

Initial attempts at this integration revealed a marked deleterious interaction between components of the biocatalytic and chemocatalyic transformations, which ultimately afforded low yields of arylated compound. This was initially overcome by separation of the bio- and chemo-catalysts using a size-selective membrane and increased loading of Pd catalyst (**Figure 27**). Although this approach did allow a modest range of arylated compounds to be obtained (**Figure 47B**), the near-stoichiometric amount of palladium catalyst required (50 mol%) significantly reduced the synthetic utility of this method. In an attempt to reduce the loading of palladium catalyst required, the productivity of the FIHals was improved through the formation of stabilised cross-linked enzyme aggregates (CLEAs). By providing increased catalyst lifetime, these preparations afforded high concentrations of intermediate aryl bromide and subsequently allowed lower loadings of Pd catalyst to be employed (**Figure 35**). This more robust methodology for integrated FIHal SMC reactions allowed a broader range of arylated, heteroarylated and vinylated compounds to be prepared (**Figure 47C**). In addition to allowing efficient SMC reactions with lower loadings of palladium catalyst, increased substrate concentration also improved the space-time yield of the reaction compared to those involving pure protein.

Both of these methods required the removal of the biocatalyst prior to the addition of the SMC reagents (**Scheme 48**), however. Although in the case of CLEAs this was demonstrated to improve the efficiency of the process overall by allowing the recycling of biocatalyst, the requirement for intermediary processing between the two reactions is not ideal.



Scheme 48: Generalised approach the FIHal-SMC cascades through removal of the biocatalyst prior to SMC.

To overcome the need for biocatalyst removal, membranes which separate compounds based on charge were explored (PDMS thimbles). Since both the bio- and chemo-catalysts carry net charges, this approach allowed the two catalysts to be effectively separated from each other without the need for intermediary processing (**Figure 42A**). By separating the cofactors NADH and FAD, in addition to enzymes, from the SMC components, this approach allowed significantly reduced loading of palladium catalyst (10 mol%) to afford good yield of arylated products compared to MWCO filtration methods using purified protein (**Figure 39**). The improved efficiency of the SMC reaction in this case also allowed heteroarylation reactions which were not possible using MWCO filtration (**Figure 42B**). Most notably, the use of PDMS membranes meant that no intermediary processing was required between the FIHal and SMC reactions – affording an increase in space-time yield compared to the previous processes requiring removal of biocatalyst by filtration. To further improve the efficiency of FIHal SMC integrated reactions, biocatalyst bromination reactions

using FlHal CLEAs were combined with SMC reactions using PDMS membranes (**Figure 40**). The higher concentration of intermediate aryl bromide afforded by FlHal CLEAs allowed palladium loading to be further reduced in some cases (2 mol%) and a broader range of products to be obtained (**Figure 42C)**.

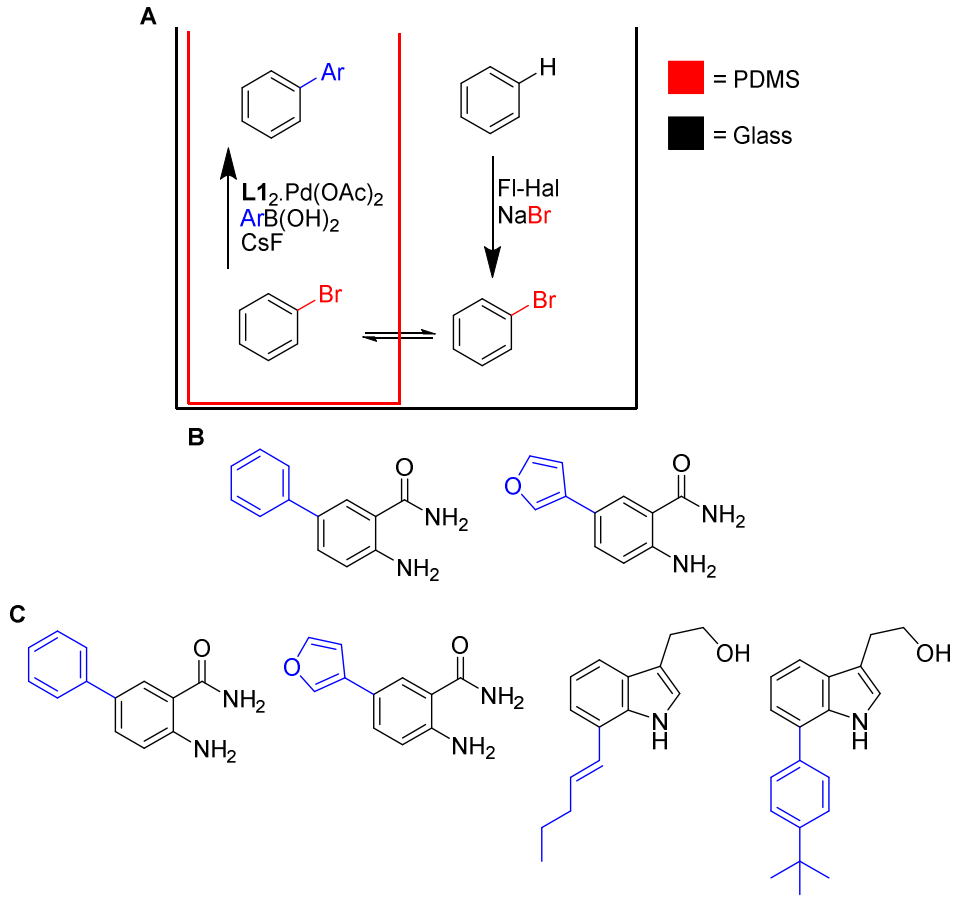
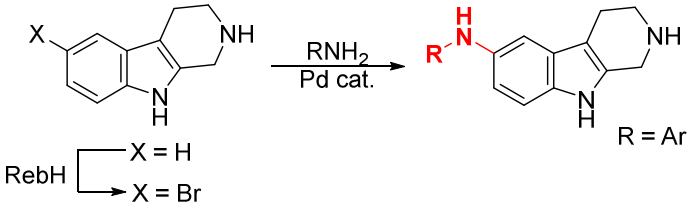


Figure 42: (**A**) Generalised approach to biocatalyst and chemocatalyst compartmentalisation using PDMS membrane to allow FlHal-SMC cascades (**B**) Scope of products synthesised with integrated FlHal-SMC reactions using PDMS compartmentalisation of purified FlHal enzymes and cofactors from the SMC reaction components. (**C**) Scope of products synthesised with integrated FlHal-SMC reactions using PDMS compartmentalisation of FlHal CLEA reactions from SMC reactions.

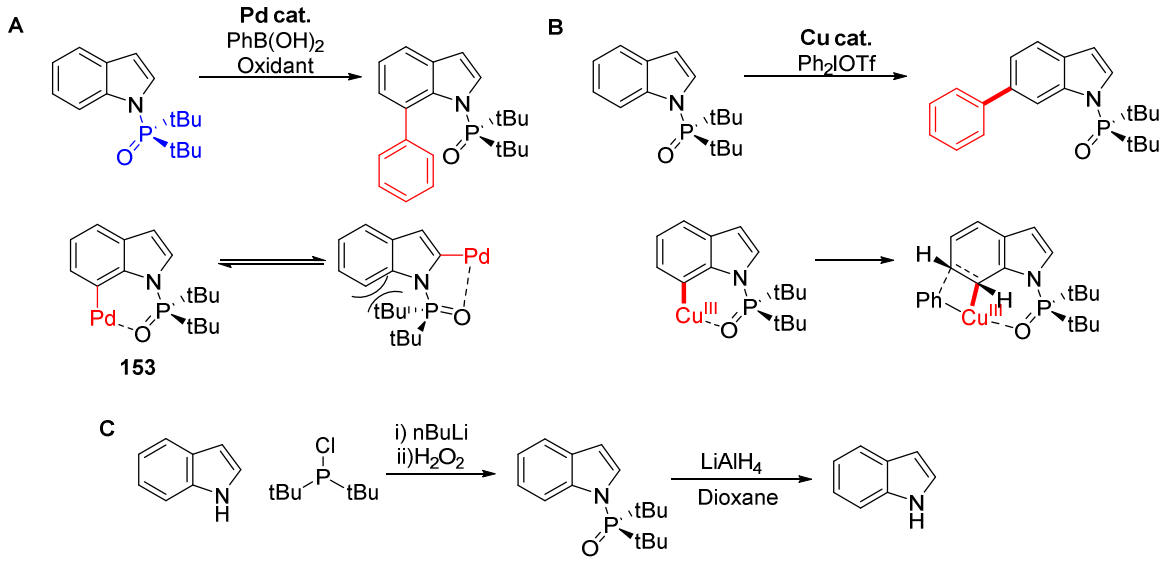
Despite allowing efficient FlHal SMC cascades without the need for any intermediary processing through effective compartmentalisation, the substrate scope of the method involving PDMS membranes was somewhat limited by the inability of substrates which carry charge at the reaction pH (e.g. substrates **106** and **139**) to effectively pass between the two compartments. The scalability of this process is also an issue, with the fabrication of PDMS thimbles for use in large scale reactions likely to be difficult. Since the PDMS monomers are inexpensive however, reactor designs other than thimbles may allow this approach to be used on larger scales. After publication of this work, other groups reported approaches to the combination of FlHals with SMCs.[207,211] One such example involved the removal of FlHal CLEAs prior to addition of the SMC reagents.[207] This process was only demonstrated on tryptophan (**31**), rather than the obvious extension to other non-natural substrates. Another report used engineered FlHal variants for the late-stage halogenation of bioactive molecules, which were then isolated by solvent extraction.[211] After removal of solvent, the

crude aryl bromide was subjected to a separate SMC reaction. Since this method requires the isolation of intermediate bromide, this cannot be classified as a single pot reaction unlike the strategy using PDMS compartmentalisation. The use of engineered FlHal variants does demonstrate the value of these enzymes for the late-stage functionalisation of complex biologically-active scaffolds, however.



Scheme 49: Example of an amination reaction facilitated by the isolation of a aryl bromide synthesised using a FlHal prior to subjection to a Buchwald-Hartwig amination in a separate reaction.[211]

Ultimately, the regioselectivity of different FlHals was exploited in this thesis to allow the controlled, regioselective, formation of C-C bonds at the C7, C6 and C5 positions of indolic substrates (**Scheme 47**). Whilst this study was being conducted, direct access to these positions of indole using C-H activation was unprecedented.[67] Around the time this approach was published however, two reports of the C7 and C6 arylation of indoles were published using a complementary strategy.[212,213] These two examples rely upon substrate-control, with a bulky directing group on the indole nitrogen employed to direct the position of metallation. In one example, the steric bulk of the directing group is thought to lead to stabilisation of C7-metallated intermediate **153** over the C2 counterpart (**Scheme 50A**).[213] Use of this directing group therefore allows the direct C7 arylation of indolic substrates.[213] The same group later reported use of this strategy for the C6 arylation of indoles.[212] After C7 insertion of Cu, it is though that a Heck-type intermediate is formed which allows C6 arylation (**Scheme 50B**).[212] These two strategies, similar to those described in **Scheme 33A**, illustrate the concept of substrate control very well. In each case either innate substrate electronic effects or intramolecular coordination between metal and substrate control the position of functionalisation. The need for such effects however, means that prior functionalisation of substrates is often required. Since these directing groups are usually not desired in the final compound, they often need to be removed also – further reducing the efficiency of the process overall. In these particular cases (**Schemes 50A** and **50B**), the conditions required for addition and removal of the directing group are quite harsh (**Scheme 50C**) and would preclude the presence of certain sensitive functionalities in substrates.



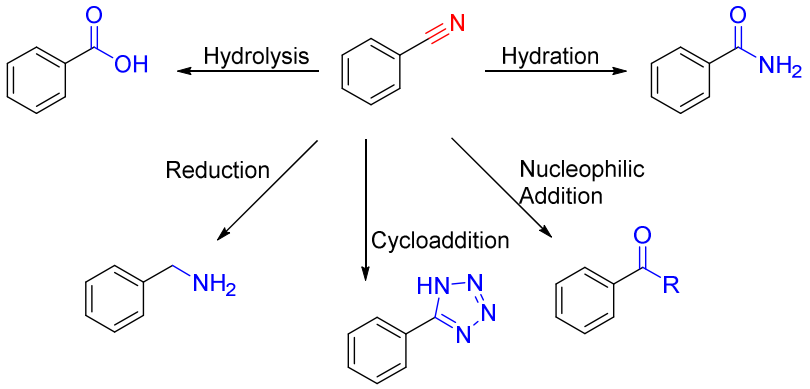
Scheme 50: (**A**) C7 arylation of indole using the bulky directing group (highlighted in blue), showing how orientation of this group determines regioselectivity of metallation.[213] (**B**) C6 arylation of indole using the same directing group showing the proposed Heck-like intermediate.[212] (**C**) Reported method for the installation and removal of the required directing group for these reactions.[212,213]

The appeal of the FlHal-SMC reactions developed herein is that directing groups on substrate are not required. The control of regioselectivity comes from how the substrate interacts with the FlHal active site, meaning that different positions of identical substrates can be activated. Although such specificity does allow good regioselectivity, it limits the substrate scope of the process somewhat. The development of structure-guided and screening approaches to modifying the substrate scope and regioselectivity of FlHals does offer further promise that these methods could be of more generality however.[155,156,158,162] Additionally the scalability of the integrated FlHal-SMC reactions is currently limited compared to the more traditional chemocatalytic methods, due to the poor productivity of the FlHals. Although the use of stabilised CLEAs has been demonstrated to circumvent this somewhat, there is still some way to go until these integrated catalysis methods can match what can be achieved by more traditional means.

4. Halogenase-Mediated Regioselective Cyanation

4.1 Introduction

The nitrile (-C≡N) is a versatile functional group, susceptible to hydration, reduction, hydrolysis, cycloaddition and nucleophilic addition – allowing access to a range of other useful moieties (**Scheme 51**).[214,215] Key to this plethora of reactivity is the highly polarised nature of the C≡N bond, as well as its significant π-character. The resulting electropositive C is available for attack by an array of nucleophiles, affording the manifold of transformations possible.



Scheme 51: Examples of arylnitrile application in synthesis.

Aryl nitriles are also of importance in their own right, with a number of pharmaceutical agents containing this moiety (**Figure 43**). The relatively small size of the C≡N group means that its introduction to a pharmacophore has minimal impact upon steric interactions with target.[216] The lack of reactivity under physiological conditions means that such groups are usually metabolically stable also.[216] The electronegative nitrogen of the C≡N can influence binding affinity however, by acting as a hydrogen bond acceptor with the biological target.[216] Such interactions have been observed by crystallography in a number of examples (**Figure 44**).[216-218]

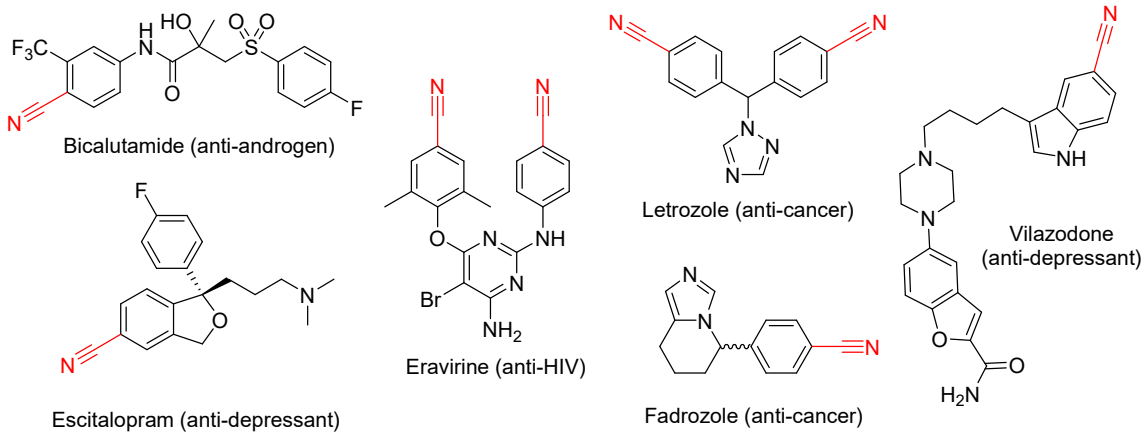


Figure 43: Nitrile-containing compounds of pharmaceutical importance.

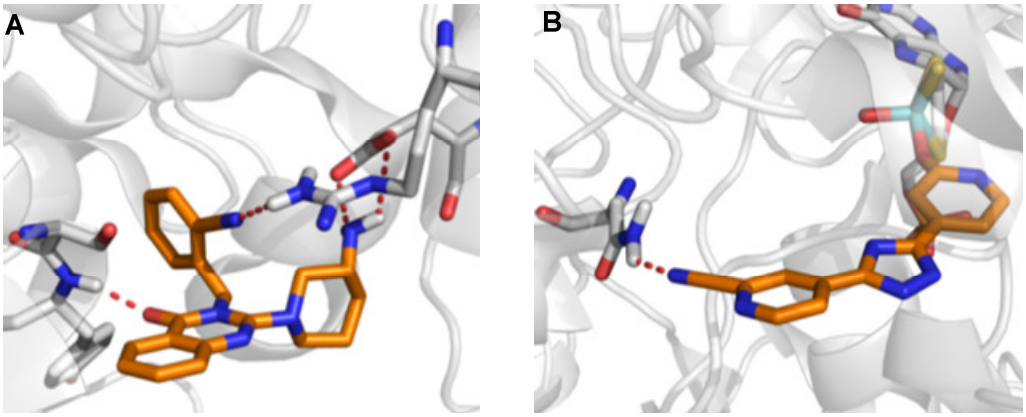
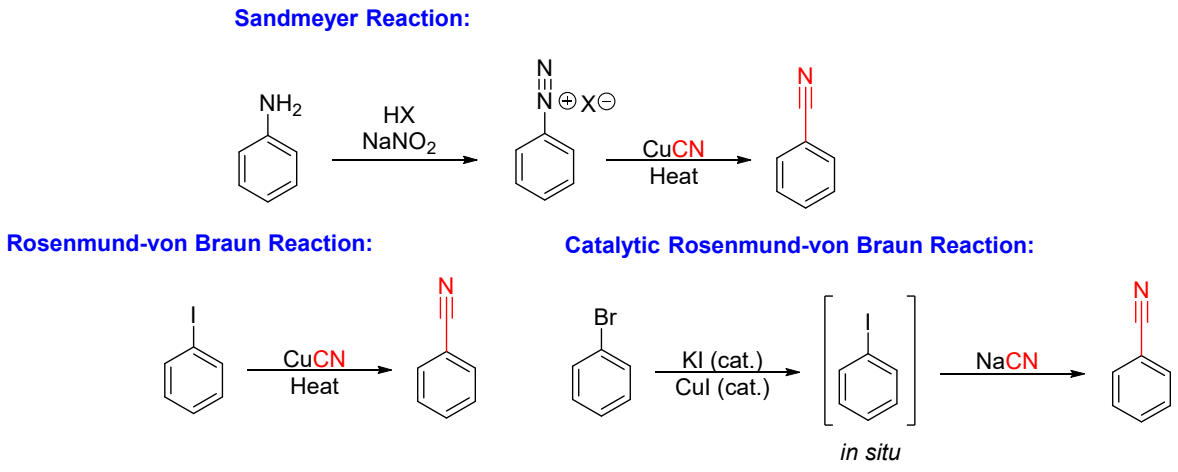


Figure 44: Crystal structures of nitrile-containing compounds bound to protein targets of medicinal importance with the proposed nitrile hydrogen bond highlighted. (**A**) Dipeptidyl Peptidase IV (PDB 2ONC).[217] (**B**) Xanthine oxidoreductase (PDB 1V97).[218]

4.1.1 Traditional Methods of Arylnitrile Synthesis

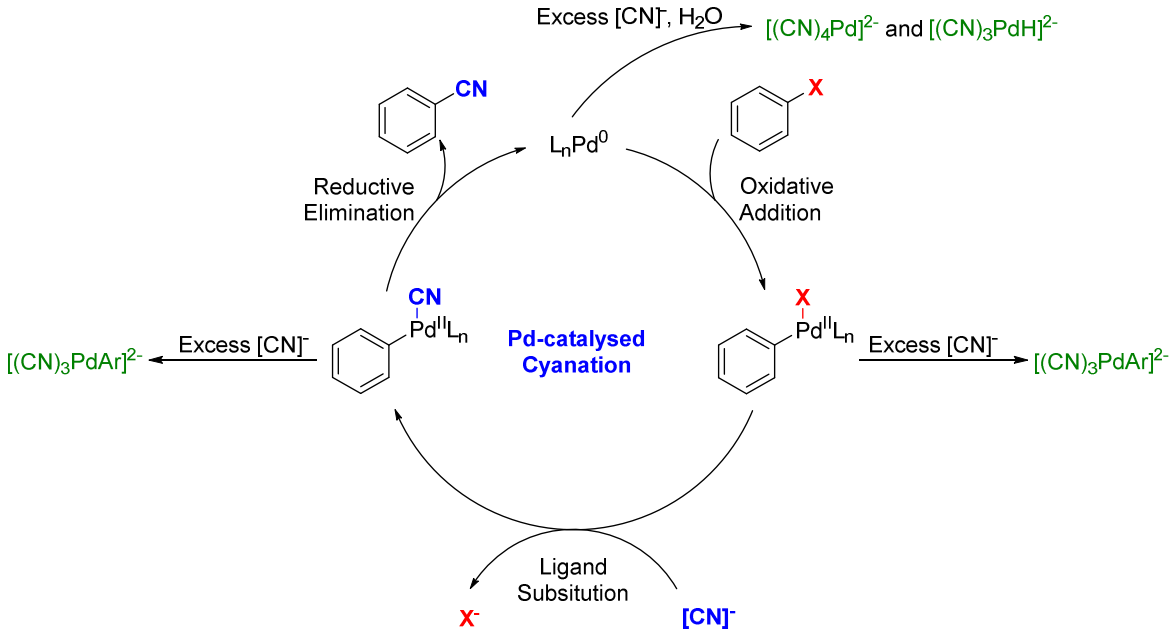
Traditional methods for the synthesis of aryl nitriles involve the interconversion of a pre-existing functional group (**Scheme 52**). As with similar methods discussed earlier for the synthesis of aryl halides, these methods require the installation of this functional group prior to cyanation – meaning that issues of regioselectivity are often encountered earlier in synthetic routes. One such example, the Sandmeyer reaction, involves the diazotisation of an aniline followed by heating with a stoichiometric amount of CuCN.[63] The Rosenmund-von Braun reaction is similar, involving the heating of aryl iodides with a stoichiometric amount of CuCN.[214] As such, both of these traditional approaches generate a stoichiometric amount of heavy metal waste, which limits their applicability on an industrial scale. A more recent modification of this reaction allows the conversion of aryl bromides, via *in situ* conversion to aryl iodides, using catalytic amounts of Cu.[219,220] Despite work on the optimisation of reaction conditions and ligands, these Cu-promoted methods typically require significant heat input to proceed in reasonable yield.[220] Such a requirement therefore limits the tolerance of certain functionalities to these conditions, reducing their generality, and is not desirable from the green perspective.



Scheme 52: Traditional methods for the synthesis of aryl nitriles.[214]

4.1.2 Palladium-catalysed Aryl nitrile Syntheses

Palladium catalysis has been investigated as an alternative to the copper-catalysed cyanation reactions above. These reactions are believed to proceed in a manner analogous to other Pd-catalysed cross couplings (**Scheme 26**), except that ligand substitution with $^{-}$CN occurs rather than transmetallation (**Scheme 53**).[214] Reductive elimination of this Pd^{II} complex then affords aryl nitrile product and regenerates Pd^{0} catalyst. Application of these reactions was initially limited by a lack of reproducibility, believed to come about from the propensity of $^{-}$CN to poison Pd-intermediates of the catalytic cycle (**Scheme 53**).[221,222]



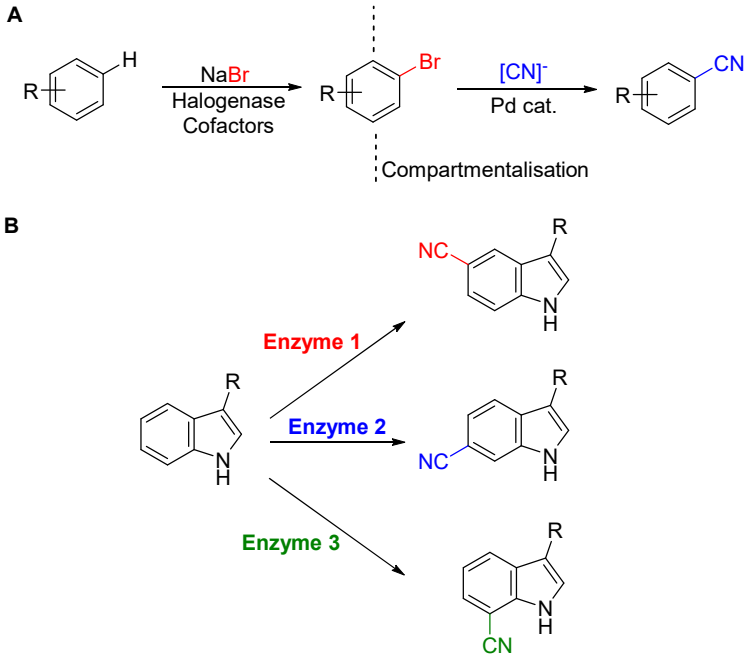
Scheme 53: Proposed catalytic cycle of the Pd-catalysed cyanation reaction, showing potential deactivation pathways. Inactive "poisoned" Pd complexes are highlighted in green.[214,221,222]

Given the possibility of catalyst poisoning by excess cyanide, much of the progress made towards reliable cyanation protocols has involved mediating the availability of cyanide during the reaction course.[214] The first demonstration of a palladium-catalysed cyanation reaction utilised KCN, and was found to accept only strongly activated aryl iodides and bromides – presumably due to a lack of available cyanide in organic solvents.[223] Subsequent methodologies for palladium-catalysed cyanation reactions using alkali cyanide salts found bulky amine co-catalysts to afford improved yields – possibly by acting as a phase-transfer agent to allow greater availability of cyanide.[224] Other methodologies have utilised tributyltin chloride for a similar purpose, with the formation of an intermediary tin cyanide complex thought to promote the availability of CN for catalysis.[225] Attention was later turned to the use of alternative cyanide sources. One popular example is $Zn(CN)_2$, which was found to allow the use of lower Pd loadings than methods using alkali cyanide sources.[226] The highly covalent nature of the Zn-CN bond is thought to limit the concentration of cyanide during the reaction, thereby reducing the potential for catalyst poisoning.[214,226] Although transformations involving $Zn(CN)_2$ are well-reported, the toxicity of this reagent limits its applicability on scale. For this purpose, $K_4[Fe(CN)_6]$ has been explored by many groups. The strong σ- and π-bonding of $^{-}$CN

to Fe^{II} in this reagent significantly reduces its toxicity compared to other sources, but typically requires heating to liberate sufficient CN from the Fe-complex for ligand substitution.[214,215,227,228]

4.1.3 Conclusions

Because of the importance of arylnitriles as synthetic intermediates and in pharmaceuticals, methods for their preparation have been subject to much investigation. Some of the most recent approaches to these compounds have involved palladium-catalysed cyanation reactions of aryl halide compounds.[214] As previously discussed, the preparation of these aryl halides with the required regioselectivity is oftentimes not straightforward and can require the use of toxic halogen sources. Given that the integration of palladium catalysis with biocatalytic halogenation had allowed a regioselective arylation, vinylation and heteroarylation to be developed (Chapter 3), it was envisaged that an analogous combination could be extended to the direct, regioselective, cyanation of aromatic substrates (**Scheme 54**).



Scheme 54: (**A**) Proposed combination of FlHals with Pd-catalysed cyanation reactions for C-H cyanation. (**B**) Regiochemistry of cyanated indoles which may be accessed using the proposed FlHal-cyanation cascade.

Like the FlHal-SMC cascade, this proposed combination (**Scheme 54A**) could be viewed as a C-H activation reaction. The current methods for C-H cyanation involve substrate-control strategies, whereby intramolecular coordination or innate substrate electronics are exploited to control the position of cyanation.[67] As with many other C-H activation reactions, the phenyl-2-pyridine moiety (**4**) has been utilised for C-H cyanation since this allows metallation at the *ortho*- position with good regioselectivity (**Figure 45A**).[229] In a recent example, the cyanation of benzyl C-H bonds has been achieved by exploiting the increased stability of radicals at this position (**Figure 45B**).[230] Methods for the C-H cyanation of indoles have also been reported, with most exploiting the innate nucleophilicity of the indole C3 for C3 cyanation (**Figure 45C**).[67,231,232] More recently, a 2-pyridine

directing group was utilised to allow the C2 cyanation of indoles (**Scheme 45D**).[233] Since current methods for the C-H cyanation of indoles are limited, it was hoped that the regiodivergent nature of the FIHals used in the FIHal-SMC reaction could allow the direct cyanation of indole positions which are not accessible by other C-H activation methods – through the combination of FIHals with Pd-catalysed cyanations in one-pot integrated reactions (**Scheme 54B**). Given the synthetic versatility of the nitrile group, it was also hoped that subsequent reactions could also be integrated to expand the scope of this reaction beyond regioselective cyanation. This chapter will describe the tolerance and integration of the FIHal and cyanation reactions, towards developing the regioselective C-H cyanation reaction (**Scheme 54A**). Attempts are also made towards a regioselective amidation reaction through the incorporation of additional biocatalysts into the cascade.

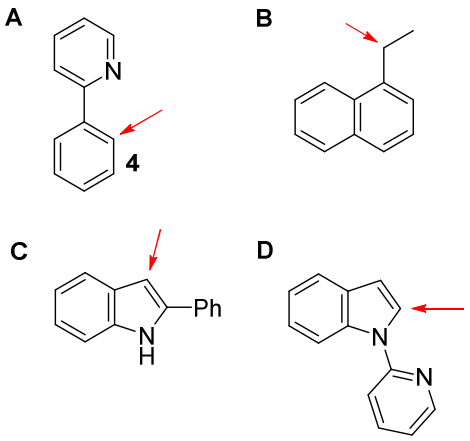


Figure 45: Motifs used in metal-catalysed C-H-cyanation reactions. The red arrow shows the position of cyanation. (**A**) 2-phenyl pyridine (**4**) for *ortho*-cyanation.[229] (**B**) Benzyl C-H cyanation.[230] (**C**) Indole C3 cyanation.[234] (**D**) Indole C2 cyanation.[233]

4.2 Results and Discussion

4.2.1 Integrated Cyanation of Anthranilamide

To assess the feasibility of an integrated FIHal-cyanation reaction, the palladium-catalysed cyanation of pure 5-bromo anthranilamide (**84**) was screened with a number palladium sources and ligands previously found effective in SMC reactions of this substrate (**Table 11**, entries 1 to 4). $K_4[Fe(CN)_6]$ and 1,8-diazabicyclo[5.4.0]undec-7-ene (**154**, DBU) were selected as cyanide source and base respectively, based on previous reports of the cyanation of heteroaryl bromides using these conditions.[227] Of these combinations, only the combination of Na_2PdCl_4 with tppts (**90**) afforded any conversion to a new product which could be identified by analytical HPLC. LC-HRMS subsequently identified *m/z* peaks consistent with the molecular formula of **155**. A base screen was subsequently conducted and found Na_2CO_3 afforded increased conversion of **84** to **155** compared to DBU (**Table 11**, entry 5).

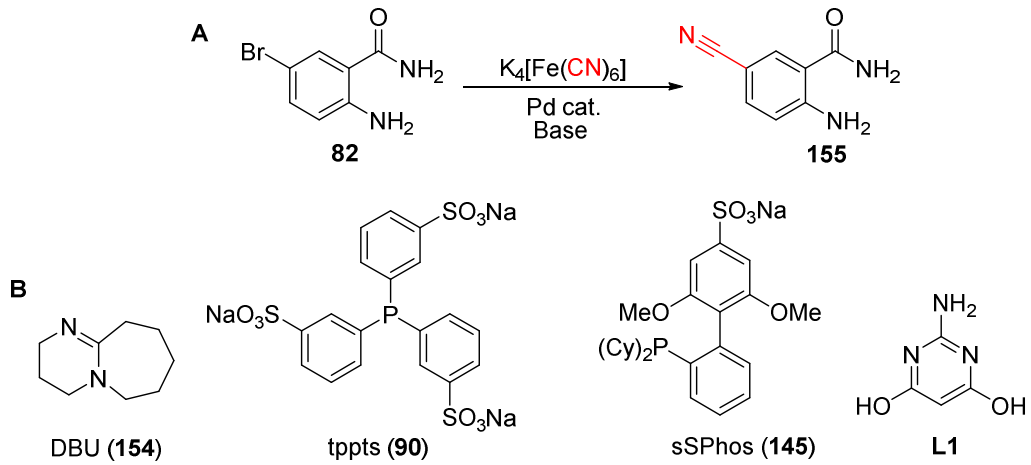
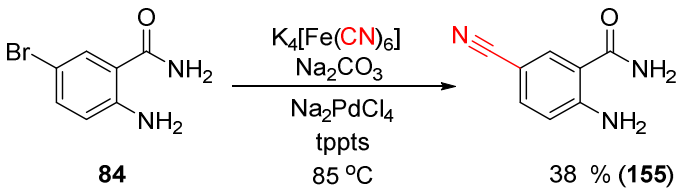


Figure 46: (**A**) Cyanation of 5-bromo anthranilamide (**84**) used for screening of reaction conditions. (**B**) Structure of DBU and ligands used during screening of cyanation conditions.

Entry	Pd catalyst	Base	Solvent	Conversion (155, %)*
1	$Pd^0(PPh_3)_4$	DBU	$^tBuOH/H_2O$	N.C.
2	$Na_2PdCl_4.tppts_2$	DBU	H_2O	31
3	$Na_2PdCl_4.sSphos_2$	DBU	H_2O	N.C.
4	$\mathbf{L1}_2.Pd(OAc)_2$	DBU	H_2O	N.C.
5	$Na_2PdCl_4.tppts_2$	Na_2CO_3	H_2O	41
6	$Na_2PdCl_4.tppts_2$	K_3PO_4	H_2O	11
7	$Na_2PdCl_4.tppts_2$	CsF	H_2O	N.C.
8	$Na_2PdCl_4.tppts_2$	NaOH	H_2O	N.C.

Table 11: Catalyst and base screening for the cyanation of **84** to **155** (**Figure 46A**). Conditions: 5-bromo anthranilamide (**84**) (3.0 mM), Pd catalyst (10 mol %), $K_4[Fe(CN)_6]$ (40 mol %), base (2.0 eq), 85 °C, overnight. *Determined by analytical HPLC using a calibration curve of **155**. †4:1 $^tBuOH:H_2O$. N.C. = no conversion.

To verify these optimisation attempts, the cyanation of **84** was upscaled to allow isolation of the proposed aryl nitrile product. The optimised conditions (**Table 11**, entry 5) allowed the isolation of **155** in 38 % yield from **84**. ^1H NMR confirmed the regioselectivity of this product, and HRMS confirmed a molecular formula consistent with the proposed constitution of **155**. The presence of the C≡N functionality is evident from both the ^{13}C NMR and infrared (IR) spectra of **155** (**Appendix 2**).



Scheme 55: Isolated yield from the cyanation of 5-bromo anthranilamide (**84**). Conditions: Na_2PdCl_4 (10 mol %), tppts (20 mol %), $K_4[Fe(CN)_6]$ (40 mol %), Na_2CO_3 (2 eq), H_2O, 85 °C, overnight.

In an attempt to improve conversion of the cyanation reaction further, the loading of $K_4[Fe(CN)_6]$ was varied (**Table 12**) considering that excess cyanide has been found to have a poisoning effect on these reactions.[221,222] A $K_4[Fe(CN)_6]$ loading of 20 mol % was found to afford higher conversion to **155** than either 40 or 60 mol % (**Table 12**, entries 1 to 3). This is logical considering the increased solubility of ^{-}CN in completely aqueous systems, meaning that poisoning of Pd intermediates is increasingly likely. Further reduction of $K_4[Fe(CN)_6]$ loading to 10 mol % afforded a reduction in conversion to cyanated product however, suggesting that ^{-}CN availability had become a limiting factor.

To assess the tolerance of these conditions with respect to the cofactors required for the FlHal bromination reaction, the cyanation of **84** was run in the presence of NADH (**103**), FAD (**29**) and NaBr (**Table 12**, entries 5 to 7). Whilst conversion to **155** was largely unaffected by the presence of NADH, FAD did appear to have a deleterious effect. Interestingly, the addition of NaBr (**Table 12**, entry 7) afforded an increase in conversion relative to the reaction without additive at 40 mol % $K_4[Fe(CN)_6]$ (**Table 12**, entry 1). This suggests that NaBr may be somehow mediating the concentration or availability of ^{-}CN in the reaction. Binding of Na^+ to liberated ^{-}CN could slow ligand substitution with the intermediary Pd complex, thereby reducing the likelihood of catalyst inhibition. Br^- could have a similar effect, acting competitively with ^{-}CN for binding to Pd hence limiting the extent of inhibition. Given the very high affinity of ^{-}CN for Pd it is unlikely that Br^- competition will have any significant effect in this regard, although it should be noted that Br^- is present at a 25-fold excess relative to ^{-}CN. Ligand substitution of Br^- for Fe-bound ^{-}CN may however increase the stability of the resulting decyano-iron complex – thereby favouring ^{-}CN liberation from Fe. Although the reason for the positive effect of NaBr upon cyanation conversion is interesting, it was considered of little consequence to the development of FlHal-cyanation cascades at the time.

Table 12: Effect of varying CN loading and addition of FlHal biotransformation components upon the conversion of **84** to **155** (**Figure 46A**) Conditions as reported using Na_2PdCl_4 and tppts as prior (**Table 11**, entry 5) except the CN loading and additive concentration was varied as stated. *Determined by analytical HPLC using a calibration curve of **155**.

Entry	$K_4[Fe(CN)_6]$ Loading (mol %)	Additive (Concentration)	Conversion (155, %)*
1	40	None	38
2	60	None	17
3	20	None	52
4	10	None	16
5	40	NADH (100 µM)	32
6	40	FAD (10 µM)	14
7	40	NaBr (30 mM)	66

With varied results relating to the tolerance of these conditions with respect to the FlHal biotransformation components, the cyanation of a PyrH CLEA biotransformation supernatant was attempted using these standard conditions (**Figure 47A**). Using 10 mol % of the Pd catalyst and

40 % $K_4[Fe(CN)_6]$ afforded only 14 % conversion to cyanated product **155**, however. The low conversion and small scale of these reactions meant it was infeasible to isolate this product, although analytical HPLC showed the proposed cyanated product to co-elute with the standard of **155** prepared above (**Figure 55B**), confirming the identity of this product. The UV spectrum of this peak (**Figure 47C**) also matched the spectrum recorded of the standard of **155** - which is distinct from bromo intermediate **84** by a strong absorbance at around 280 nm that is absent in UV spectra of **84**.

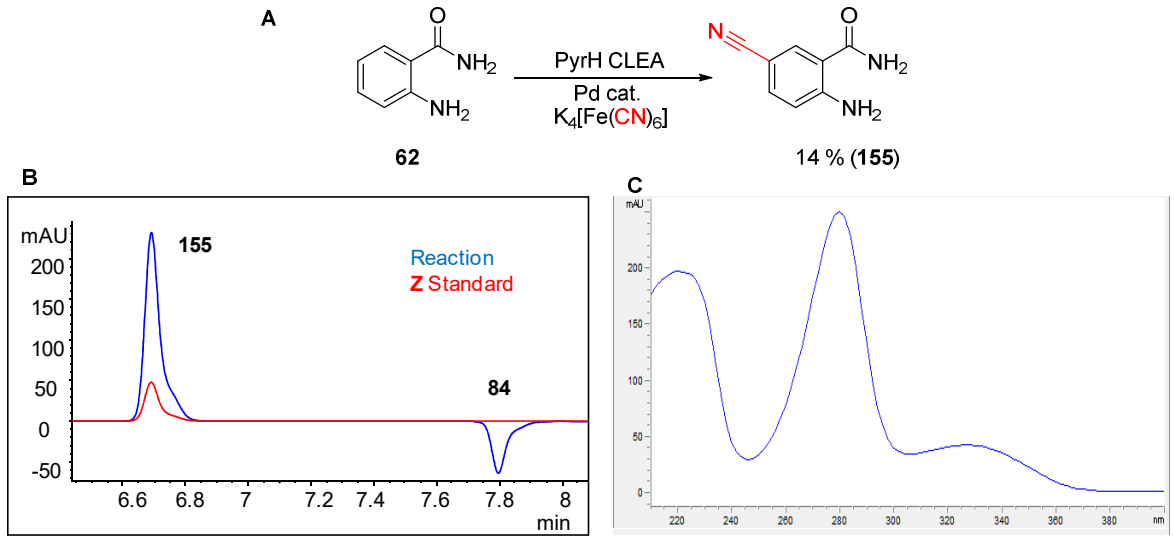


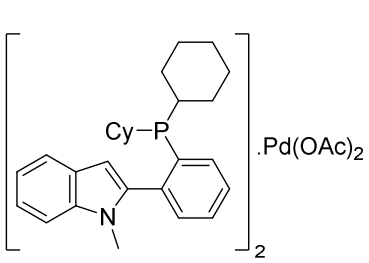
Figure 47: (**A**) Conversion of an attempted FlHal-cyanation reaction of anthranilamide (**62**) using PyrH CLEAs. (**B**) HPLC chromatogram of the above cyanation overlaid with a standard of **155** showing UV absorbance measured at 280 nm. (**C**) UV spectrum of the proposed cyanated product from this integrated cyanation.

In an attempt to overcome this low conversion, the loading of both Pd catalyst and $K_4[Fe(CN)_6]$ was varied in the cyanation of PyrH CLEA biotransformation supernatants containing **84** (**Table 13**). An increased loading of CN afforded a slight increase in conversion to **155** (**Table 13**, entry 2), however above 70 mol % no **155** could be detected (**Table 13**, entry 1). Below 50 mol % of CN, conversion to **155** was significantly reduced (**Table 13**) suggesting that ligand substitution to Pd had become a yield-limiting factor. This would also explain the improved conversion upon increasing CN loading, suggesting that $^{-}$CN may bind to a biotransformation component. Increased Pd catalyst loading afforded a slight increase in conversion at a range of CN loadings (**Table 13**), although none afforded conditions amenable to upscale of the reaction with this catalyst.

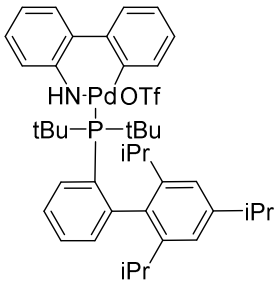
Table 13: Effect of altering CN and Pd catalyst loading upon conversion of **62** to **155** in an integrated cyanation reaction of **62** using PyrH CLEAs. Conditions as prior (**Table 11**, entry 5) except CN and Pd loading is varied as stated. *Determined by analytical HPLC using a calibration curve of **155**. N.C. = no conversion.

Entry	$K_4[Fe(CN)_6]$ Loading (mol %)	Na_2PdCl_4/tppts loading (mol %)	Conversion (155, %)*
1	100	10	N.C
2	70	10	20
3	50	10	13
4	40	10	4
5	30	10	8
6	20	10	7
7	10	10	8
8	20	20	19
9	50	20	20
10	100	20	14

At this point it was considered that a significant change in the catalyst system may be required. Pd-catalysed cyanation reactions in completely aqueous media are not commonplace because of the high solubility of $^{-}$CN, meaning there is increased potential for catalyst poisoning (**Scheme 53**). Additionally, the tppts ligand employed in these cases (**90**) was considered potentially unsuitable for these reactions since the net negative charge of the ligand could cause electrostatic repulsion of incipient $^{-}$CN – thereby inhibiting ligand substitution. The group of Buchwald have reported a number of Pd-catalysed cyanation protocols using the bulky palladacycle catalysts **156** and **157**.[215,228] The steric bulk of the large ligands is thought to contribute to the effective reductive elimination of Pd^{II}-CN intermediates and hence promote effective cyanation cycles. The hydrophobic nature of these ligands means that they are employed in mixtures of organic solvent and water, which leads to a system with limited solubility of the cyanide source and hence reduces the possibility of catalyst poisioning. Following these reports, the cyanation of pure aryl bromide (**84)** was again attempted on an analytical scale using these ligands and mixtures of water and organic solvent. In each of these cases however, poor conversion of **84** to cyanated product **155** was detected (**Table 14**).



CMPhos.Pd(OAc)$_2$ (**156**)



^{t}BuXPhos-Pd-G3 (**157**)

Figure 48: Bulky palladacycle catalysts screened for cyanation of aryl bromides.

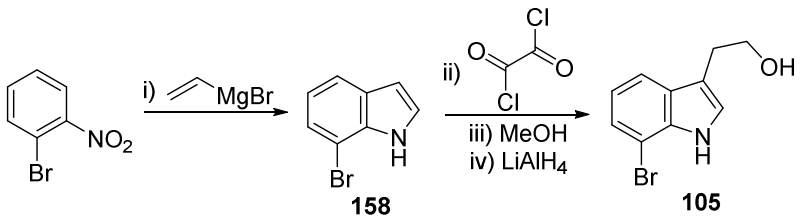
Table 14: Screening of bulky palladacycle catalysts (**Figure 48**) for the cyanation of 5-bromo anthranilamide (**84**) to **155** (**Figure 46A**). Conditions: Pd catalyst (10 mol %), $K_4[Fe(CN)_6]$ (50 mol %), Na_2CO_3 (2 eq), 5:1 water:organic, 80 ^oC, overnight. *Determined by analytical HPLC using a calibration curve of **155**. N.C. = no conversion.

Entry	Pd catalyst	Solvent	Base	Conversion (155, %)*
1	CMPhos.Pd(OAc)$_2$	H_2O	Na_2CO_3	1
2	CMPhos.Pd(OAc)$_2$	H_2O:ACN	Na_2CO_3	<1
3	CMPhos.Pd(OAc)$_2$	H_2O	None	N.C
4	CMPhos.Pd(OAc)$_2$	H_2O:ACN	None	2
5	CMPhos.Pd(OAc)$_2$	H_2O:THF	None	2
6	^{t}BuXPhos-Pd-G3	H_2O	Na_2CO_3	11
7	^{t}BuXPhos-Pd-G3	H_2O:ACN	Na_2CO_3	6
8	^{t}BuXPhos-Pd-G3	H_2O:THF	Na_2CO_3	21
9	^{t}BuXPhos-Pd-G3	H_2O:THF	None	21

Consistently poor conversions using **84** as substrate suggested that the reason may lay with the substrate itself rather than the specific catalytic conditions employed. Given that the 5- position of 5-bromo anthranilamide (**84**) is quite electron-rich due to the amine moiety, it could be that reductive elimination of a formally negatively-charged fragment ($^-$CN) to this position is retarded by unfavourable electrostatic interactions. Even the bulky catalysts **156** and **157** appeared inefficient at promoting this reductive elimination step, which could alternatively be because the polar substrate **84** is incapable of entering the organic phase where **156** and **157** reside. The polar amide and aniline moieties of **84** could also contribute to electrostatic repulsion of $^-$CN during ligand substitution, reducing the likelihood of productive cyanation cycles. Given that $Pd^0(PPh_3)_4$ and Na_2PdCl_4.tppts$_2$ were both effective at catalysing SMC reactions of **84**, it is unlikely that the oxidative addition step shared by both reactions is yield-limiting in these cases.

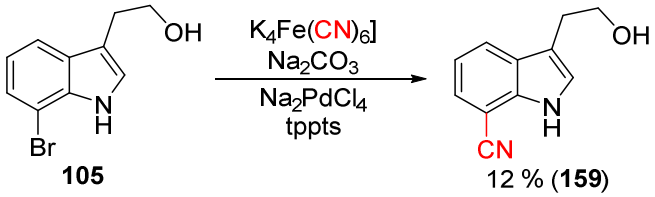
4.2.2 Integrated Cyanation of Indoles

With an aversion to further increasing Pd loadings to promote the cyanation of **84**, other substrates were investigated. The indolic substrate tryptophol (**104**) was selected because of its successful integration into FlHal-SMC reactions previously, and because the introduction of further indolic substrates would allow realisation of the regiodivergent indole cyanation reaction (**Scheme 54B**). Given that RebH, a tryptophan 7-halogenase, was known to be capable of efficiently affording 7-bromo tryptophol (**105**), a standard of **105** was prepared from 7-bromo indole (**158**) according to literature protocols to facilitate the screening and optimisation of cyanation reaction conditions (**Scheme 56**).[235,236]



Scheme 56: Synthesis of 7-bromo tryptophol (**105**) as a standard for use in the screening of cyanation conditions. Conditions: i) Vinyl magnesium bromide (3.0 eq), THF, - 42 °C, 1 hr, 30 % yield; ii) $(COCl)_2$ (3.0 eq), Et_2O, 0 °C to rt, 1 hr; iii) MeOH (5.0 eq), rt, 1 hr; iv) $LiAlH_4$ (4 eq), THF, 0 °C to rt, overnight, 18 % yield over three steps.

With a reasonable amount of **105** in hand, focus was turned to the cyanation of this substrate. Initial screening using the same conditions as 5-bromo anthranilamide (**Table 11**, entries 1 to 4) did not reveal any obvious conversion to the expected cyanated product **159** by LC-HRMS however. Use of Na_2CO_3 as base in place of DBU allowed the isolation of **159** in 12 % yield upon upscale. As with **155**, the regioselectivity of this product was identified by ^{1}H NMR, whilst the presence of the C≡N group was confirmed by HRMS, ^{13}C NMR and IR spectra (**Appendix 2**). Although yields of cyanated product were very poor, this product was used as an analytical standard of **159** for further screening and optimisation attempts.



Scheme 57: Isolated yield of **159** from the cyanation of 7-bromo tryptophol (**105**) using conditions previously optimised from 5-bromo anthranilamide (**84**). Conditions: $K_4[Fe(CN)_6]$ (40 mol %), Na_2CO_3 (2.0 eq), Na_2PdCl_4 (10 mol %), tppts (**90**, 20 mol %), H_2O, 85 °C, overnight.

With the hope of identifying conditions capable of affording better conversion to cyanated product **159**, similar screening reactions were run using the bulky palladacycle catalysts **156** and **157** based on Buchwald's previous reports.[215,228] Whilst the CMPhos-based catalyst **156** did not afford any appreciable improvement in conversion of **105** to **159** under a number of conditions (**Table 15**, entries 1 to 5), ^{t}BuXPhos-Pd-G3 allowed high conversions to **159** to be achieved using a 5:1 mixture of water and THF (**Table 15**, entries 8 and 9). The requirement for this solvent mixture is reported in the literature - indeed significant deviation from this ratio is strongly discouraged as it is considered key for controlling the amount of $^-$CN present during catalysis.[228] Also in agreement with the literature report, it was found that no base is necessary for the cyanation in this case (**Table 15**, entry 9) – with $^-$CN thought to act in its place to activate Pd^{II} pre-catalyst.[228] It should be noted that the literature report utilised $Zn(CN)_2$ as cyanide source,[228] however this reagent was avoided in our attempts because of its increased toxicity compared to $K_4[Fe(CN)_6]$. Despite this deviation, **159** could be isolated in 73 % yield upon upscale using the optimised conditions (**Table 15**, entry 9).

Table 15: Screening of bulky palladacycle catalysts (**Figure 48**) for the cyanation of 7-bromo tryptophol (**105**) to **159** (**Scheme 57**). Conditions: Pd catalyst (10 mol %), $K_4[Fe(CN)_6]$ (50 mol %), Na_2CO_3 (2 eq), 5:1 water:organic, 80 ^oC, overnight. *Determined by analytical HPLC using a calibration curve of **159**. N.C. = no conversion. †Isolated yield = 73.

Entry	Pd catalyst	Solvent	Base	Conversion (159, %)*
1	CMPhos.Pd(OAc)$_2$	H_2O	Na_2CO_3	13
2	CMPhos.Pd(OAc)$_2$	H_2O:ACN	Na_2CO_3	9
3	CMPhos.Pd(OAc)$_2$	H_2O	None	5
4	CMPhos.Pd(OAc)$_2$	H_2O:ACN	None	2
5	CMPhos.Pd(OAc)$_2$	H_2O:THF	None	N.C
6	^{t}BuXPhos-Pd-G3	H_2O	Na_2CO_3	21
7	^{t}BuXPhos-Pd-G3	H_2O:ACN	Na_2CO_3	45
8	^{t}BuXPhos-Pd-G3	H_2O:THF	Na_2CO_3	70
9	^{t}BuXPhos-Pd-G3	H_2O:THF	None	78†

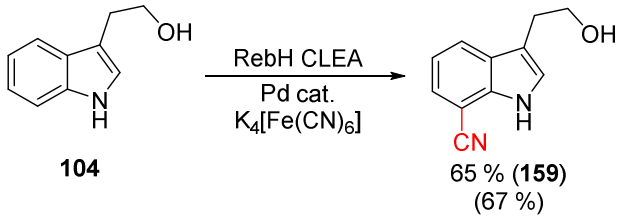
Encouraged by these results, the tolerance of this system with respect to the FlHal biotransformation cofactors was determined to assess the feasibility of the desired FlHal-cyanation cascade (**Scheme 54**). The cyanation was found to proceed well in the presence of NaBr, FAD and NADH (**Table 16**). Notably, conversion to cyanated product was not reduced upon introduction of FAD using these conditions, unlike the analogous reaction using Na_2PdCl_4 and tppts (**Table 12**, entry 6). This could be because the increased steric bulk of the ligands in **157** compared to tppts (**90**) may effectively protect the catalytic Pd centre from coordination to FAD through steric repulsion. The organic co-solvent employed may also help to prevent this interaction by limiting the solubility of FAD.

Table 16: Effect of different FlHal biotransformation components upon the conversion of **105** to **159** using the ^{t}BuXPhos-Pd-G3 catalysed cyanation conditions described above (**Table 15**). *Determined by analytical HPLC using a calibration curve of **159**.

Entry	$K_4[Fe(CN)_6]$ Loading (mol %)	Additive (Concentration)	Conversion (159, %)*
1	50	NaBr (30 mM)	79
2	50	NADH (100 µM)	76
3	50	FAD (10 µM)	72
4	100	NaBr (30 mM)	74

Given that these cyanation conditions appeared tolerant to the cofactors required for effective FlHal bromination, the cyanation of a RebH biotransformation supernatant containing **105** was attempted (**Scheme 58**). Although it would be desirable to use the PDMS-based compartmentalisation strategy used for FlHal-SMC reactions since this does not require intermediary processing, it was considered that the hydrophobic nature of **157** may cause flux of the catalyst into the PDMS membrane – preventing it from effectively catalysing the cyanation. Indeed the flux of phosphine-based palladium complexes into PDMS has been noted previously.[36] Gratifyingly however, the

cyanation of **104** using a RebH CLEA biotransformation supernatant was found to proceed in good conversion (67 %) and allowed the isolation of **159** in 65 % yield from **104**. In should be noted that in this case no intermediate bromide **105** could be detected by analytical HPLC after the cyanation.

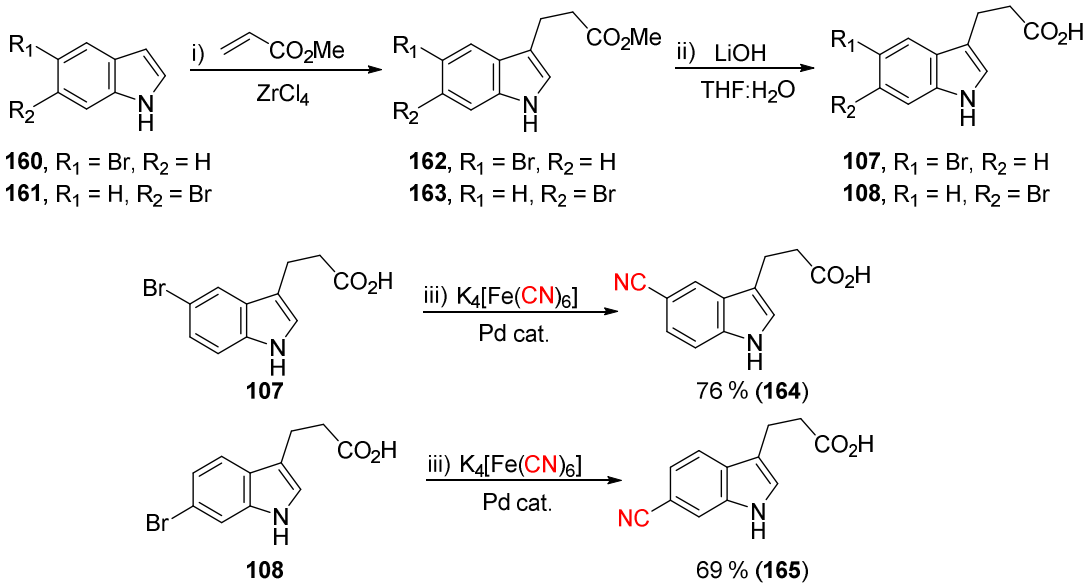


104

65 % (**159**)
(67 %)

Scheme 58: Integrated 7-cyanation of tryptophol (**104**) using cyanation of a RebH CLEA biotransformation supernatant. Conversion determined by analytical HPLC using a calibration curve of **159** is shown in brackets. RebH biotransformation conditions as previously described with substrate concentration 3.0 mM. Cyanation conditions: ^{t}BuXPhos-Pd-G3 (**157**, 10 mol %), $K_4[Fe(CN)_6]$ (50 mol %), THF:KPi (1:5), 85 °C, overnight.

With an effective 7-cyanation in hand, attention was turned to the desired 5- and 6-cyanation reactions. Given the effective arylation of indole-3-propionate (**106**) using the FIHal-SMC cascade, this substrate was selected for the integrated cyanation also. To assess the effectiveness of these conditions for cyanation of this substrate, the intermediate aryl bromides **107** and **108** were synthesised according to literature protocols (**Scheme 59**).[237,238] Although yields from the Michael addition of both 5- and 6-bromo indole (**160** and **161**) to methyl acrylate were poor, subsequent hydrolysis of these esters (**162** and **163**) afforded **107** and **108** in great enough quantities to attempt the cyanation. Using the same conditions applied for the cyanation of 7-bromo tryptophol (**Table 15**), the 5- and 6-cyano indole-3-propionic acids **164** and **165** could be afforded in good isolated yields (**Scheme 59**).



160, R_1 = Br, R_2 = H
161, R_1 = H, R_2 = Br

162, R_1 = Br, R_2 = H
163, R_1 = H, R_2 = Br

107, R_1 = Br, R_2 = H
108, R_1 = H, R_2 = Br

107

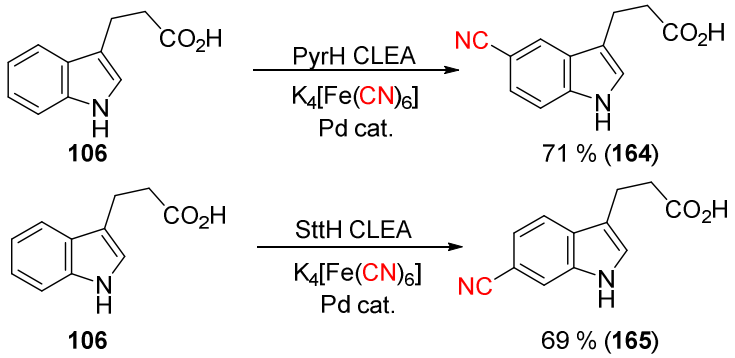
76 % (**164**)

108

69 % (**165**)

Scheme 59: Synthesis of brominated indole-3-propionic acid standards **107** and **108**, as well as the 5- and 6-cyano indole-3-propionic acid derivatives **164** and **165**. Conditions: i) methyl acrylate (2.0 eq), $ZrCl_4$ (0.5 eq), CH_2Cl_2, rt, 3 hrs; ii) LiOH.H_2O (15 eq), THF:H_2O (1:1), rt, overnight; iii) $K_4[Fe(CN)_6]$ (50 mol %), ^{t}BuXPhos-Pd-G3 (**157**, 10 mol %), THF:H_2O (1:5), 80 °C, overnight.

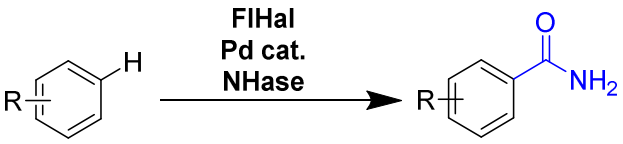
Given that the same cyanation conditions used for the integrated 7-cyanation appeared effective for the cyanation of **107** and **108** (**Scheme 59**), the integrated 5- and 6- cyanation of indole-3-propionic acid (**106**) was attempted. As with the FlHal-SMC arylation reported previously (Chapter 3) PyrH and SttH CLEAs were used to generate intermediate bromides **107** and **108** respectively. Subsequent cyanation of the supernatant from this biocatalytic bromination using the conditions optimised herein allowed the 5- and 6-cyano indole-3-propionic acids to be isolated in 71 % and 69 % yield respectively from indole-3-propionic acid (**106**). These integrated reactions represented the first direct, regioselective, cyanation of indole C5, C6 and C7 C-H bonds.



Scheme 60: Isolated yield from the integrated 5- and 6-cyanation of indole-3-propioniate (**106**) using PyrH and SttH CLEAs respectively. Conditions as previously described for the 7-cyanation of tryptophol (**104**, **Scheme 58**).

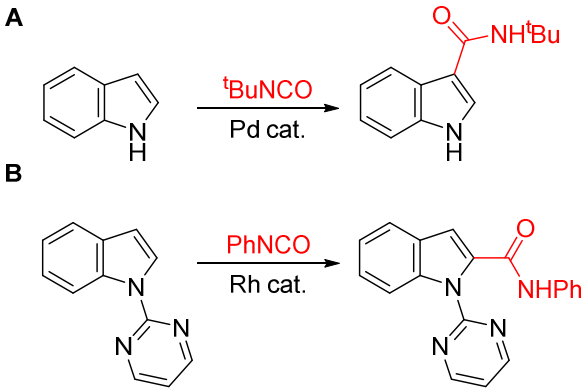
4.2.3 Integrated Amidation

As previously discussed, aryl nitriles are versatile functional groups for interconversion to other functionalities (**Scheme 51**). It was hoped that this reactivity could be exploited in the integrated cyanation reaction – allowing the scope of functional groups installed to be extended beyond nitriles. One such reaction involves the hydration of nitriles to the corresponding amides. Traditional methods for this conversion rely upon the heating of nitrile in strongly basic media, which often results in hydrolysis of the subsequent amide to acid – limiting material efficiency and necessitating by-product separation. For this reason nitrile hydratases (NHases) have been exploited for many years as a biocatalytic alternative.[239,240] NHases will be discussed in greater detail in a subsequent chapter (Chapter 5), but it should be noted at this point that they catalyse the hydration of nitriles in aqueous media, without significant heat input, whilst avoiding any hydrolysis to acid product. Their highly productive and chemoselective nature has been exploited to allow the production of bulk amide chemicals on an industrial scale.[239]



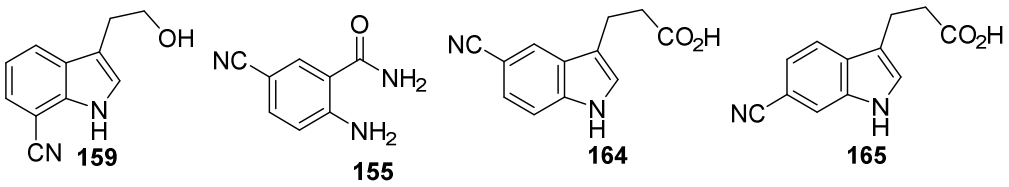
Scheme 61: Proposed C-H amidation enabled by combination of the FlHal-cyanation reaction with nitrile hydratases (NHase).

It was hoped that integration of NHase enzymes into the integrated FlHal cyanation reaction would allow the conversion of the resultant nitriles to the corresponding amides – thereby representing a formal C-H amidation reaction (**Scheme 61**). At the time of writing, the direct C-H amidation of indoles was limited to the C3 and C2 positions (**Scheme 62**).[241,242] A similar rationale to the cyanations discussed above is utilised whereby substrate control determines the position of metallation and subsequent amidation (**Figure 62**), except isocyanate equivalents are used in place of $^{-}$CN.[243] It was therefore hoped that such direct amidations could be expanded to the C7, C6 and C5 positions through integration of a NHase into the FlHal-cyanation reaction.



Scheme 62: Previous work on the C3 (**A**)[241] and C2 (**B**)[242] C-H amidation of indoles through isocyanide insertion.

To determine the feasibility of this cascade, 7-cyano tryptophol (**159**) was incubated with a Fe-centered NHase kindly provided by the Bornscheuer group.[244] At a biocatalyst loading of 0.1 mol % (typical for this type of enzyme, see Chapter 5) very little consumption of nitrile starting material **159** was detected by analytical HPLC. Significantly increasing the loading of biocatalyst afforded almost complete conversion of starting nitrile **159** to a new product, however. HRMS of this analytical scale reaction found *m/z* peaks consistent with the molecular formula of **166**, encouraging the upscale of this reaction which allowed isolation of **166** in 64 % yield. Whilst the smaller nitrile **155** was also efficiently converted by the NHase at this loading, as determined by analytical HPLC and HRMS, the cyano indole-3-propionic acid derivatives **164** and **165** did not appear to be converted at all.



Substrate	NHase Loading (mol %)	Conversion (%)*
159	0.1	8
159	0.3	33
159	0.8	61
159	1.6	98[†‡]
155	0.8	38
155	2.0	91[‡]
164	0.8	N.C.
164	2.0	N.C.
165	0.8	N.C.
165	2.0	N.C

Table 17: Hydration of integrated cyanation products by NHase. Conditions: Substrate (3.0 mM), NHase from *Rh. Equi.*, rt, overnight.*Based on consumption of starting material by analytical HPLC using the relevant calibration curves. [†]Isolated yield = 64 %. [‡]Amide detected by HRMS.

To rationalise this substrate specificity, each substrate was docked into the proposed binding site of the NHase homology model (Chapter 5) using AutoDock. Whilst poses with the C≡N of **159** coordinated to the Fe^{III} centre of the NHase active site, required for hydration, were returned (**Figure 49B**) a significant number of poses showed the OH group of **159** to be coordinated to the Fe^{III} centre (**Figure 49A**). Docking of **155** into the homology model active site in the same way solely returned poses whereby the amide oxygen atom of **155** was coordinated to the Fe^{III} centre. These results suggest that the need for high biocatalyst loadings for efficient hydration of these substrates could arise from these interactions inhibiting effective binding of CN to the Fe^{III} centre. Docking of **164** and **165** similarly only returned poses where the free carboxylic acid moiety was coordinated to Fe^{III}, rather than the nitrile, which is unsurprising given the full negative charge of this group.

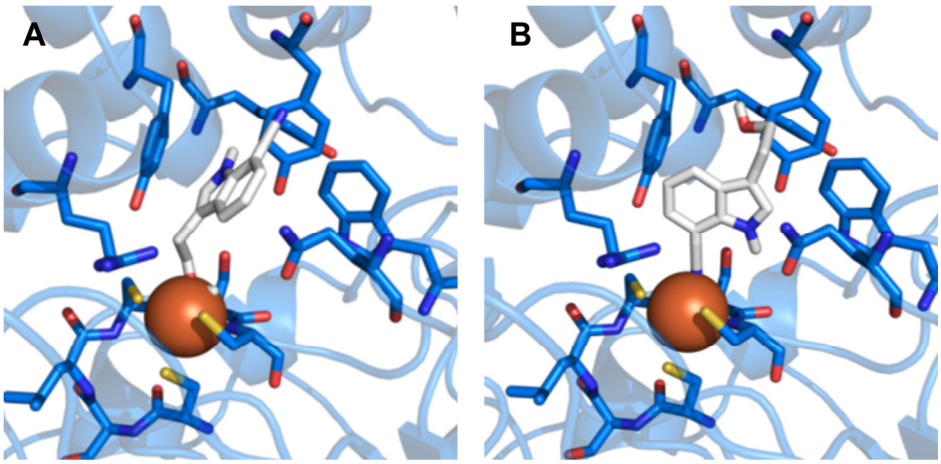
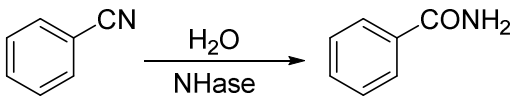


Figure 49: Poses of **159** bound to the NHase active site returned by AutoDock 4.2. (**A**) Shows the potentially inhibitory interaction of the polar OH group with Fe^{III} centre. (**B**) Shows the catalytically-active pose with the -C≡N group bound to the axial position of Fe^{III}. Fe^{III} is represented in both by an orange sphere. Details of how the homology model was generated and how the active site residues contribute to catalysis are discussed in Chapter 5.

Given that the 7-cyano indole **159** could be generated through the FIHal-cyanation cascade and hydrated by the NHase with good efficiency, an attempt at the direct C7 amidation of **104** was attempted. Since the activity this type of NHase is dependent upon catalysis at an Fe^{III} centre, it is unsurprising that their strong inhibition by $^{-}$CN has been reported.[245] Given that attempts at the FIHal-SMC cascade suggested a strong inhibition of various biocatalysts by Pd also, it seemed reasonable to assess the tolerance of the NHase to each of the cyanation reaction components. To allow this the hydration of benzonitrile (**167**) to benzamide (**168**), a well reported transformation by these enzymes, was used as a test reaction.[244]

This found NHase activity to be completely abolished in the presence of THF and the Pd catalyst **157** (**Table 18**, entries 1 and 2). Surprisingly however, NHase activity appeared tolerant to the presence of $K_4[Fe(CN)_6]$ (**Table 18**, entry 3). Given the tight coordination of $^{-}$CN to Fe, it is likely that no $^{-}$CN was liberated from $K_4[Fe(CN)_6]$ under these reaction conditions – and therefore no $^{-}$CN was available for inhibition of the NHase. It has been reported previously that Pd-catalysed cyanation reactions using $K_4[Fe(CN)_6]$ as cyanide source require heating to liberate $^{-}$CN for ligand substitution.[215] Similar tolerance assays were therefore run with pre-heating of cyanide source and palladium catalyst prior to addition of NHase to more accurately mimic the aftermath of a Pd-catalysed cyanation (**Table 18**, entries 4 to 6). In these cases an inhibitory effect of $K_4[Fe(CN)_6]$ on NHase activity was noted (**Table 18**). The inhibitory effect is likely small because of the potential for $^{-}$CN to re-combine with the Fe-centre derived from $K_4[Fe(CN)_6]$ without interaction with the Fe^{III} centre of the NHase.



167 **168**

Entry	Additive (Concentration)	Pre-incubation†	Remaining Activity (168, %)*
1	$K_4[Fe(CN)_6]$ (1.5 mM)	N	94
2	^{t}BuXPhos-Pd-G3 (0.3 mM)	N	22
3	THF (1:5)	N	N.C.
4	$K_4[Fe(CN)_6]$ (1.5 mM)	Y	36
5	^{t}BuXPhos-Pd-G3 (0.3 mM)	Y	31
6	$K_4[Fe(CN)_6]$ and ^{t}BuXPhos-Pd-G3 (1.5 mM and 0.3 mM)	Y	27

Table 18: Effect of various Pd-catalysed cyanation components upon conversion of benzonitrile (**167**) to benzamide (**168**) by the Fe-centred NHase. Conditions: PhCN (**167**, 10 mM), NHase from *Rh. Equi.* (10 μM), rt, overnight. †Pre-incubation refers to the heating of cyanation components in reaction buffer at 80 °C for 1 hour, before cooling to 20 °C and addition of PhCN (**167**) and NHase and incubation at room temperature overnight. *Based upon analytical HPLC using a calibration curve of benzamide **168**. N = No. Y = Yes.

Nonetheless, the inhibitory effect from presence of Pd catalyst **157** and $K_4[Fe(CN)_6]$ meant that some means of compartmentalising the NHase from these components must be devised. For this purpose, a PDMS thimble was chosen. Since the $[Fe(CN)_6]^{4+}$ complex is of high net charge, it should be incapable of dissolving sufficiently well in PDMS to move between the two compartments of such a reaction. Although the catalyst **157** is very hydrophobic and therefore likely to dissolve in PDMS, its limited solubility in water was considered to be sufficient to prevent flux between the two compartments – instead becoming entrapped in the PDMS membrane (**Figure 50**). Therefore if a thimble containing NHase were to be added once the cyanation reaction had completed, it was postulated that both the cyanation and hydration reactions may proceed sequentially without the need to isolate the intermediate nitrile.

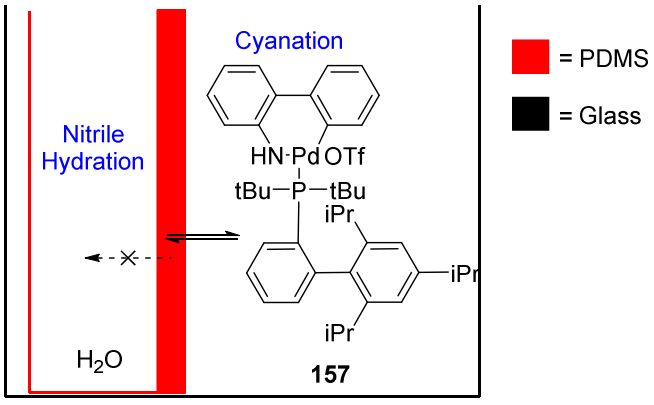
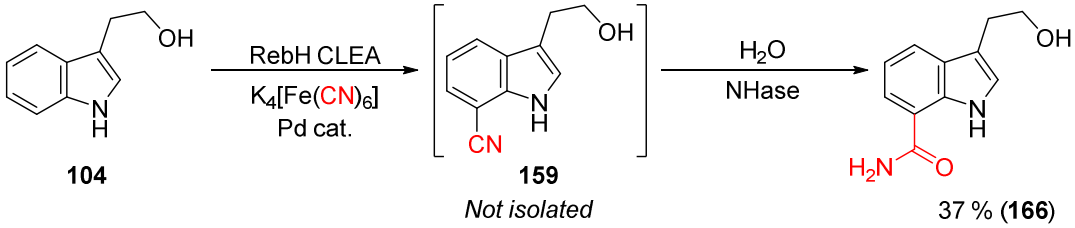


Figure 50: Schematic representation of ^{t}BuXPhos-Pd-G3 (**157**) entrapment in a PDMS membrane through partition into the hydrophobic polymer, but not out into the aqueous phase.

An initial attempt at this amidation reaction, with the NHase added within a PDMS thimble after completion of the cyanation reaction, showed no consumption of nitrile **159** or presence of **166** by analytical HPLC. The lack of NHase activity was presumed to be because of the intolerance of this NHase to the THF co-solvent essential for the cyanation reaction (**Table 15**). THF has been found to be very soluble in PDMS,[200] meaning that it would be able to efficiently flux between the two sides of the PDMS thimble since it is also water-miscible. The residual THF from integrated cyanation of **104** was therefore removed under a stream of nitrogen prior to addition of the NHase-containing thimble, allowing isolation of the 7-amide tryptophol product **166** in 37 % yield from **104**.



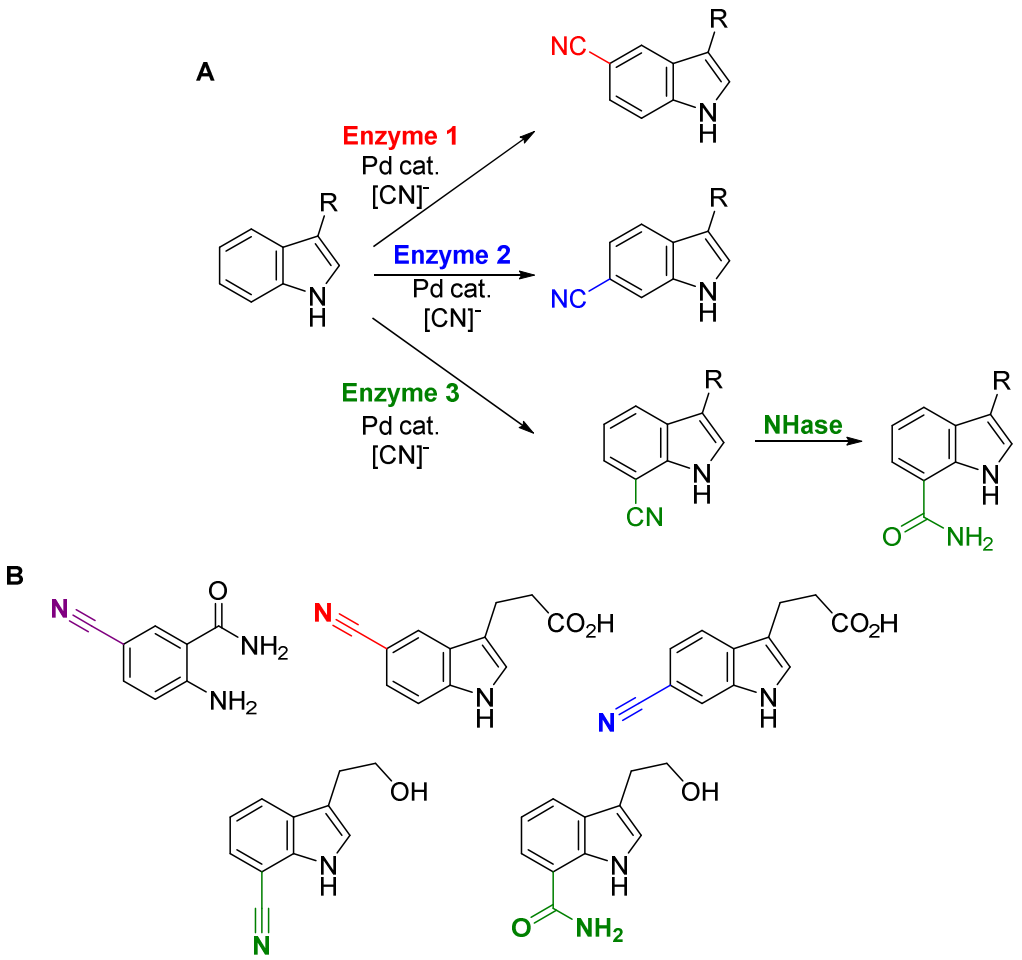
104 **159** *Not isolated* 37 % (**166**)

Scheme 63: Isolated yield from the integrated amidation of tryptophol (**104**). Following the integrated cyanation of **104** using a RebH CLEA as described above, THF was removed with nitrogen prior to addition of the NHase (2 mol %) in a PDMS thimble.

4.3 Conclusions and Outlook

Aryl nitriles (Ar-C≡N) are versatile intermediates towards a number of other useful functional groups.[214] Traditional methods for their preparation typically require the conversion of a pre-existing functional group - the installation of which often suffers from poor regioselectivity. Additionally methods for the direct cyanation of aromatic C-H bonds are currently very limited, both in terms of substrate scope and the number of regiochemistries which can be accessed. These issues are particularly pronounced with the indole heterocycle, with methods currently limited to C3 and C2 cyanation. This chapter describes the combination of flavin-dependent halogenases (FlHals), for the regioselective bromination of aromatic scaffolds, with palladium-catalysed cyanation reactions to afford a regioselective C-H cyanation overall (**Scheme 64A**).

Efficient one-pot FlHal-cyanation reactions were enabled through the use of FlHal CLEAs. As demonstrated in Chapter 3, these stabilised biocatalyst preparations allowed high concentrations of intermediate aryl bromide to be generated by affording prolonged catalyst lifetime compared to purified enzymes. The heterogeneous nature of CLEAs also facilitated their facile removal prior to addition of palladium catalysts and other cyanation reaction components - thereby overcoming the issues of incompatability between FlHals and palladium-catalysed reactions found whilst developing the FlHal-SMC cascade (Chapter 3).



Scheme 64: (**A**) General scheme of the regioselective C5, C6 and C7 cyanation reactions enabled by the combination of FIHals and Pd-catalysed cyanation chemistry into one-pot integrated reactions. The C7 cyanation was also exentended to an amidation through application of an additional NHase enzyme. (**B**) Scope of cyanated and amidated products obtained using the FIHal-cyanation reactions developed herein.

Removal of the CLEA allowed intermediate bromide to be efficiently cyanated with the palladium-catalysed conditions, without the need for isolation or purification of the intermediate aryl bromide. By avoiding such steps, the overall efficiency of the process is improved relative to a traditional telescoped process with intermediary workup procedures. By employing FIHals of different substrate scope, the cyanation of both benzamides and indoles was achieved (**Scheme 64B**). Most importantly however, this approach allowed the direct cyanation of the indole C5, C6 and C7 positions (**Scheme 64B**) – positions which are currently beyond tradition methods for indole C-H cyanation (**Figure 45**).

It was hoped that a cascade bromination-cyanation reaction could be achieved by using PDMS compartmentalisation to separate the biocatalytic and chemocatalytic components from each other based on charge. The hydrophobic nature of the ligands employed, and the need to exclude air from the cyanation reaction, meant that this was not possible however. PDMS compartmentalisation was used to allow the effective integration of a nitrile hydratase (NHase) into the FIHal-cyanation regime however, allowing the formal C-H amidation of the indole C7 position.

This combination stresses the potential value in a regioselective, direct, C-H cyanation of the indole scaffold, since the nitrile group is a valuable synthetic intermediate. The amidation method described in this chapter could not be extended to the 5- and 6-amidation however, since the cyanated intermediates **164** and **165** did not appear to be substrates for the NHase employed. The charged nature of these intermediates would also likely preclude their efficient flux through the PDMS membrane. Although integration of this particular enzyme into the cyanation cascade was not entirely successful, there is potential for a manifold of other reactions given the versatile reactivity of the nitrile group.

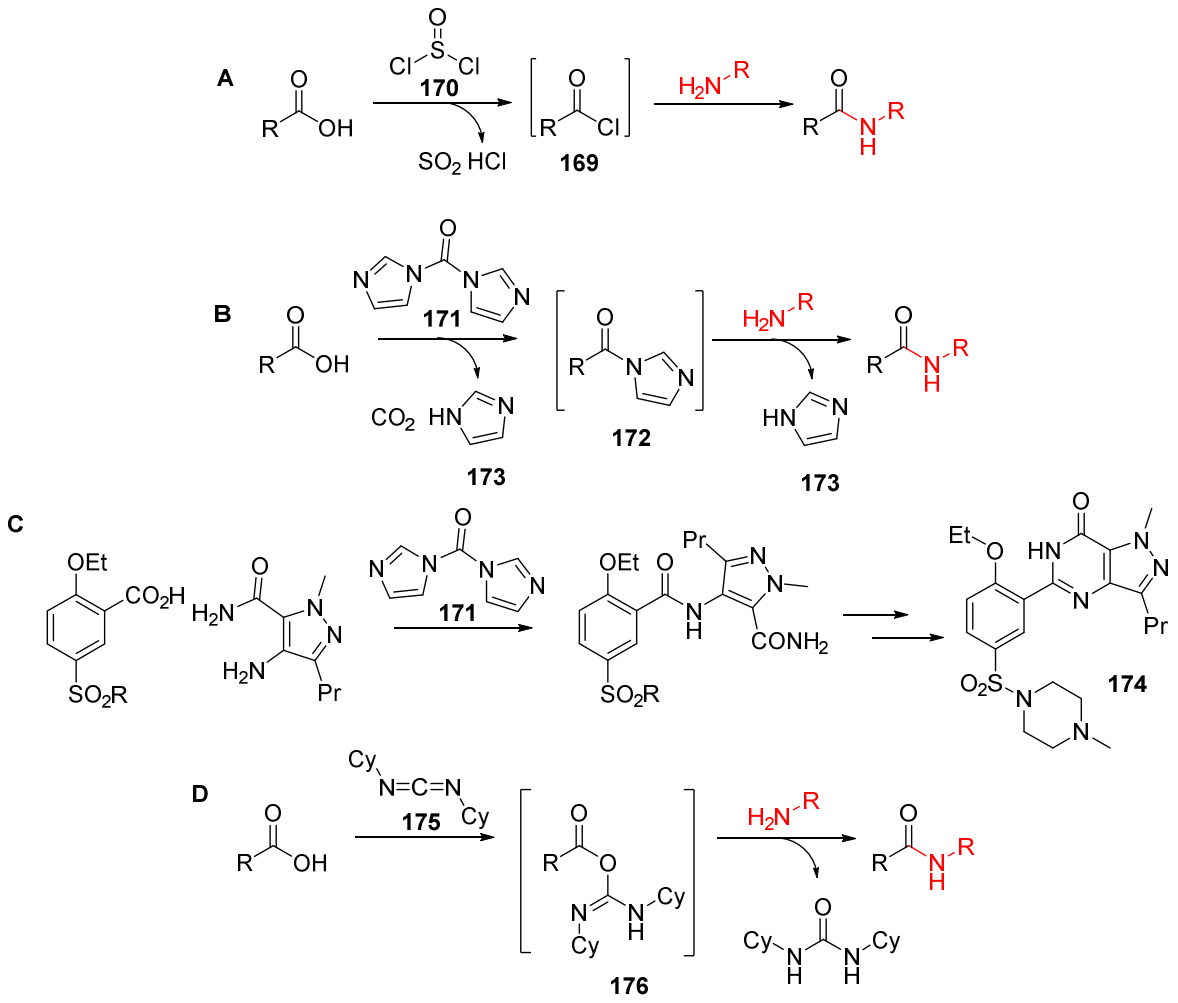
5. Nitrile Hydration and C-N Bond Formation Cascade

5.1 Introduction

The development of methods for the synthesis of amides which do not require poor atom economy reagents has been a long-standing challenge in synthetic organic chemistry.[246-248] The need for such methods has been emphasised by industrialists a number of times,[246,247] due to the prevalence of amide bond-formation steps in the synthesis of drug molecules and polymers.[249,250] Amide bonds are prevalent in compounds of pharmaceutical importance partially because of the duality of the amide bond – acting as both a hydrogen bond donor and acceptor. The ability of the amide to form interactions in this way is key to stabilisation of protein secondary structures, by allowing strong hydrogen bonding interactions between the backbone of residues in α-helices and β-sheets. Such interactions also contribute to the interaction of small molecule and peptide drugs with their protein targets. Additionally the delocalisation of electrons around amide bonds confers increased stability and rigidity to the C-N linkage.

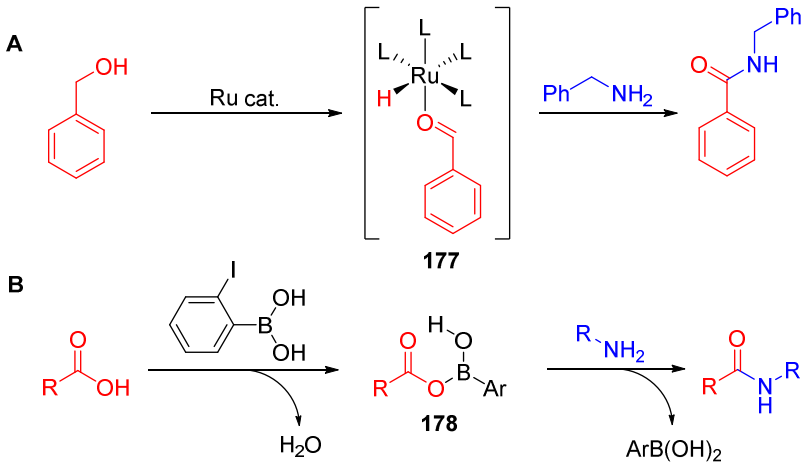
A typical strategy for the preparation of amides involves the activation of carboxylic acids, followed by treatment with a nucleophilic amine.[248] Delocalisation of electrons around the carboxylic acid group leads to reduced electrophilicity of the carbonyl C, therefore direct treatment with a basic amine would lead to salt formation rather than amide bond formation. Prior activation of the carboxylic acid can create a carbonyl C of increased polarity with a better leaving group however, facilitating attack of the nucleophilic amine to form amide. A classic activation approach involves the conversion of carboxylic acids to acyl chlorides (**169**), which possess significant polarisation of the carbonyl C, using thionyl chloride (**170**) - thereby facilitating nucleophilic attack of amine with loss of chloride (**Scheme 65A**). This reaction is widely used for the synthesis of amides because the by-products (SO_2 and HCl) are both gaseous, allowing their facile removal through basic scrubbing.[246]

Alternative activation methods involve the formation of reactive esters with facile leaving groups.[248] Treatment of carboxylic acid with 1,1-carbonyldiimidazole (**171**) for example leads to the generation of an imidazole ester (**172**), which can then undergo facile attack with an incipient nucelophile because of the stability of the imidazole (**173**) leaving group (**Scheme 65B**). Whilst **171** is preferable over $SOCl_2$ for these reactions, due to the toxicity of $SOCl_2$ and its by products relative to **171**, activation of the acid generates an equivalent of imidazole and CO_2. Reaction of the activated ester then generates a second equivalent of imidazole. These by-products reduce the atom economy of the reaction significantly, although the favourable handling properties of **171** mean that is has been utilised in the synthesis of Sildenafil (**174**) in preference over $SOCl_2$ (**Scheme 65C**).[251] A number of other reagents which act in a manner analogous to **171** are known, for example dicyclohexylcarbodiimide (DDC, **175**) which generates a reactive acylurea intermediate (**176**, **Scheme 65D**).



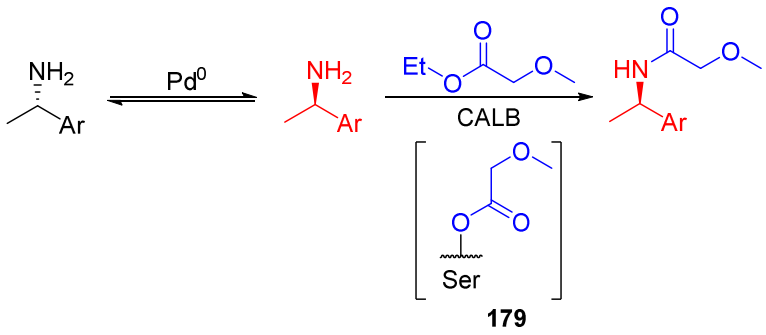
Scheme 65: Methods of amide synthesis involving activation of a carboxylic acid using; (**A**) thionyl chloride (**170**) to generate an acyl chloride intermediate (**169**) and (**B**) 1,1'-carbonyldiimidazole (**171**, CDI) to generate an imidazole ester intermediate (**172**); (**C**) Synthesis of Sildenafil (**174**) using **171** for the key amide bond formation step.[251] (**D**) Use of *N*,*N*'-dicyclohexylcarbodiimide (**175**, DCC) for activation of a carboxylic acid to generate an acyl urea intermediate (**176**).

To address the issues of atom economy with these activation approaches, there has been significant interest in catalytic methods for the preparation of amides – driven by the desire for industrial processes which do not generate stoichiometric waste.[246-248] One such example uses a ruthenium catalyst to directly couple alcohols and amines to give secondary amides (**Scheme 66A**).[252] The ruthenium catalyst is proposed to generate an activated ruthenium-aldehyde complex (**177**) from the alcohol, which is then susceptible to attack by the amine – forming secondary amide. This method, along with similar Ru-catalysed methods, offer the advantage of directly coupling the two partners with catalytic activation, meaning that H_2 is the only stoichiometric by-product.[248,252] Catalytic methods for the activation of carboxylic acids are also known. One such approach involves the use of boronic acids as organocatalysts to generate acyl boronate intermediates (**178**), which are then susceptible to attack by a nucleophile (**Scheme 66B**).[253] Since this attack leads to regeneration of the boronic acid, the only by-products from this reaction are H_2O and catalyst.



Scheme 66: Catalytic methods for amide synthesis using (**A**) ruthenium catalysis to activate alcohols[252] and (**B**) organocatalysis to generate reactive boronate esters.[253]

Other catalytic strategies for the conversion of carboxylic acids to amides have involved the use of biocatalysis.[254] The most common strategy involves using a lipase enzyme, which generates an acyl-enzyme adduct (**179**) with starting carboxylic acid. This adduct is then susceptible to attack by a nucleophilic amine – generating the secondary amide product (**Scheme 67**). The enantiospecificity of these enzymes allows the establishment of dynamic kinetic resolution processes, when integrated with metal catalysts for racemisation of the amine starting material (**Scheme 67**).[255,256] Biocatalysis has also been employed for the hydration of nitriles to primary amides (5.1.1).[239]

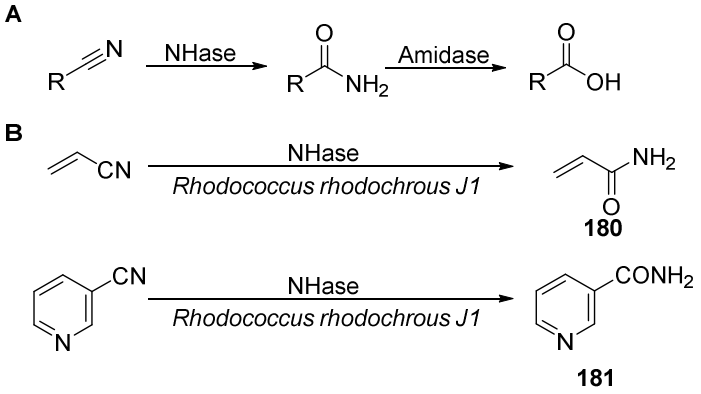


Scheme 67: Dynamic kinetic resolution of racemic amines to chiral amides through use of a Pd-catalysed racemisation and lipase-catalysed amine acylation.[256] CALB = *Candida antartica* lipase B.

5.1.1 Nitrile Hydratases (NHases)

Organisms which possess nitrile hydratases (NHases) have been used for the production of amides on an industrial scale for decades.[239,240] These enzymes, found in a range of bacteria and fungi, are responsible for the hydration of nitriles to amides which are often subsequently hydrolysed to acids by an associated amidase enzyme (**Scheme 68A**).[239] Through the correct use of inducers however, it was found that the amide hydrolysis reaction could be suppressed in these organisms – affording a biocatalyst capable of the selective hydration of nitriles to amides without formation of acid by-product.[257,258] These biocatalysts were subsequently applied to the production

of acrylamide (**180**, a common feedstock for the polymer industry) and nicotinamide (**181**, a food additive) – both produced in kiloton quantities using these methods annually (**Scheme 68B**).[239] The previous methods for production of these compounds involved the heating of nitrile with base, leading to formation of acid by-product, whilst the NHase biocatalyst allowed high purity amide products to be obtained by avoiding amide hydrolysis.



Scheme 68: (**A**) general activity of nitrile hydratase (NHase) and amidase enzymes. (**B**) Industrial routes to acrylamide (**180**) and nicotinamide (**181**) using NHase-expressing organisms.[239]

Despite their use being reported for many years, the mechanism of hydration by NHases remained elusive until recently. Mechanistic and crystallographic studies were initially hampered by difficulty in the recombinant expression of the NHases in an active form.[239,259] Crystallography of the NHases eventually confirmed that these enzymes exist as a heteromer (**Figure 51A**) – with the α-subunit thought to be responsible for catalysis whilst the β-subunit is involved in substrate binding.[260,261] Two distinct classes of NHase were also identified – those which bind Fe^{III} and those which bind Co^{II}.[245,257,262] In both cases, the metal ion was found to be bound to a cysteine residue, as well as two further post-translationally modified cysteine residues (**Figure 51B**) – a cysteine-sulfinic acid (Cys-SO_2H, **182**) and cysteine-sulfenic acid (Cys-SOH, **183**).[260,261]

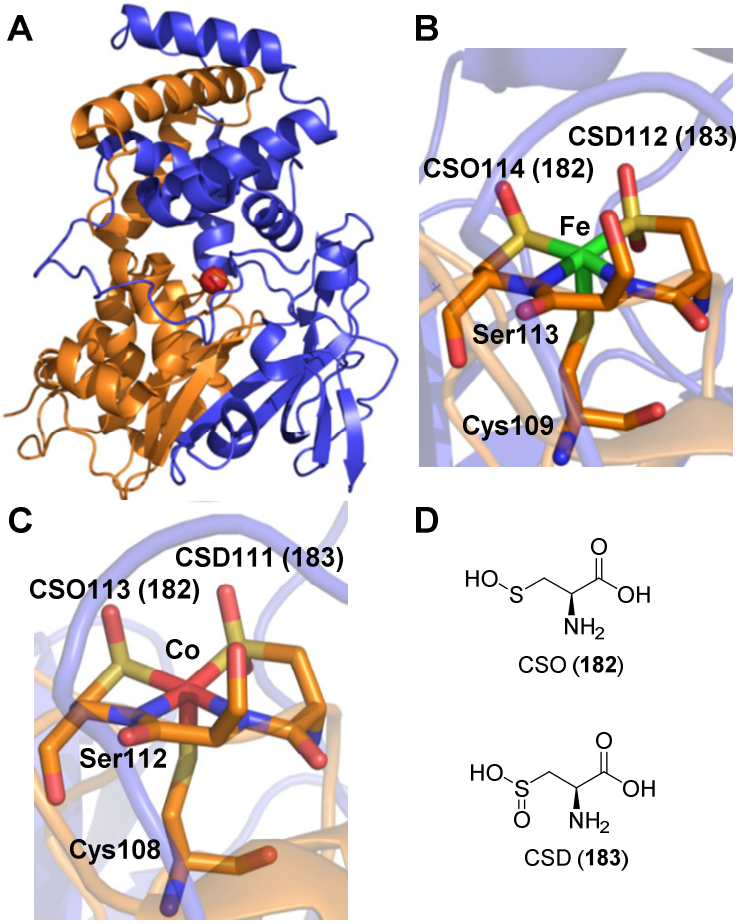


Figure 51: (**A**) Secondary structure of the Co-dependent NHase from *Pseudonocardia thermophilia* with the α-subunit shown in orange, the β subunit shown in blue and Co^{II} shown as a red sphere (PDB 4OB0). (**B**) Crystal structure showing the coordination geometry around Fe^{III} in the Fe-centred NHase from *Rhodococcus erythropolis* (PDB 3X26). (**C**) Crystal structure showing the coordination geometry around Co^{II} in the Co-centred NHase from *Pseudonocardia thermophilia* (PDB 1IRE). (**D**) Structure of the post-translationally modified cysteine residues cysteine-sulfenic acid (CSO, **182**) and cysteine-sulfinic acid (CSD, **183**).

This non-canonical binding of Fe^{III} and Co^{II} sparked further intrigue about the mechanism of nitrile hydration by these enzymes. Crystallisation of a Co^{II}-centred NHase with phenyl boronic acid found the B of $PhB(OH)_2$ to be covalently bound to the cystine-sulfenic acid residue (**Figure 52A**), suggesting that $CysSO^-$ may act as a nucleophilie to attack the nitrile C and form a covalent intermediate.[263] Time-resolved X-ray crystallography subsequently provided evidence for such a cyclic intermediate (**184**) after binding of nitrile substrate to the Fe^{III} centre of a NHase (**Figure 52B**).[264] FTIR (Fourier-transform infrared) spectroscopy was also used to determine the position of $H_2{}^{18}O$ attack during this reaction, and found the $CysSO^-$ to be attacked - evidenced by a downshift of the $S\text{-}O^-$ vibration of 24 cm^{-1}.[264] Given these data, the mechanism shown in **Figure 52C** was proposed.

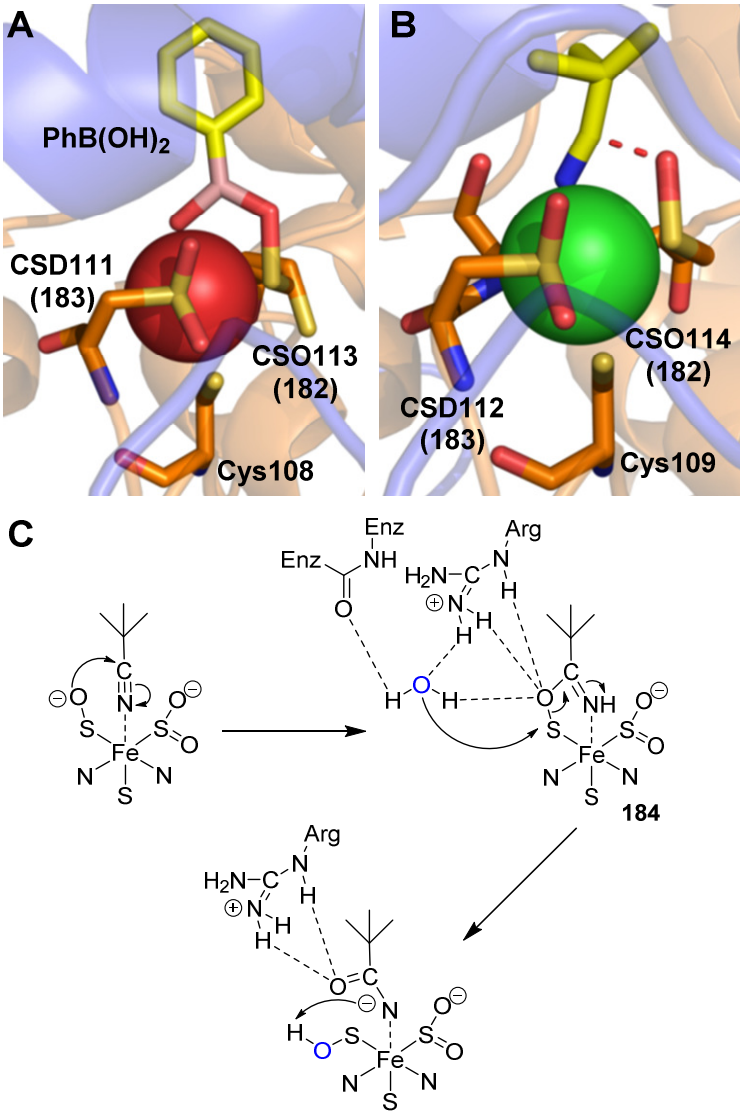
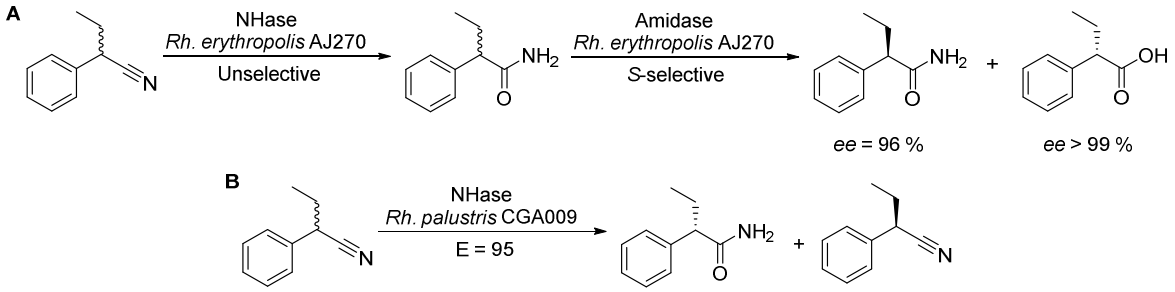


Figure 52: (**A**) Crystal structure of the NHase from *Pseudonocardia thermophilia* co-crystallised with phenyl boronic acid ($PhB(OH)_2$) showing the covalent adduct between B and CSO (PDB 4OB0).[263] (**B**) Bent nitrile intermediate observed by time-resolved X-ray crystallography during the reaction course of trimethylacetonitrile with the NHase from *Rhodococcus erythropolis* (PDB 3X26).[264] The proposed covalent bond between nitrile C and CSO is shown as a red dotted line. (**C**) Proposed mechanism of nitrile hydration by NHases.

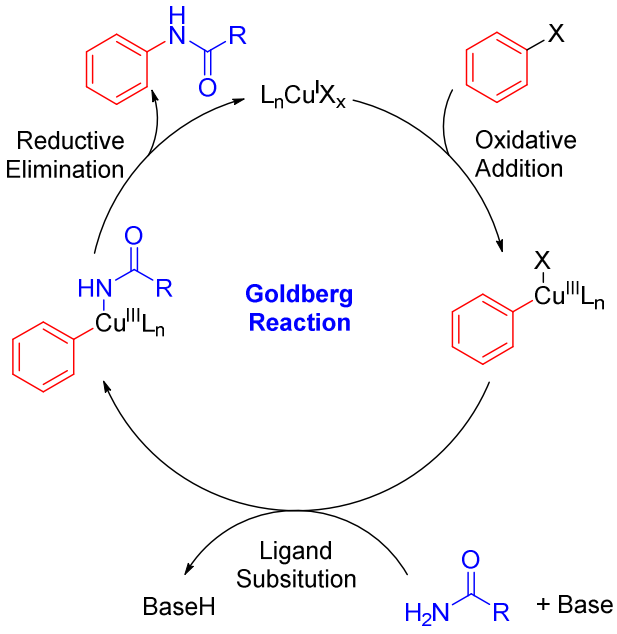
Since their initial discovery and characterisation, there has been significant work focussed on the application of these enzymes as biocatalysts despite a lack of mechanistic information at the time. In general, the Fe^{III}-centred NHases hydrate aliphatic substrates more readily than the Co^{II}-centred NHases which prefer aromatic nitriles.[244,265] Typically the Fe-centred NHases display poor stereospecificity (E<20),[265,266] although some associated amidases from the same species have been demonstrated to preferentially hydrolyse the *S*-amide.[267,268] The wild-type organisms, expressing both NHase and the associated amidase, have therefore been used together for the preparation of *R*-amides and *S*-acids from a single racemic nitrile (**Scheme 69A**).[269,270] Conversely, a number of Co-centred NHases have been shown to selectively hydrolyse the *S*-enantiomer of nitriles with bulky, non-polar functionality α to the CN group (**Scheme 69B**).[265]



Scheme 69: (**A**) Whole cell hydration and hydrolysis of a substituted 2-propyl nitrile using an unselective NHase and *S*-selective amidase to afford amide and acid with high enantiomeric excess (*ee*).[268] The E factor of nitrile hydration using the purified NHase from this organism was subsequently found to be 7,[265] suggesting that enantiopurity of the acid and amide products results from the *S*-selective amidase. (**B**) Enantioselective hydration of a racemic nitrile using a NHase.[265]

5.1.2 Copper-catalysed C-N Bond Formations

Copper catalysis has been explored for the synthesis of secondary amides directly from primary amides, thereby avoiding the need for activation of carboxylic acid precursors. The earliest examples of this strategy are now termed the "Goldberg" reaction and involve the Cu^{I}-catalysed coupling of *N*-nucleophiles (including amides) with aryl halides (**Scheme 70**).[271] This reaction is thought to proceed via the oxidative addition of Cu^{I} across the C-halide bond to generate a Cu^{III} intermediate (**Scheme 70**). Ligand substitution of amide onto this complex then generates a Cu^{III} complex bearing both organic fragments. Reductive elimination of the two fragments from this intermediate then forms the C-N bond and re-generates Cu^{I} catalyst.[272]

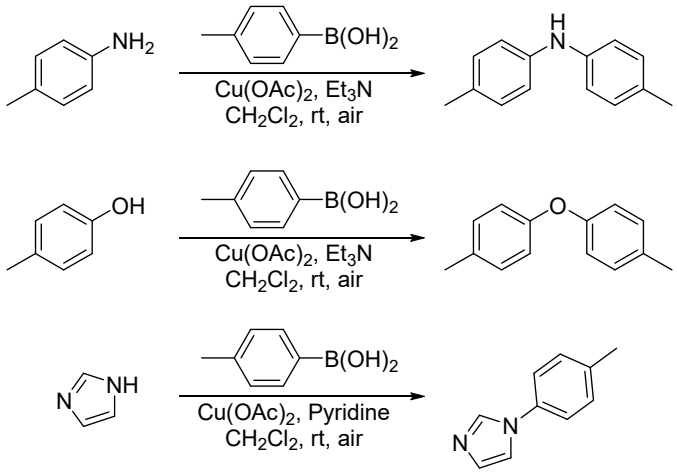


Scheme 70: Proposed mechanism for the Goldberg coupling of amides and aryl halides.

Because of the requirement for oxidative addition, Goldberg reactions typically require very high temperatures for good yield of secondary amide product. Since persistence of Cu^{III} intermediates is required, an inert atmosphere is also typically necessary. The poor nucleophilicity of the amide N

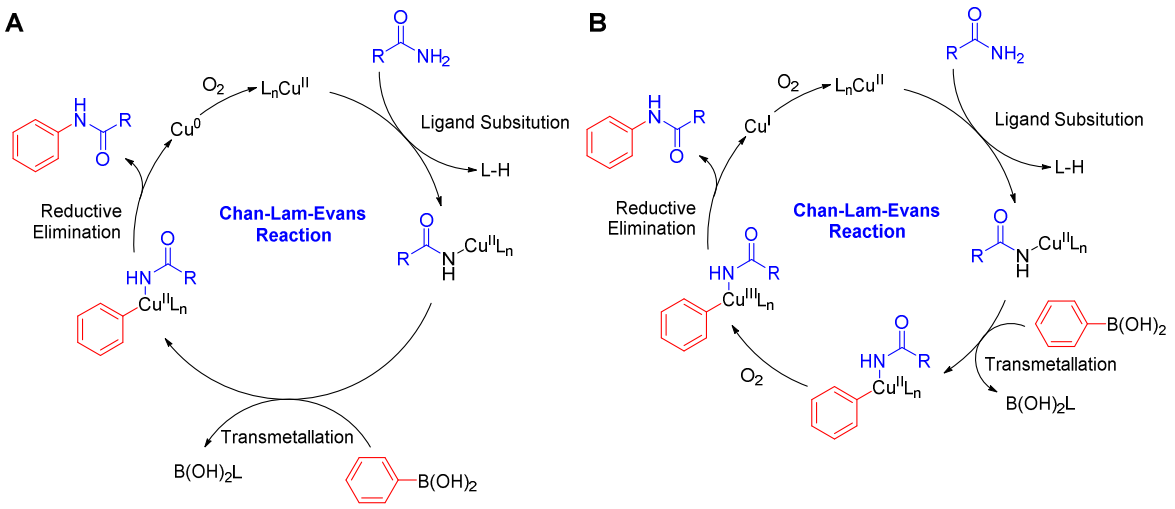
also often necessitates a strong inorganic base to promote ligand substitution. Despite these drawbacks, this method received much attention because of the relatively inexpensive metal catalyst and ability to synthesise enamides using vinyl halides – which would be difficult using classical activation strategies because of the poor nucleophilicity of enamines.[271]

A related process was subsequently developed, whereby Cu^{II} promotes the coupling of amines, imidazoles and phenols with boronic acids – termed the Chan-Lam-Evans (CLE) reaction (**Scheme 71**).[273-275] Since initial reports which required the use of stoichiometric quantities of Cu^{II}, protocols involving catalytic Cu^{II} have been reported.[276] The scope of the reaction has also been improved significantly to include a range of nucleophilic coupling partners, including amides.[277,278] Unlike the Goldberg reaction, CLE reactions typically afford high yields at or near ambient temperature (ca. 40 °C), with weak organic bases (e.g. Et_3N) under open air rather than inert atmosphere. Given the preferable operating conditions, the CLE reaction has attracted much attention as an alternative to the Goldberg reaction for the synthesis of secondary amides.



Scheme 71: Seminal stoichiometric Cu^{II}-promoted couplings of nucleophilic N and O groups with aromatic boronic acids.[273-275]

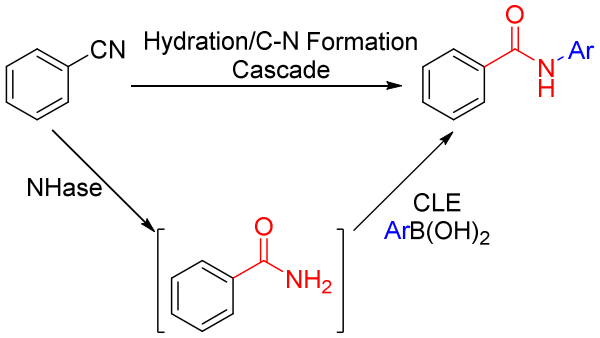
The mechanism of the CLE reaction is still disputed, however a number of working mechanisms have been proposed (**Scheme 72**).[276] It is believed that amide first coordinates to the Cu^{II} species through ligand substitution, followed by transmetallation of boronic acid with this intermediate. It is then thought that O_2 may cause oxidation of this Cu^{II} intermediate to Cu^{III}, which would then undergo facile reductive elimination to afford secondary amide product and Cu^{I} (**Scheme 72B**). If Cu^{I} is generated, then O_2 may be responsible for oxidation back to the catalytically active Cu^{II}. Alternatively, it is also proposed that reductive elimination may occur directly from the Cu^{II} intermediate, affording Cu^{0} which is then oxidised back to Cu^{II} by O_2 (**Scheme 72A**).



Scheme 72: Two proposed catalytic cycles of the Chan-Lam-Evans reaction proceeding via (**A**) Cu^{II} and Cu^{0} or (**B**) Cu^{II}, Cu^{III}, Cu^{I}.

5.1.3 Conclusions

Amide bonds are prevalent moieties in compounds of importance in many areas of the chemical industry.[248] As such, green methods for the synthesis of amides has been an active interest in the field of synthetic organic chemistry of late.[247] The NHase catalysed hydration of nitriles to amides has been exploited on an industrial scale for many years for the synthesis of important primary amides, because these enzymes allow selective hydration to amide whilst avoiding hydrolysis to acid by-product.[239] The scope of NHases is limited to the synthesis of primary amides, however. It was therefore envisioned that the combination of this mild hydration process with Cu-catalysed CLE reactions, for the arylation and vinylation of the resulting amides, in single pot integrated reactions would expand the scope of NHases to secondary aryl- and vinyl-amides (**Scheme 73**).

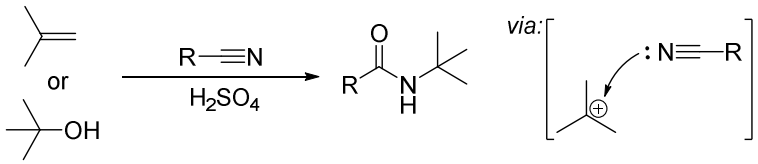


Scheme 73: Proposed hydration-arylation cascade through combination of NHase enzymes and Chan-Lam-Evans (CLE) coupling for the synthesis of secondary amides from nitriles.

Biocatalytic methods for the synthesis of secondary amides by acylation of an amine are well established, with dynamic kinetic resolution processes possible through integration of metal catalysts for racemisation of the amine donor (**Scheme 67**).[256] Given the poor nucleophilicity of enamines however, such processes are unable to synthesise enamides.[256] CLE reactions are capable of enamide synthesis however, since the vinyl donor is a vinyl boron species. It was therefore hoped that the integrated NHase-CLE approach (**Scheme 73**) would allow the synthesis

of enamides, beyond the scope of current biocatalytic methods, directly from nitriles without the need for isolation or purification of the intermediate amide.

The direct conversion of nitriles to secondary amides is analogous to the Ritter reaction (**Scheme 74**).[60] A typical Ritter reaction involves the reaction of a nitrile with a carbocation equivalent to form secondary amide. The requirement for such carbocation equivalents can significantly reduce the scope of this process however, as only compounds capable of generating such a centre are substrates. Typically the carbocation is generated by treatment of a tertiary alcohol or alkene with a strong acid, which limits the functional group tolerance of the Ritter methodology. Since the CLE reaction does not rely upon such equivalents, it was hoped that the substrate scope and functional group tolerance of the integrated NHase-CLE reaction may be broader than the Ritter methodology.



Scheme 74: Synthesis of secondary amides via the Ritter reaction of nitriles with alkenes or tertiary alcohols, showing the requirement for generation of a stable carbocation.

This chapter will describe the attempts made to combine the NHase-catalysed nitrile hydration with Cu-catalysed CLE reactions. A number of additional Cu-catalysed C-N bond formation reactions are also probed for their utility in arylation and vinylation of the intermediate amides, as well as Pd-catalysed methods.

5.2 Results and Discussion

5.2.1 Nitrile Hydratase Biocatalysis

Since use of the wild-type NHase-expressing organisms for nitrile hydration requires addition of extraneous inducers, typically nitriles or amides themselves,[240] a NHase which could be recombinantly expressed in *E. coli* was desired. A previously-reported construct containing the α- and β-subunits of a Fe^{III}-centred NHase from *Rhodococcus equi* TG328-2 (called Equi_NHase hereafter) was therefore obtained from the group of Uwe Bornscheuer.[244] It has been reported by many groups that expression of functional NHases requires an additional "activator" protein, the role of which is still not fully understood;[239] therefore an additional construct was also kindly provided encoding the proposed NHase activator from this organism. Co-transformation of *E. coli* BL21 with both constructs allowed expression of Equi_NHase, which could subsequently be purified by IMAC. Presence of the α- and β-subunits is evidenced by the presence of two bands at approximately 23 and 24 kDa respectively in SDS-PAGE analysis of the IMAC eluent (**Figure 53**).

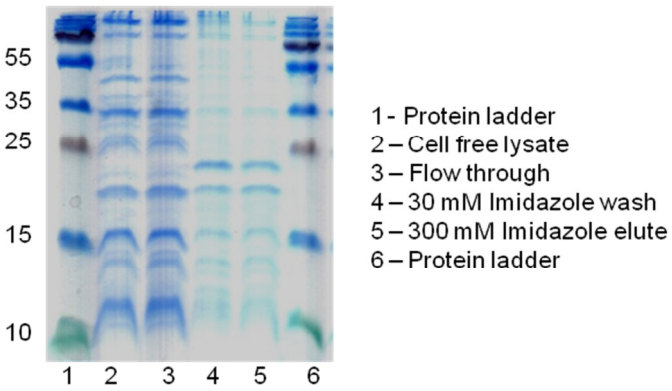


Figure 53: SDS-PAGE analysis of Equi_NHase purification from *E. coli* BL21 by IMAC. Protein mwt: 23.3 kDa (αNHase); 24.1 kDa (βNHase). The ladder used is PageRuler^{TM} Prestained Protein Ladder from ThermoScientific^{TM}.

Activity of this NHase was subsequently verified using the hydration of benzonitrile (**167**) to benzamide (**168**) as test reaction (**Figure 54A**). A new peak was detected by analytical HPLC which was not present in controls containing either no enzyme (**Figure 54B**) or enzyme which had been heat deactivated prior to addition of substrate (**Figure 54C**). This new peak overlaid with a standard of **168** and LC-MS subsequently confirmed the presence of amide through the detection of *m/z* peaks consistent with the formula of **168**.

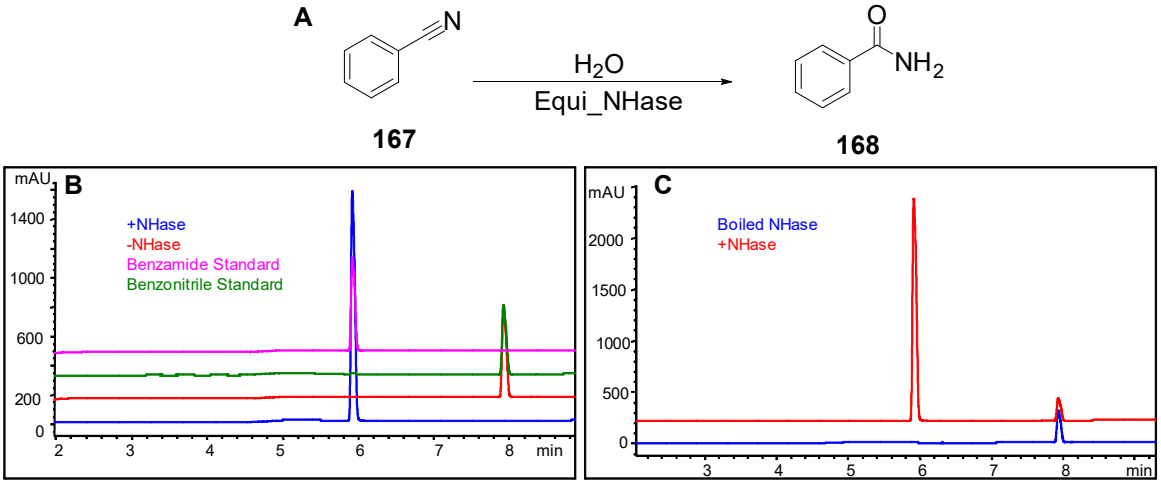


Figure 54: (**A**) Test reaction used to determine activity of purified Equi_NHase. (**B**) HPLC UV chromatogram monitored at 223 nm resulting from the test reaction in **A** with and without Equi_NHase added, overlaid with standards of **167** and **168**. (**C**) HPLC UV chromatogram monitored at 223 nm resulting from the test reaction in **A** using heat denatured NHase. Conditions: substrate (10 mM), Equi_NHase (8 μM), KPi (pH 8.0), 30 ^{o}C , 10 min.

5.2.1.1 NHase Substrate Scope

With the knowledge that an active NHase enzyme had been prepared, the scope of nitriles which could be hydrated by this enzyme was subsequently determined to assess the potential substrate scope of the integrated NHase-CLE process. The group of Bornscheuer had previously reported a small panel of standard NHase substrates to be efficiently hydrated by this enzyme - with a

preference for small aliphatic nitriles.[244] A panel of additional substrates was subsequently selected to determine the breath of this substrate scope (**Figure 55**).

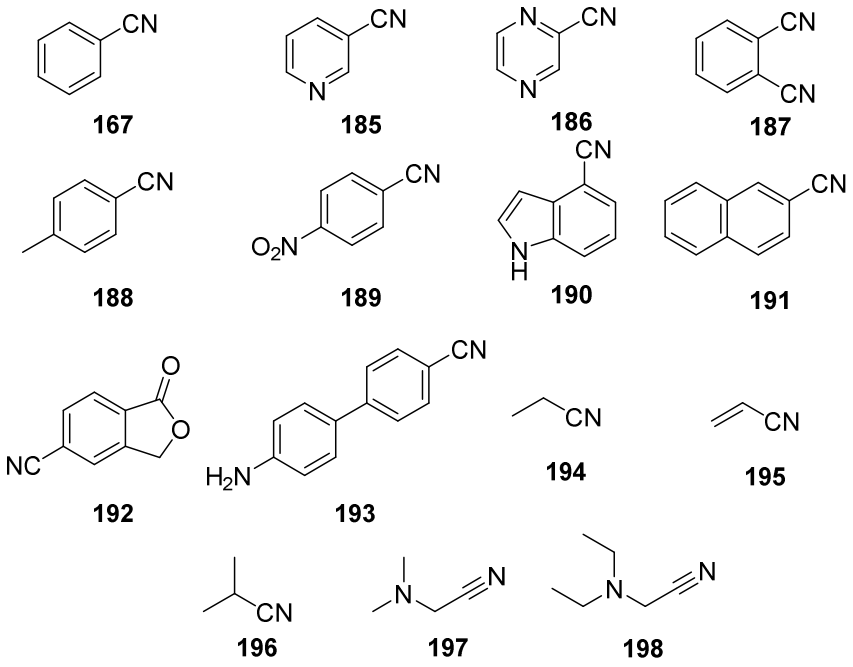
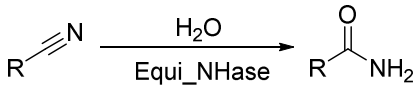


Figure 55: Nitrile substrates tested with Equi_NHase for hydration activity.

Each substrate was assayed with the purified NHase on an analytical scale, with associated controls containing no NHase enzyme. Analytical HPLC and GC-MS were used to identify the presence and abundance of any peaks unique to the reactions containing NHase. Conversions were then calculated based upon calibration curves of either the substrate nitrile or amide product depending upon commercial availability.

In addition to benzonitrile (**167**) and 2-cyanopyridine (**185**) which had been shown to be good substrates for this NHase,[244] the aromatic nitriles **186** to **191** also appeared to be converted under these conditions (**Table 19**). The pyrazine nitrile (**186**) was less efficiently converted than the analogous pyridine nitrile (**185**), suggesting that the additional coordinating group may lead to non-ideal binding of substrate to the enzyme. The electron-rich 4-methylbenzonitirle (**188**) appeared to be converted more efficiently than the electron-deficient 4-nitrobenzonitrile (**189**). This may be because the electron-donating effect of the 4-methyl group in **188** leads to an increase in the electron density of the nitrile N, promoting binding to the active site Fe of the NHase, whilst the electron-withdrawing nature of the 4-NO_2 group in **189** retards this binding. Alternatively this could be related to interactions between the 4-subsituents of these substrates with other substrate-binding residues. The bicyclic substrates **190** and **191** were hydrated significantly less efficiently by this NHase, suggesting that their steric bulk may limit efficient binding to the NHase active site or that the limited water solubility of these nitriles reduces the amount of substrate available for hydration. Interestingly the phthalide nitrile (**192**) did not appear to be converted by this NHase at all despite being a similar steric size to the indolic (**190**) and napthyl (**191**) substrates, suggesting that the presence of an additional coordinating group into substrate may impact binding to the

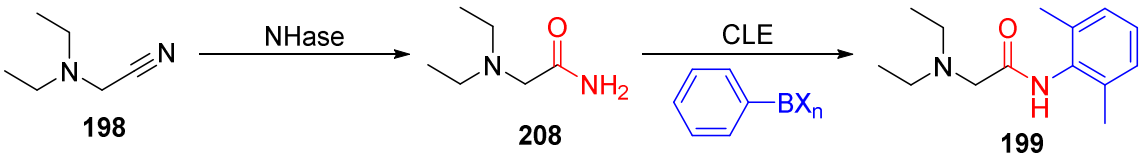
NHase active site. It is also possible that the aniline moiety of **193** has a similar effect, or that the increased steric bulk of **193** resulting from the biaryl linkage may prevent efficient binding.



Substrate	Conversion (%)	Method
167	93 (± 1.5)	A
185	89 (± 3.1)	A
186	23 (± 2.1)	A
187	9 (± 3.0)	B
188	94 (± 3.2)	A
189	64 (± 5.5)	A
190	15 (± 2.0)	B
191	11 (± 1.5)	B
192	N.C.	B
193	N.C.	B
194	>99 (± 1.7)	C
195	98 (± 1.2)	C
196	>99 (± 0.5)	C
197	92 (± 1.5)	C
198	87 (± 5.9)	C

Table 19: Percentage conversion of the substrates in **Figure 55** to amide product by Equi_NHase. Conditions: substrate (10 mM), Equi_NHase (10 μM), KPi (pH 8.0), 30 ^{o}C, 1 hr. Average of three runs. Method A: Conversion determined by analytical HPLC using a calibration curve of amide product. Method B: Conversion determined by analytical HPLC using a calibration curve of nitrile substrate. Method C: Conversion determined by GC-MS based on remaining starting material compared to butanoic acid internal standard. Average of three runs. N.C. = No conversion.

Unsurprisingly the small aliphatic substrates **194** to **198** appeared to be hydrated more efficiently by the NHase than the larger aromatic substrates. **197** was assayed for hydration by the NHase because of its similarity to the intermediate **198** in the proposed integrated NHase-CLE synthesis of lidocaine (**199**, **Scheme 75**) – a local anaesthetic highlighted in the World Health Organisation list of essential medicines. The hydrogen bond accepting group of this substrate not appear to impact upon the efficiency of **197** hydration, and therefore the lidocaine precursor **198** was synthesised. GC-MS analysis of analytical scale NHase assays showed little deterioration of hydration efficiency of the larger substrate **198** by Equi_NHase, adding promise to the proposed integrated NHase-CLE synthesis (**Scheme 75**).



Scheme 75: Proposed synthesis of the local anaesthetic lidocaine (**199**) using the integrated NHase-CLE reaction.

These NHase-catalysed hydrations were subsequently upscaled to allow isolation and characterisation of the proposed amide products (**Figure 56**). In each case the amide product was isolated in good yield, and confirmed to be a primary amide by ^{1}H and ^{13}C NMR in addition to HRMS and IR spectroscopy. Of particular interest was the product from hydration of the di-nitrile **187**, since it has previously been reported that NHases are capable of the selective mono-hydration of di-nitriles.[239] In this case, the product was indeed identified as the mono-amide product **201** using the methods previously described. ^{13}C NMR and IR spectroscopy are of particular value in this case, since the $CONH_2$ and C≡N groups give distinctive peaks in both spectra.

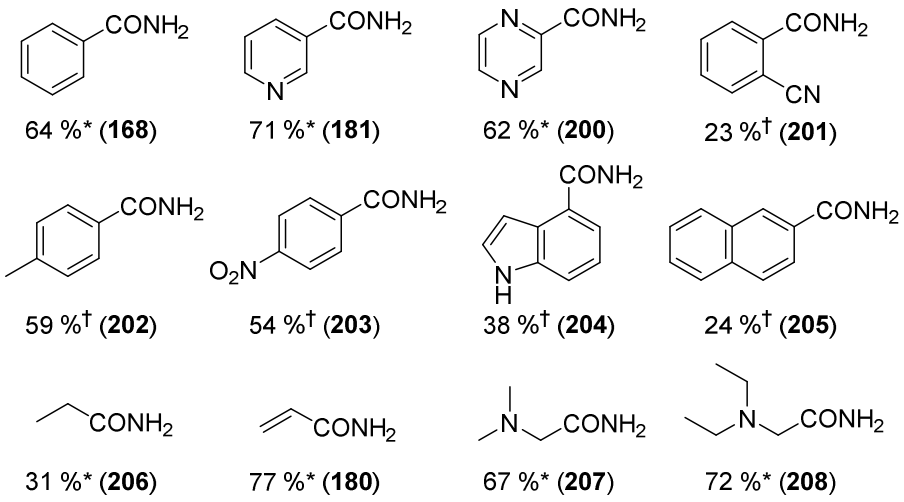


Figure 56: Isolated yield of amide products obtained from the hydration of the nitilres in **Figure 54** by Equi_NHase. Conditions: substrate (25 or 50 mM), Equi_NHase (8 μM), rt, overnight. *[Substrate] = 50 mM. †[Substrate] = 25 mM.

5.2.2.2 NHase Homology Model

In order to rationalise the observed substrate preference, and potentially allow engineering of Equi_NHase to accept further substrates of interest, structural information about this enzyme was desired. Lacking a crystal structure of this particular enzyme, a homology model was constructed using the resolved crystal structure of a related Fe-centred NHase. A database search of resolved crystal structures using ExPASy revealed the α-subunit of Equi_NHase to share very high sequence identity (83 %) with the α-subunit of the Fe-centred NHase from *Rhodococcus erythropolis* AJ270. Similarly the β-subunit of Equi_NHase was found to share 70 % sequence identity with the β-subunit of the Fe-centred NHase from *Rhodococcus erythropolis* N-771. The resolved crystal structures of these two NHases (PDB 2QDY and 3WVD respectively) were subsequently used to prepare a homology model of Equi_NHase using ExPASy. The high

sequence identity of the target and template enzymes allowed a homology model with 99 and 97 % coverage of the α- and β-subunits of Equi_NHase to be prepared. Alignment of these two models to the overall secondary structure of the *Rhodococcus erythropolis* N-771 NHase crystal structure afforded a workable approximation of the Equi_NHase structure (**Figure 57A**).

Analysis of the Equi_NHase homology model revealed a loop similar to the Fe-binding region of the template α-unit – with three cysteine residues capable of coordination to the Fe-centre (**Figures 57B** and **57C**). The two cysteine residues proposed to coordinate at the equatorial positions of Fe (αC116 and αC118) are likely to be post-translationally oxidised to the cysteine-sulfenic acid (**182**) and cysteine-sulfinic acid (**183**) residues respectively. Residues analogous to the substrate-binding residues of the β-subunit were also identified (βY72, βY77, βY81) – together suggesting a similarly small and hydrophobic substrate-binding region. With a proposed active site and binding residues identified, the Equi_NHase homology model was used to rationalise the observed substrate preference using AutoDock to suggest binding modes of these nitriles.

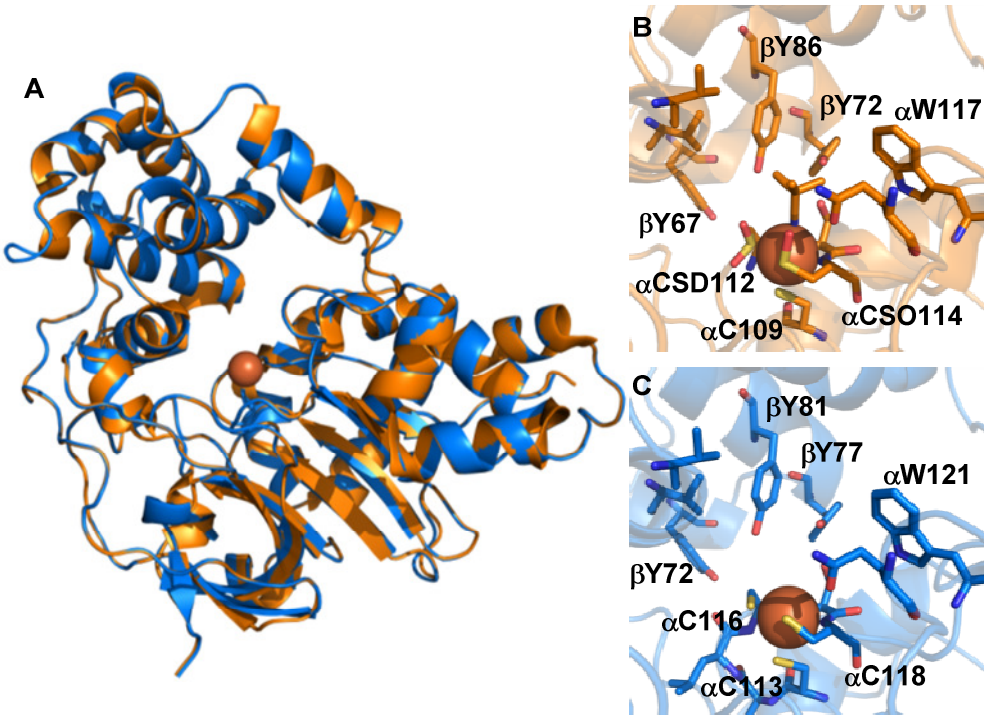


Figure 57: (**A**) Overlay of the Equi_NHase homology model secondary structure (blue) with the crystal structure of NHase from *Rhodococcus erythropolis* N-771 (orange, PDB 3WVD). (**B**) Substrate and Fe-binding residues of the NHase from *Rhodococcus erythropolis* N-771 (PDB 3WVD). (**C**) Equi_NHase homology model showing residues proposed to be involved in Fe and substrate binding. In all structures, Fe is represented as an orange sphere.

As expected, docking of the small aromatic nitrile **185** solely returned poses whereby the substrate was coordinated to the Fe centre via the nitrile N (**Figure 58A**). Docking of the larger substrate **191** found that although coordination of the nitrile N to the Fe centre is possible, the steric bulk of the hydrophobic tyrosines (βY77 and βY81) appeared to cause the nitrile to coordinate at a non-axial geometry (**Figure 58B**). Indeed placing substrate at the optimal geometry for coordination to Fe resulted in the clash of **191** with βY81. This suggests that the unfavourable steric repulsion of these

residues with the larger substrates **190** and **191** could cause their inefficient hydration by Equi_NHase. Docking of the biaryl nitrile **193** failed to return any poses of substrate with the nitrile coordinated to Fe, suggesting that this substrate is too large to effectively coordinate to Fe.

Interestingly docking of the phathalide nitrile **192** solely returned poses of substrate with the carbonyl oxygen coordinated to Fe (**Figure 58C**), suggesting that this competitive coordination may inhibit hydration of this substrate by Equi_NHase. The coordination of carbonyl residues in this manner would explain the apparent substrate specificity of the NHases, resulting in their selective mono-hydration of di-nitriles. To assess if inhibition of the mono-nitrile product **201** may be responsible for the mono-hydration of **187**, 2-cyano benzamide (**201**) was also docked into the Equi_NHase active site in the same way. Similarly this returned poses with the substrate coordinated to Fe via the carbonyl O rather than nitrile (**Figure 58D**), suggesting this could lead to inhibition of the NHase and hence mono-hydration of **187**. Much more detailed analysis of this effect is required to determine how valid this conclusion is, as mono-hydration could also result from poor catalyst lifetime of the NHase coupled with a slow second hydration. The coordination of amides to the Fe-centre of NHases in this manner could however explain the observed mono-hydration activity of many NHases. Binding modes of additional nitrile substrates and amide products generated in the same manner are shown in **Appendix 3** for comparison.

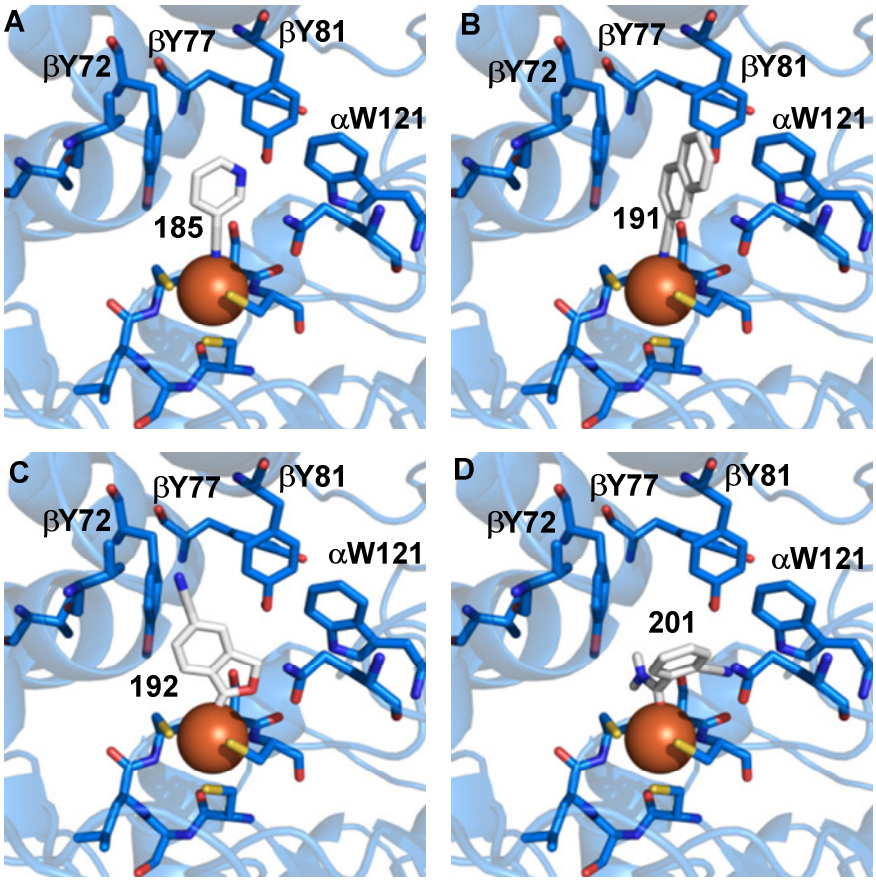
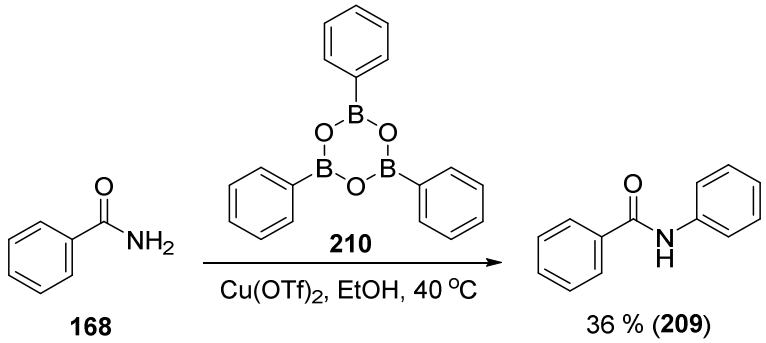


Figure 58: Binding modes of various nitrile substrates to the Equi_NHase homology model proposed by AutoDock 4.2. (**A**) Nicotinonitrile (**185**), (**B**) 2-napthonitrile (**191**), (**C**) 5-cyano phthalide (**192**) and (**D**) 2-cyano benzamide (**201**).

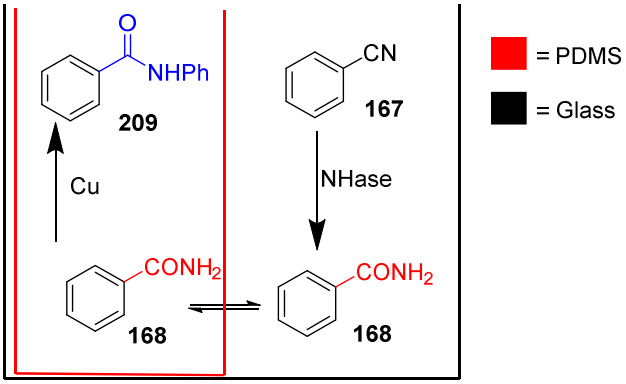
5.2.2 Chan-Lam Evans Couplings for Cascade Reactions

With a workable biocatalytic nitrile hydration in hand, conditions for the efficient CLE coupling of these amides in single pot transformations were sought. Given that boronic acids had been found to form adducts with the active site CSO residue of NHases,[263] and the finding that transition metals inhibit enzyme activity in Chapters 3 and 4, it was envisioned that PDMS compartmentalisation of the NHase and CLE reactions may be required. Accordingly the test NHase-CLE reaction was conducted using the conversion of benzonitrile (**167**) to *N*-phenyl benzamide (**209**) and a number of literature conditions for the CLE coupling of amides. Although conditions employing the boroxime **210** were efficient at producing *N*-phenyl benzamide (**209**) in isolation (**Scheme 76**),[279] when combined with the NHase reaction through PDMS compartmentalisation no arylated product could be detected – even with stoichiometric Cu source (**Table 20**, entries 1 and 2). It was presumed that this was because the very hydrophobic nature of the **210** causes its flux into the PDMS membrane. Therefore the more polar trifluoroborate salt **98** was selected – although this failed to provide any evidence of arylated product (**Table 20**, entry 3).



Scheme 76: Preparation of the *N*-phenyl benzamide (**209**) standard. Conditions: benzamide (100 mM), $Ph_3B_3O_3$ (**210**, 0.5 eq), $Cu(OTf)_2$ (10 mol %), EtOH, 40 °C, 24 hrs.

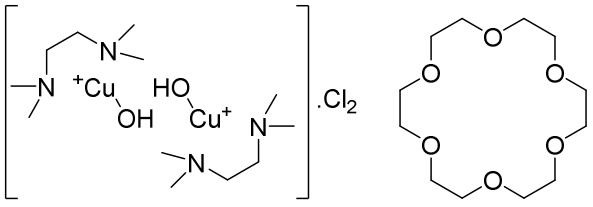
Attention was therefore turned to other reported conditions for the CLE coupling of amides.[274,277] These methods did not efficiently promote the CLE reaction when PDMS compartmentalisation was employed however (**Table 20**, entries 4 and 5). It was noted that many of these literature protocols involved high concentrations of amide (typically 1 – 5 M), whilst the integrated NHase-CLE reactions would provide a maximum 50 mM concentration of intermediate amide. Considering that the amide N is poorly nucleophilic, it was considered that the low concentration of intermediate amide could limit the efficiency of amide coordination to Cu^{II} and hence reduce CLE efficiency. Since increasing the concentration of intermediate amide by two orders of magnitude was considered unlikely in the timescale of this project, methods of increasing the nucleophilicity of the amide N were considered.



Entry	Solvent	Cu source (mol %)	Base	Ar Donor	% Yield (209)
1	EtOH	$Cu(OTf)_2$ (10)	None	**210**	N.C.
2	EtOH	$Cu(OTf)_2$ (100)	None	**210**	N.C.
3	EtOH	$Cu(OTf)_2$ (100)	None	$PhBF_3K$	24
4	CH_2Cl_2	$Cu(OAc)_2$ (100)	None	$PhBF_3K$	27
5	CH_2Cl_2	$Cu(OAc)_2$ (100)	DIPA	$PhBF_3K$	N.C.

Table 20: Attempted one-pot NHase-CLE reactions using stoichiometric Cu sources. Conditions: benzonitrile (50 mM), Equi_NHase (10 μM), Ar Donor (2 eq), base (2 eq), rt overnight then 40 °C, open air, overnight. Yield after chromatography. DIPA = diisopropylamine.

One approach involved the use of strong bases, which may be capable of transiently deprotonating the amide NH. After screening a number of bases and copper sources, potassium *tert*-butoxide (^{t}BuOK) appeared to afford reasonable conversion of **168** to **209** with $Cu(OAc)_2$ and Cu-TMEDA (**211**) as stoichiometric Cu source. It has been reported previously that the bacisity of ^{t}BuOK is impacted significantly by solvent effects and the presence of chelating agents such as 18-crown-6 (**212**),[280-282] therefore a screen of solvents was conducted with the hope of increasing the extent of amide deprotonation and hence allowing catalytic quantities of Cu catalyst to be employed (**Table 21**). This subsequently found dioxane in combination with 18-crown-6 to afford excellent conversion of **168** to **209** - allowing **209** to be isolated in 87 % yield.



Cu-TMEDA (**211**) 18-crown-6 (**212**)

Figure 59: Structures of Cu-TMEDA (**211**) and 18-crown-6 (**212**)

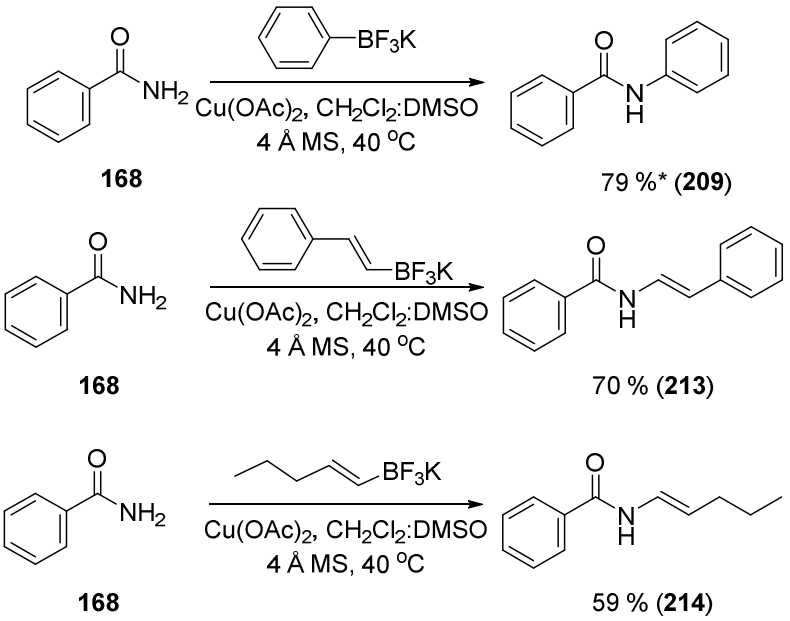
Attempts at integrating these optimised conditions (**Table 21**, entry **1**) with the Equi_NHase-catalysed hydration of **167** however failed to afford any conversion of **167** to intermediate amide **168** or arylated product **209**. Visual inspection of the reaction after incubation overnight found that a significant quantity of dioxane had fluxed through the PDMS membrane from the CLE

compartment into the NHase compartment – presumably causing deactivation of the NHase. The high solubility of dioxane in PDMS has been previously reported,[200] and therefore the analogous conditions using DMSO (**Table 21**, entry 6) were employed since DMSO is reported to be significantly less soluble in PDMS. Although intermediate amide **168** could be detected in this instance, no arylated product **209** was detected. Since the thimbles used in these reactions are not sealed at one end, it was considered that the lack of reactivity may be because of the mechanical movement of water vapour into the CLE compartment through this open end – resulting in deactivation of potassium *tert*-butoxide. The mechanical movement of water into the CLE compartment was likely exacerbated by use of the highly water-miscible solvents dioxane and DMSO. Since Cu-TMEDA (**211**) had also afforded reasonable conversion to arylated product in CH_2Cl_2 (**Table 21**), a less water-miscible solvent, the integrated NHase-CLE reaction was also attempted using this catalyst. The hydrophobic nature of the diamine ligand of (**211**) appeared to promote flux of the Cu catalyst through the PDMS membrane, resulting in the detection of no arylated product **209**.

Table 21: Conversion to *N*-phenyl benzamide (**209**) from the CLE of benzamide (**168**) in isolation (method A) or the integrated NHase-CLE reaction of benzonitrile (**167**, method B). Conditions as prior except potassium *tert*-butoxide (2 eq) was also used. *Conversion determined by analytical HPLC using a calibration curve of **209**. †Isolated yield = 61 %. ‡Isolated yield = 87 %.

Entry	Solvent	Cu source (mol %)	Additive	Method	Conversion (209, %)*
1	CH_2Cl_2	Cu-TMEDA (10)	None	A	24
2	CH_2Cl_2	Cu-TMEDA (100)	None	A	39
3	CH_2Cl_2	$Cu(OAc)_2$ (100)	None	A	69†
4	CH_2Cl_2	$Cu(OAc)_2$ (10)	**212**	A	46
5	CH_2Cl_2	$Cu(OAc)_2$ (10)	None	A	35
6	DMSO	$Cu(OAc)_2$ (10)	**212**	A	70
7	DMSO	$Cu(OAc)_2$ (10)	None	A	66
8	Dioxane	$Cu(OAc)_2$ (10)	**212**	A	92‡
9	Dioxane	$Cu(OAc)_2$ (10)	None	A	43
10	Dioxane	$Cu(OAc)_2$ (10)	**212**	B	N.C.
11	DMSO	$Cu(OAc)_2$ (10)	**212**	B	11
12	CH_2Cl_2	Cu-TMEDA (100)	None	A	11

As an alternative to promoting amide nucleophilicity by deprotonation, attempts were made to utilise the PDMS membrane to flux the amide product from the aqueous biotransformation medium into a small amount of an organic solvent – thereby effectively concentrating it. Using previously reported conditions,[278] the effective CLE coupling of benzamide (**168**) at 500 mM to various aryl and vinyl trifluoroborate salts was achieved using 10 mol % $Cu(OAc)_2$ – affording the secondary amides **209**, **213** and **214** in good isolated yields (**Scheme 77**). The arylation of **168** at 50 mM however, afforded only 10 % isolated yield of product using the same conditions and potassium phenyl trifluoroborate (**Scheme 77**).



Scheme 77: Isolated yields from the arylation and vinylation of benzamide (**168**) using $Cu(OAc)_2$ and potassium trifluoroborate salts. Conditions: benzamide (500 mM), potassium trifluoroborate salt (3.0 eq), $Cu(OAc)_2$ (10 mol %), CH_2Cl_2:DMSO (1:1), 4Å molecular sieves (MS), 40 °C, overnight.*Isolated yield = 10 % when [**168**] = 50 mM.

The NHase-CLE reaction was subsequently attempted using these conditions, with the CLE components placed inside a PDMS thimble. Despite a number of attempts at varying the amount of Cu source and molecular sieves employed, only 10 % arylated product **209** could be isolated when using stoichiometric Cu (**Table 22**). Determining the flux of amide through PDMS into the organic solvent subsequently showed that a maximum concentration of 52 mM was achieved inside the thimble, suggesting that this approach to improving CLE yield was infeasible.

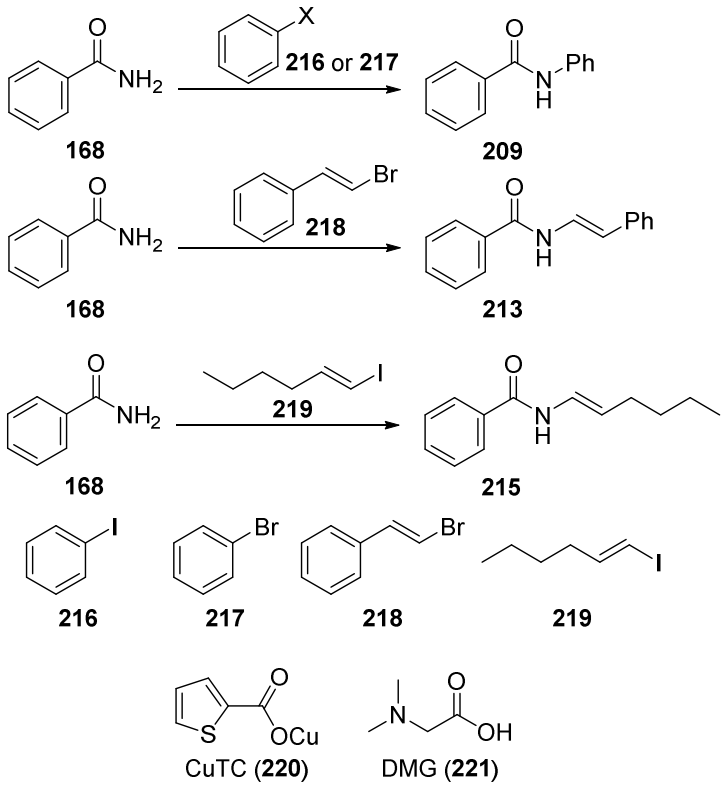
Table 22: Isolated yield of **209** from the CLE coupling of benzamide (**168**) to potassium phenyl trifluoroborate when compartmentalised in a PDMS thimble. Conditions: benzamide (50 mM total in 10 mL H_2O) and $Cu(OAc)_2$, potassium phenyl trifluoroborate (3.0 eq) and 4 Å MS in CH_2Cl_2:DMSO (1:1), 40 °C, overnight. N.C. = no conversion. Yield after chromatography.

Cu Catalyst Loading (mol %)	4 Å Molecular Sieves Loading (mg)	% Yield (209)
10	375	N.C.
100	375	10
10	750	N.C

5.2.3 Goldberg Couplings for Cascade Reactions

As various CLE reaction conditions had failed to provide efficient NHase-CLE arylations, the Goldberg reaction was explored as an alternative. Following a number of literature conditions however,[283-285] only trace amounts of arylated or vinylated benzamide were detected when benzamide (**168**) was used at 50 mM (**Table 23**). Similar to CLE-catalysed reactions, significantly

improved yields were afforded when higher amide concentrations were used (**Table 23**), suggesting that amide nucleophilicity was a limiting factor.



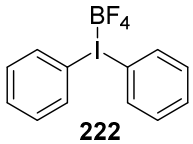
[Amide]	Halide	Solvent	Base	Temperature (°C)	Cu catalyst (mol %)	% Yield
50	**218**	DMF	^{t}BuOK	100	CuTC (15)	N.C. (**213**)
50	**218**	DMF	^{t}BuOK	Reflux	CuTC (15)	N.C. (**213**)
50	**217**	DMF	^{t}BuOK	Reflux	CuTC (15)	N.C. (**209**)
50	**216**	Dixoane	K_3PO_4	Reflux	CuTC (15)	N.C. (**209**)
50	**216**	Dixoane	K_3PO_4	Reflux	CuI.DMG (15)	N.C. (**209**)
500	**219**	NMP	Cs_2CO_3	90	CuTC (15)	55 (**215**)
50	**219**	NMP	Cs_2CO_3	90	CuTC (15)	15 (**215**)

Table 23: Isolated yield from the screening of various literature conditions for the Goldberg coupling of benzamide (**168**) to various vinyl and aryl halides. Conditions: benzamide (50 or 500 mM), halide (2.0 eq), base (2.0 eq), Cu source, solvent and heating as specified overnight. CuTC = copper(I)thiophene-2-carboxylate (**220**). DMG = *N,N*-dimethyl glycine (**221**). NMP = *N*-methyl-2-pyrrolidone. N.C. = no conversion. Yield after chromatography.

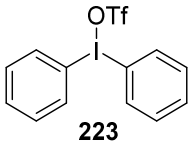
5.2.4 Copper-Catalysed Coupling to Iodonium Salts for Cascade Reactions

A number of reports suggest the coupling of amides to iodonium reagent to be efficient.[286-288] Given that amide nucleophilicity appeared to be limiting the effective integration of NHase and CLE reactions, it was considered that the “hyperleaving” ability of the aryl iodide group in these reagents may compensate and allow for effective NHase-arylation cascades. Screening a number of literature conditions for such couplings found the combination of CuI, DBU and ACN to allow

isolation of **209** in 48 % yield using the diphenyliodonium triflate **223** (**Table 24**). Attempts to integrate these conditions with the NHase reaction using PDMS compartmentalisation however failed to afford any arylated product. Given the water sensitivity of iodonium reagents, integrated reactions were also run with molecular sieves added, however this failed to afford any arylated product (**Table 24**).



222



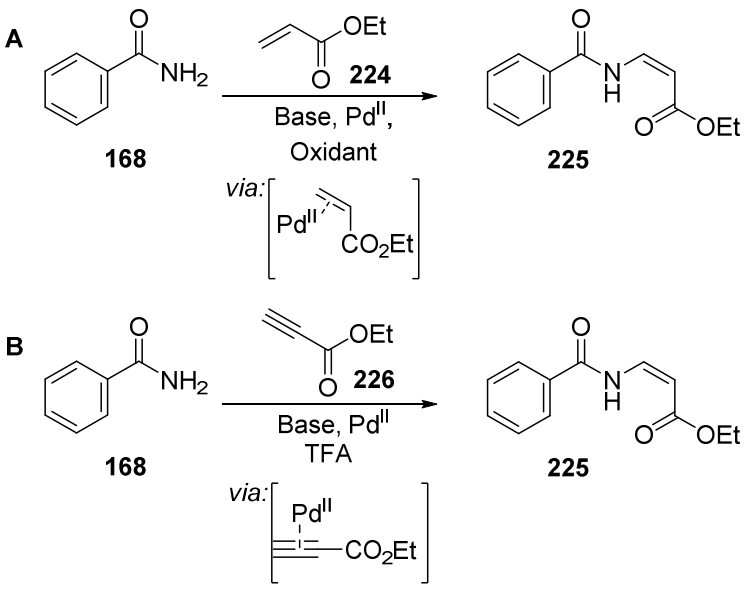
223

Iodonium Salt	Cu Source	Base	Solvent	Conversion* (%, 209)
Ph_2IBF_4	CuI	Na_2CO_3	CH_2Cl_2	N.C.
Ph_2IBF_4	CuI	K_2CO_3	Toluene	N.C.
Ph_2IBF_4	CuI	DBU	ACN	1
Ph_2IBF_4	CuI	^tBuOK	DMSO	3
Ph_2IBF_4	$Cu(OTf)_2$	None	MeOH	N.C.
Ph_2IBF_4	$Cu(OTf)_2$	None	CH_2Cl_2	N.C.
Ph_2IOTf	CuI	Na_2CO_3	CH_2Cl_2	21
Ph_2IOTf	CuI	K_2CO_3	Toluene	46
Ph_2IOTf	CuI	DBU	ACN	$53^\dagger$
Ph_2IOTf	CuI	^tBuOK	DMSO	59
Ph_2IOTf	$Cu(OTf)_2$	None	MeOH	N.C.
Ph_2IOTf	$Cu(OTf)_2$	None	CH_2Cl_2	N.C.
Ph_2IOTf	None	K_2CO_3	DMF	N.C.

Table 24: Conversion of benzamide (**168**) to *N*-phenyl benzamide (**209**) obtained from the Cu-catalysed coupling of **168** to various diphenyl iodonium salts. Conditions: benzamide (50 mM), Cu source (10 mol %), base (2.0 eq), 40 ^oC, N_2, overnight. *Determined by analytical HPLC using a calibration curve of **209**. †Isolated yield = 48 %.

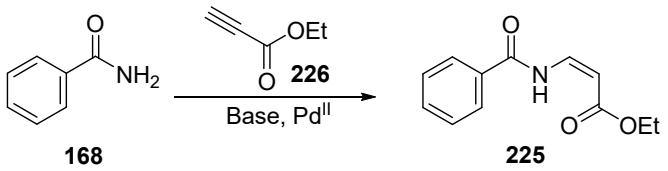
5.2.5 Pd-Catalysed Couplings for Cascade Reactions

A number of palladium-catalysed enamide syntheses have also been reported which involve the coupling of amide to an alkene or alkyne (**Scheme 78**).[289,290] In both reactions, it is thought that amide attacks an electron-deficient σ-Pd intermediate. It was therefore considered that the electron-deficient nature of this intermediate may promote more efficient attack by relatively dilute intermediate amide.



Scheme 78: Palladium-catalysed synthesis of enamides by coupling of primary amides with (**A**) electron-deficient alkenes and (**B**) electron-deficient alkynes.

The coupling of **168** to **224** (**Scheme 78A**) afforded poor yields of **225** when attempted in isolation from the NHase at 50 mM benzamide, whilst the coupling of **168** to **226** (**Scheme 78B**) appeared significantly more efficient (**Table 25**). Attempts to conduct this coupling within a PDMS thimble afforded some conversion to enamide **225** (**Table 25**). The thimble appeared to deteriorate over the reaction course, however. As TFA (trifluoroacetic acid) has previously been reported to degrade PDMS,[200] the same reactions were run using acetic acid in place of TFA. No enamide product was detected however, suggesting that acetic acid may not be a strong enough acid to promote protodemetallation.

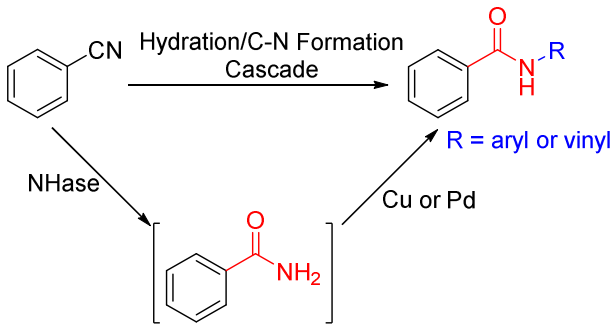


Acid	Method	% Yield (225)
TFA	A	42
TFA	B	26
AcOH	A	N.C.
AcOH	B	N.C.

Table 25: Isolated yields of **225** from the coupling of benzamide (**168**) to ethyl propiloate (**226**). Method A: coupling conducted in isolation from NHase reaction components. Method B: A PDMS thimble containing $Pd(OAc)_2$, alkyne and base was added to a solution of benzamide (**168**) in water. Conditions: benzamide (**168**, 50 mM), ethyl propioloate (**226**, 1.5 eq), NaOAc (2.0 eq), acid (5.0 eq), $Pd(OAc)_2$ (10 mol %), toluene, 70 ^{o}C, overnight. TFA = trifluoroacetic acid. Yield after chromatography.

5.3 Conclusions and Outlook

The development of methods for the atom economical synthesis of amides has received much recent attention due to the prevalence of amide bond-formation reactions in industrial synthetic routes.[248] The synthesis of primary amides has been achieved on an industrial scale for many years by the biocatalytic hydration of nitriles using nitrile hydrate (NHase) enzymes.[239] This chapter describes the attempts made to combine this biocatalytic method for the synthesis of primary amides with Cu- and Pd-catalysed reactions for arylation and vinylation to secondary amides. It was hoped that this combination would allow the synthesis of secondary amides and enamides without the requirement for stoichiometric activating reagents, hence providing an atom economical synthesis of secondary amides (**Scheme 79**).



Scheme 79: Attempted combination of biocatalytic nitrile hydration with Cu- and Pd-catalysed couplings for the synthesis of secondary amides and enamides.

The substrate scope of a previously reported NHase (Equi_NHase) was initially explored, and found to efficiently hydrate a series of aromatic and aliphatic nitriles to the corresponding amides. The substrate preference of this particular enzyme was also briefly rationalised by the preparation of a homology model of Equi_NHase and substrate docking. As with some substrates described in Chapter 4, this suggested that substrates possessing highly coordinating groups may not be efficiently hydrated by NHases because of the competitive coordination of this group with the catalytic Fe centre. This observation may explain the observed selective mono-hydration activity of NHases with di-nitriles, as the polar oxygen of the resultant amide bond may act as a competitive inhibitor of the second nitrile hydration reaction. Significantly more work would be required to determine the validity of this statement, which was considered beyond the scope of this thesis.

After establishing a biocatalytic nitrile hydration, the combination of this reaction with Chan-Lam-Evans couplings (CLE) was investigated. After screening many literature conditions for the CLE coupling of amides, it was found that the limited nucleophilicity of the amide N and the relatively low concentration of intermediate amide afforded by the NHase may be a limiting factor for efficient CLE-NHase reactions. Subsequently, strong bases were employed with the hope of transiently deprotonating the intermediate amide, thereby promoting its nucleophilicity. Although this method was successful at the arylation of amides, the integration of this process with the NHase reaction was unsuccessful – affording poor efficiency of amide arylation. This was presumed to be because the PDMS compartmentalisation strategy used may allow the movement of water into the CLE

compartment, resulting in deactivation of the strong base. Although a number of attempts at employing free and immobilised preparations of Equi_NHase in organic solvents were attempted (not reported), none were found to allow the efficient hydration of nitriles in microaqueous reaction systems – suggesting a thermodynamic requirement for a vast excess of water in the nitrile *hydration* reaction.

Attention was subsequently turned to the Cu^{I}-catalysed Goldberg coupling of amides with aryl and vinyl halides, as an alternative to the CLE reaction, although these reactions were found to be inefficient at effective couplings at the concentration of amide afforded by the NHase. To overcome the poor nucleophilicity of the amide N, a number of alternative reactions using strongly electrophilic reagents were used in the hope of priming attack from the amide N. Although the Cu-catalysed coupling of amides to iodonium salts appeared efficient in isolation, integration with the NHase reaction was unsuccessful – presumably due to the water sensitivity of the iodonium reagents. Although Pd-catalysed couplings of amides and alkynes were also found to be efficient in isolation, their integration with the NHase reaction using PDMS compartmentalisation was also unsuccessful because of the requirement for strong acids to aid protodemetallation (to which the PDMS membrane appeared unstable). In conclusion although a robust biocatalytic nitrile hydration method was established, its integration with various metal-catalysed amide arylation and vinylation reactions in single pot transformations was unsuccessful.

6. Summary and Outlook

The combination of biocatalytic and chemocatalytic reactions into single pot transformations, termed integrated catalysis, can allow transformations to be designed which would be beyond the scope of either type of catalysis in isolation.[3,4] Of particular interest is exploitation of the selectivity inherent to biocatalysis combined with chemocatalysts which extend these transformations beyond the repertoire of enzymes. By using enzymes to confer selectivity in this way, the efficiency of a process can be significantly improved by reducing the need for by-product separation and resolution steps.[5] In this thesis, the regioselectivity of the flavin-dependent halogenases is utilised to selectively install a reactive handle for further functionalisation reactions by palladium catalysis.

In an attempt to deliver biocatalysts more amenable to combination with chemocatalysis in this way, a thermophilic tryptophan halogenase (Th_Hal) was investigated. This enzyme was found to possess increased thermostability and catalyst lifetime compared to the previously-characterised mesophilic tryptophan halogenases. The increased thermostability of Th_Hal, quantified by the melting temperature (T_m), was found to be of the same order of magnitude obtained by random mutagnesis of the mesophilic tryptophan halogenases (**Figure 60**).[164,165] Since enzyme evolution by random mutagenesis requires a significant screening effort, this finding suggests that the discovery of naturally-occurring thermostable enzyme variants could be a viable alternative to such approaches without the need for time- and labour-consuming screening efforts. Alternatively such naturally-occurring thermophilic enzymes could be used as the basis for random mutagenesis approaches – affording variants of further increased thermostability. The need for flavin-dependent halogenases of increased stability is thought to be key to the development of biocatalytic halogenation protocols suitable for application in synthesis.[76,78] Since these enzymes operate via an electrophilic aromatic substitution step, an energetically unfavourable process, their kinetic values are typically in the region of 1 min^{-1}. Combined with poor stability of the FlHals, this leads to biocatalysts of poor productivity. By increasing catalyst lifetime, it is therefore hoped that FlHals with improved productivity can be obtained.

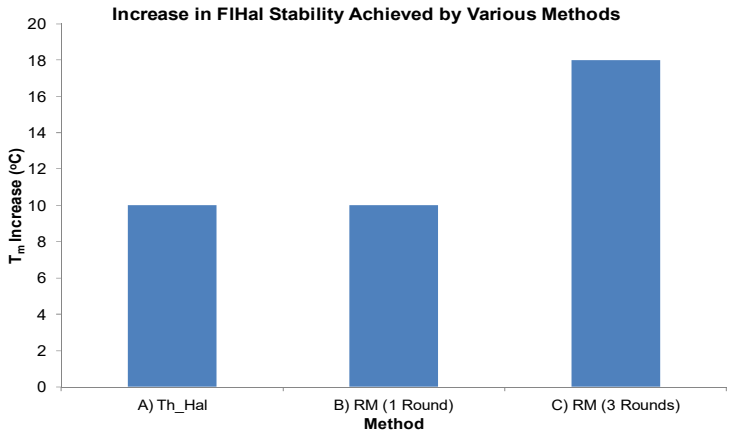


Figure 60: Comparison of various methods used to improve FlHal stability. A) Discovery of thermophilic FlHal variants (this work). B) A single round of random mutagenesis (RM).[165] C) Three rounds of random mutagenesis (RM).[164] A) Is compared to the closest homologue SttH whilst B) and C) are compared to wild-type.

Despite the increased stability of Th_Hal, its utility as a biocatalyst was limited because of a seemingly narrow substrate scope compared to the mesophilic tryptophan halogenases. This effect has been reported previously, whereby the rigidity thought to confer increased stability can reduce binding efficiency of non-natural substrates by limiting the conformational flexibility required to bind these substrates.[164] This observation highlights the advantage of using directed evolution and random mutagenesis to engineer biocatalysts. Whilst nature has delivered a thermostable variant capable of productive tryptophan (**31**) halogenation, directed evolution would deliver an enzyme capable of catalysing halogenation of the desired substrate. In future work, it may therefore be useful to use directed evolution to broaden the substrate scope of Th_Hal – ideally affording a stable FlHal with a synthetically relevant substrate scope. Since Th_Hal is a more thermostable starting point, the generally denaturing effect of directed evolution may be less pronounced and hence afford enzymes more relevant for biocatalysis.[164]

The crystal structure of Th_Hal was also solved, which revealed a number of structural features thought to confer increased stability. These observations could be exploited in the evolution of mesophilic FlHals towards increased catalyst lifetime, by targeting random mutagenesis to the regions thought to contribute to the increased stability of Th_Hal. Others have reported attempts at using targeted mutagenesis to improve the productivity of FlHals by introducing random mutations to the flavin-binding portion of a RebH variant, although productivity improvements were minimal.[174] Given the energetic demand of the electrophilic aromatic substitution step, it is unlikely that a change in the efficiency of C4a-hydroperoxyflavin (**28**) generation would have a meaningful impact upon the halogenation step – explaining the lack of productivity improvement. By using these methods to target areas which contribute to the stability of Th_Hal however, enzymes capable of withstanding the time required for the slow halogenation reaction may be afforded and could prove to be more productive biocatalysts.

Although significant work is likely required before the FlHals reach synthetic maturity, demonstrations that they can be used to regioselectiviely halogenatate different positions of similar aromatic substrates mean this work is worthwhile. Aryl halides are versatile synthetic intermediates which can be further functionalised through Pd-catalysed cross-coupling reactions.[46] This thesis exploits this fact to allow regioselective C-C bond formation reactions to be conducted - using FlHals to regioselectivley install the reactive C-Br moiety (**Scheme 80**). Particular emphasis is placed upon conducting these transformations in a one-pot manner, whereby the intermediate aryl bromide is not isolated or purified between the two reactions. This is an important aspect from the industrial perspective, since such intermediary workup and purification process can significantly reduce the efficiency of a process overall (**Figure 61**).[3] By necessitating the consumption of auxillary chemicals and solvents, as well as energy demand and reactor capacity, workup steps are seen to reduce the space-time yield of a synthetic route. Although an attractive prospect, the combination of biocatalysts and chemocatalysts into one-pot reactions is often not straightforward due to the mutual deactivation of the two types of catalyst and differences in their optimal operation conditions.

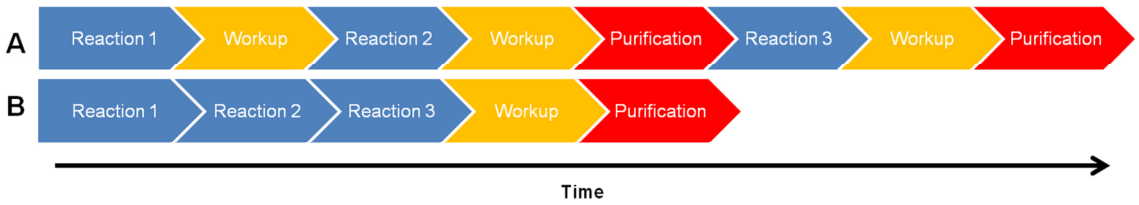
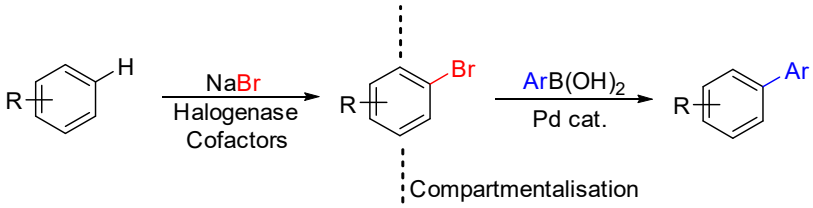


Figure 61: Schematic representation of the impact that workup and purification processes between individual reactions has upon the overall efficiency of a process (**A**) compared to syntheses which do not require these steps (**B**).

The successful combination of FlHals with Pd-catalysed cross-couplings (**Scheme 80**) was achieved in a number of ways. After noting the mutual incompatibility of the the FlHals and Pd-catalysts, as well as deactivation of the Pd-catalysts by the various co-factors required for FlHal halogenation, the successful combination of biocatalytic halogenation with Pd-catalysed SMC reactions was achieved by employing an increased Pd-loading and compartmentalisation of the FlHal enzymes using molecular-weight cut-off filtration. Whilst good yields of arylated products could be obtained, and a number of different aromatic scaffolds effectively arylated through the use of FlHals with differing substrate specificity, the scope of this process was quite narrow with respect to boronic acid coupling partners (**Figure 62A**).



Scheme 80: General representation of the reactions achieved through the combination of FlHal bromination of Pd-catalysed SMC reactions.

The high loading of palladium catalyst required for effective arylation and narrow substrate scope of FlHal-SMC reactions using MWCO compartmentalisation was proposed to be due to a combination of dilute aryl bromide resulting from the FlHal reaction and inhibitory effects from the co-factors not removed by MWCO filtration. Therefore in an attempt to improve this process, productivity of the FlHals was improved through the formation of FlHal-Fre-ADH CLEAs in the first instance. The improved stability of these materials allowed higher concentrations of intermediate aryl bromide to be achieved – allowing effective FlHal-SMC reactions with lower loadings of Pd catalyst. Additionally, a broader range of boronic acid coupling partners could be effectively utilised in FlHal-SMC reactions using FlHal CLEAs – allowing a broader range of arylated products to be obtained (**Figure 62B**). The heterogeneous nature of the CLEAs also allowed the efficient upscale of these reactions to the mmol scale, as well as recycling of the biocatalyst over multiple rounds of halogenation.

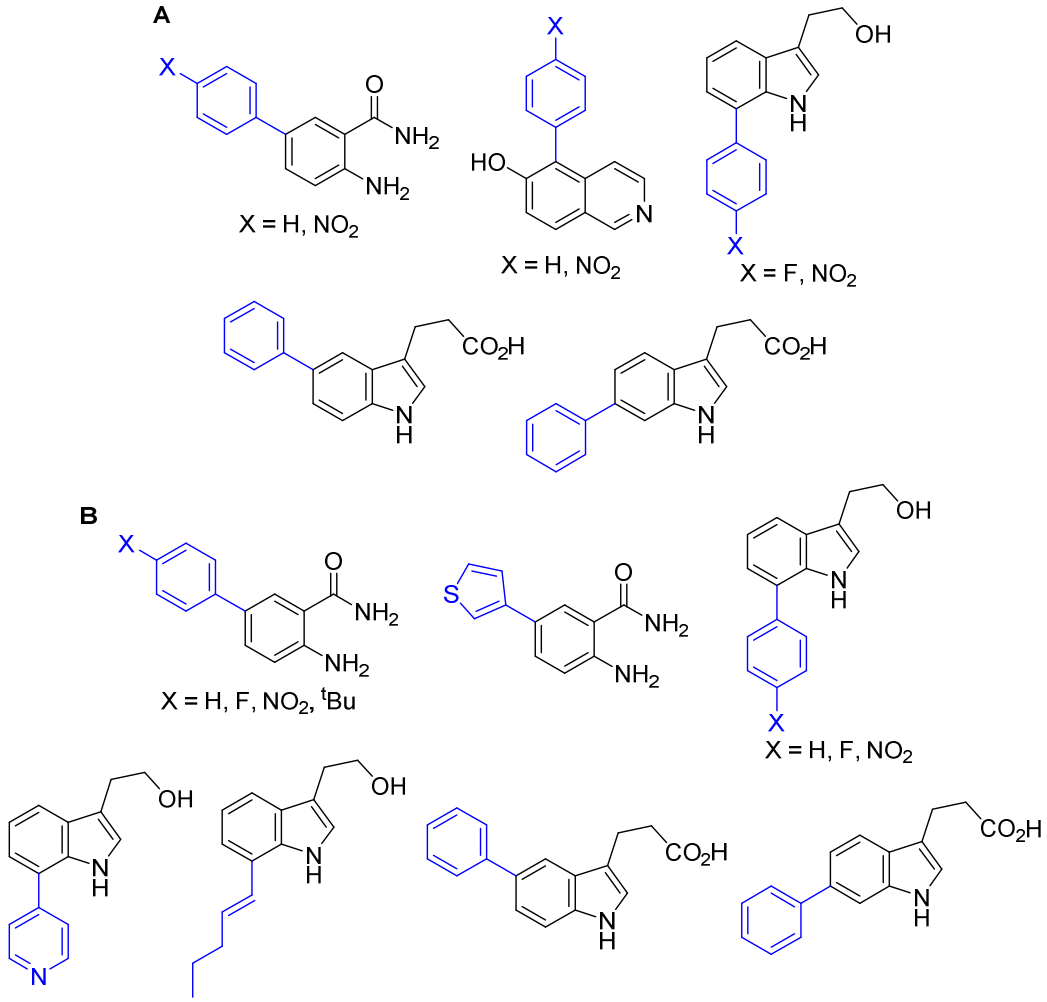


Figure 62: (**A**) Arylated products isolated from the integrated FlHal-SMC reactions using MWCO filtration compartmentalisation of purified enzymes. (**B**) Arylated, heteroarylated and vinylated products obtained from integrated FlHal-SMC reactions using heterogeneous FlHal-Fre-ADH CLEAs.

Although cross-linking of the enzyme appeared to reduce the extent of its deleterious effect on SMC efficiency, its removal prior to cross-coupling was still required to obtain good yield of arylated product. The heterogeneous nature of this biocatalyst means its removal is more facile compared pure protein, however this process still requires intermediary processing which has a deleterious effect upon space-time yield (**Figure 64**).[3] To circumvent the need for biocatalyst removal, integration of the FlHal and SMC reactions was also achieved using PDMS membranes. As these membranes separate components based on charge, they allowed the charged biocatalytic and chemocatalytic reaction components to be effectively separated whilst the non-polar intermediate aryl bromide is able to move between the two compartments.

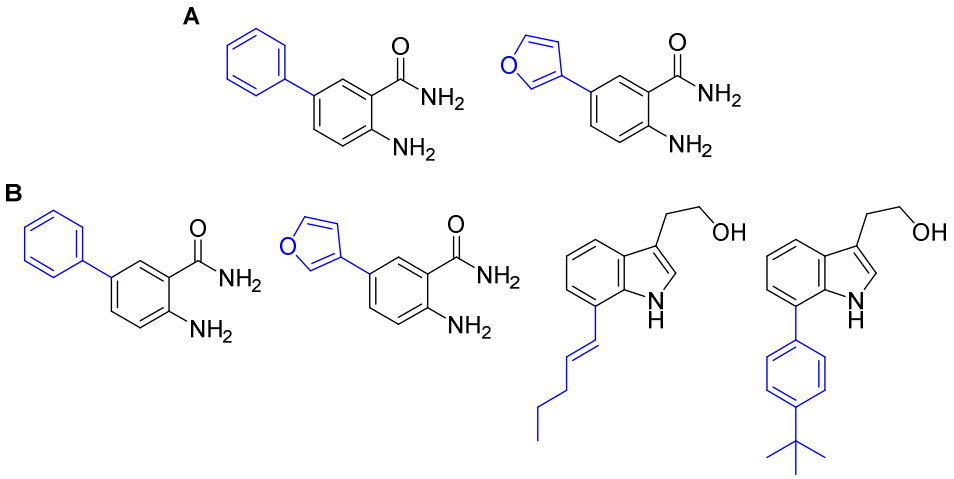


Figure 63: (**A**) Products obtained from FlHal-SMC reactions using PDMS compartmentalisation of purified FlHal enzymes and cofactors from SMC reaction conditions. (**B**) Products obtained from FlHal-SMC reactions using PDMS compartmentalisation of FlHal-Fre-ADH CLEAs and cofactors from SMC reaction conditions.

Separation of the FlHal and SMC reactions using PDMS membranes allowed integrated FlHal-SMC reactions to occur in a one-pot manner without the need for intermediate processing – representing an improvement in space-time yield relative to either of the processes requiring removal of the biocatalyst. The space-time yield of this process was further improved by the use of CLEAs – allowing greater quantities of arylated product to be generated and in some cases a further reduction in loading of the Pd catalyst. Notably, the compartmentalisation of co-factors from SMC reaction components appeared to afford SMC reactions of improved efficiency. As well as allowing a reduction in palladium loading compared to integrated reactions using MWCO filtration (10 mol % compared to 50 mol %), SMC reactions with more challenging electron-rich coupling partners were enabled (**Figure 63A**) – allowing preparation of some compounds which were inaccessible using the MWCO approach with pure protein. The space-time yield of this process was further improved by the use of CLEAs – allowing greater quantities of arylated product to be generated and in some cases a further reduction in loading of the Pd catalyst. Despite allowing improvements in space-time yield, it should be noted that PDMS compartmentalisation limited substrate scope – preventing the use of charged or highly polar substrates due to the need for diffusion of intermediate aryl bromide through the hydrophobic PDMS polymer.

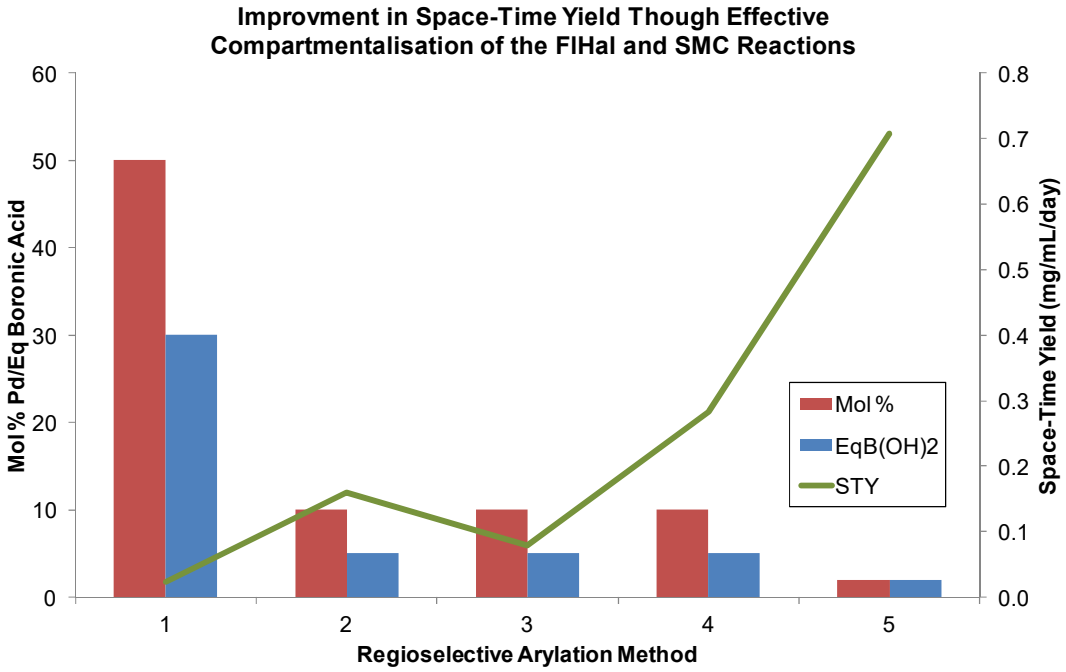
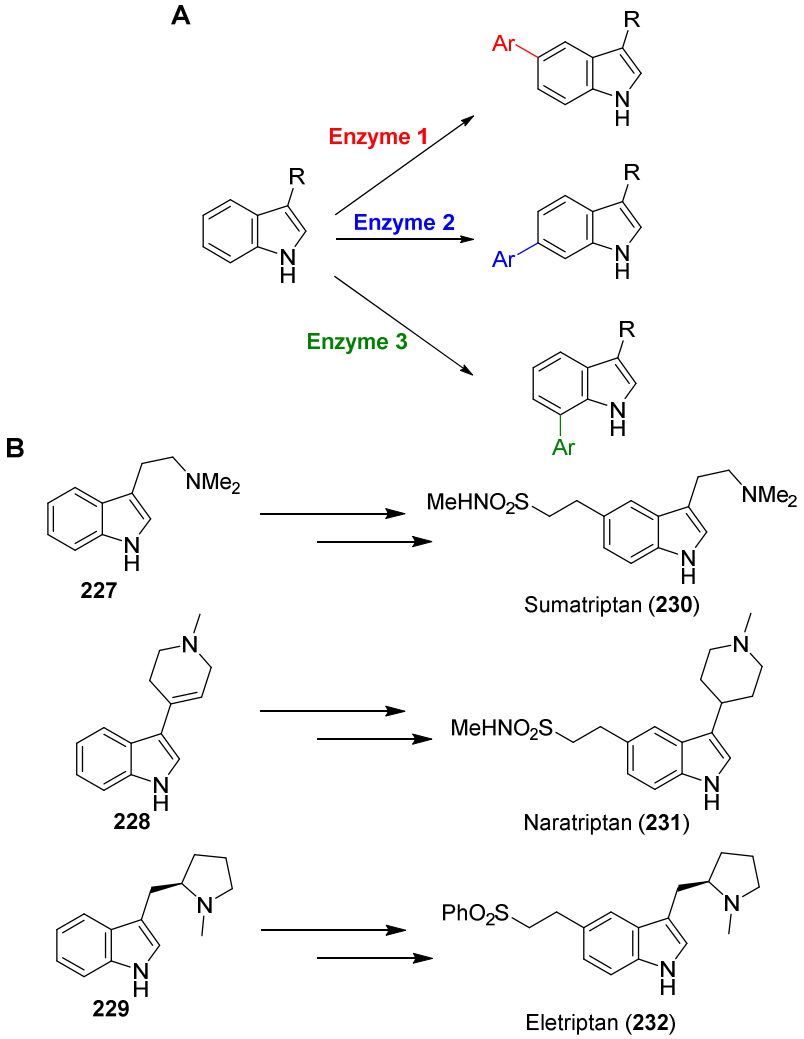


Figure 64: Space-time yield, Pd loading required and equivalents of boronic acid required in each method for conducting the integrated FlHal-SMC reactions. Methods: 1) Filtration of the FlHal biotransformation through a MWCO filter prior to SMC reaction; 2) Removal of the FlHal CLEA prior to SMC reaction; 3) Recycling of a FlHal CLEA through multiple rounds of halogenation prior to SMC coupling of the combined filtrate; 4) Compartmentalisation of a bromination catalysed by pure FlHal from the SMC reaction using PDMS membranes; 5) Compartmentalisation of a FlHal CLEA-catalysed bromination from the SMC reaction using PDMS membranes.

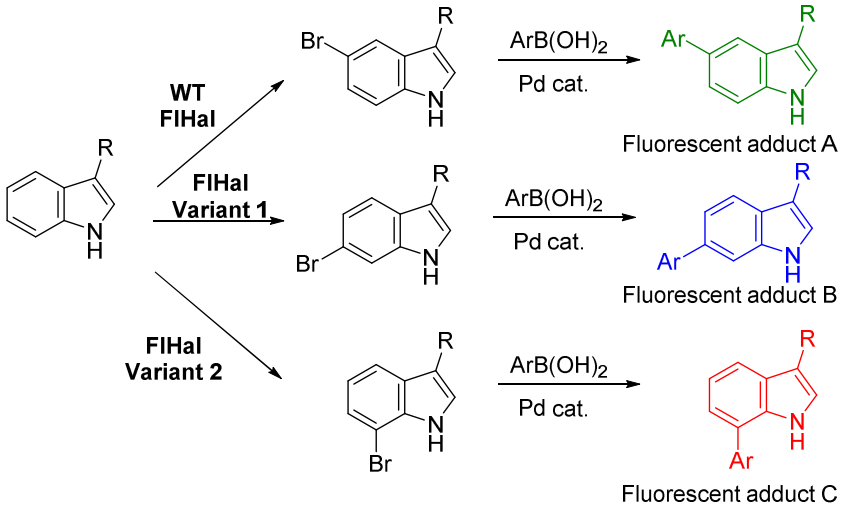
The utility of this combination is best exemplified by the regioselective arylation of the C7, C6 and C5 positions of the indole ring. The selective manipulation of these positions eluded traditional methods of C-H activation until recently,[67] and arylation of the indole C7 and C6 positions requires the use of directing groups to control the position of metallation – a fundamentally atom uneconomic strategy since these groups often need to be removed after the C-H activation has occurred (**Scheme 50**) .[212,213] By using FlHals of different regioselectivity however, these positions could be selectively activated by using different catalysts without the need for such directing groups (**Scheme 81A**). Although the catalyst-controlled method reported herein is attractive, its utility is limited by the substrate scope of the FlHals - with a limited number of substrates which can be arylated and vinylated in this manner. Current work in our group on these integrated reactions is therefore focussed on the engineering of FlHals to allow a broader range of substrates to be effectively arylated and vinylated. Additionally, since this thesis had demonstrated that the FlHals can be integrated with Pd-catalysed cross couplings through effective catalyst compartmentalisation, current work is also focussed on the integration of additional Pd-catalysed reactions with the FlHals – allowing a broader range of regioselective bond formation reactions to be conducted. For example, current work is focussed on mutagenesis of the FlHals to allow halogenation of the substrates **227** to **229**, as it is hoped that integrated FlHal-Heck reactions with these substrates may allow preparation of pharmaceutically-relevant compounds (**Scheme 81B**).



Scheme 81: (**A**) Overview of the regio-selective and regio-divergent C-H arylation of indolic substrates using FIHal-SMC cascades. (**B**) Proposed synthetic routes to the anti-migraine compounds **230** to **232** using integrated FIHal-Pd reactions.

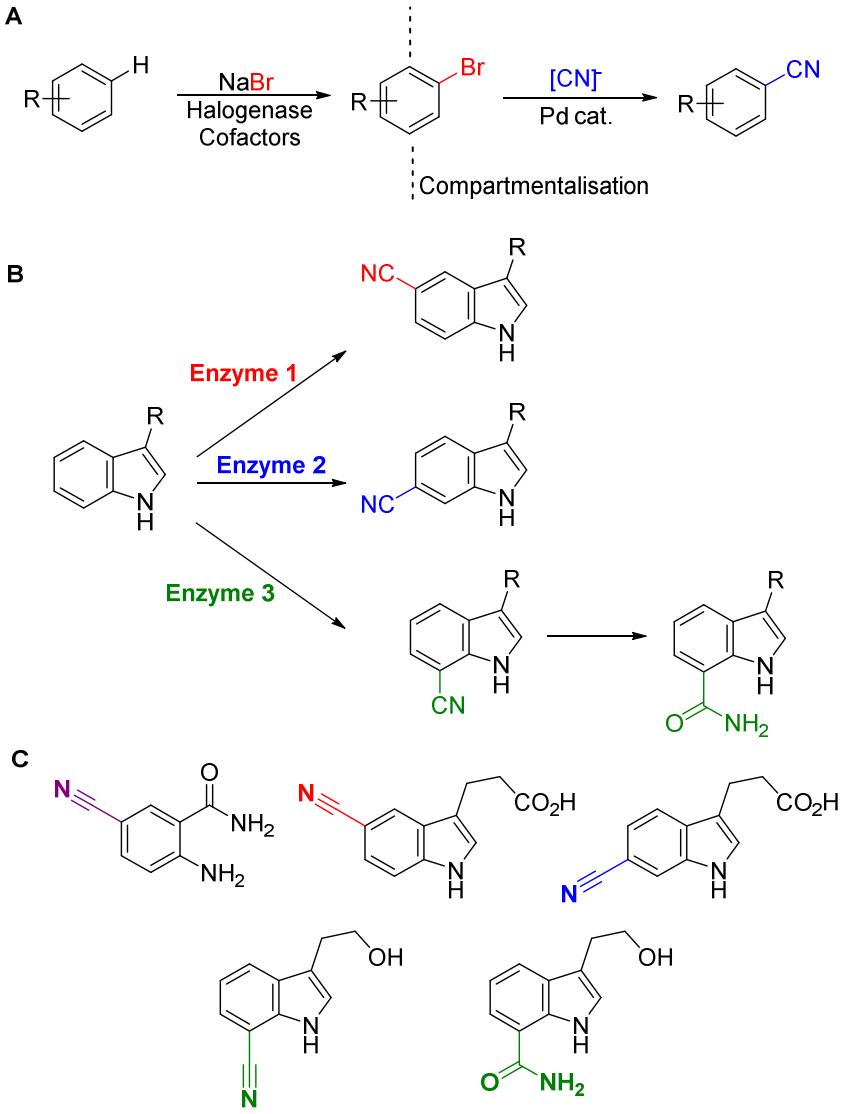
In addition to exploring the synthetic utility of combining FIHals with Pd-catalysed cross-couplings, work in our group is also focussed upon using such combinations to develop reactions which can be used to detect and monitor FIHal activity in a high-throughput manner. It is hoped that Pd-catalysed functionalisation of aryl halide FIHal products can induce an easily-detected change in spectrophotometric properties. It has been demonstrated that such reactions can allow the high-throughput screening of FIHal variants generated by random mutagenesis, affording enzymes with increased thermostability.[165] It is envisioned that selection of an appropriate coupling partner however may afford adducts with distinct spectrophotometric properties depending upon the position of functionalisation (**Scheme 82**). This would therefore allow the detection of FIHal variants of different regioselectivity directly by UV or fluorescence spectroscopy – without the need for laborious isolation and purification of reaction products – therefore allowing directed evolution of FIHal regioselectivity rather than activity or stability. Others have reported the engineering of FIHal regioselectivity using MS-based screening methods.[162] Although it was demonstrated that regioselectivity can be efficiently switched in this manner, the number of variants which can be processed is limited. Our group have previously used structure-guided mutagenesis to alter the

regioselectivity of FlHals – introducing mutations to alter the positioning of substrate relative to the active site lysine.[155,156] The utility and generality of structure-guided approaches are ultimately limited however, because of the requirement for structural information and because mutations which may alter regioselectivity are not always obvious. A high-throughput method for determining FlHal regioselectivity would allow evolution without structural information about the enzyme and could allow the introduction of mutations which would not be obvious based on structural analysis.



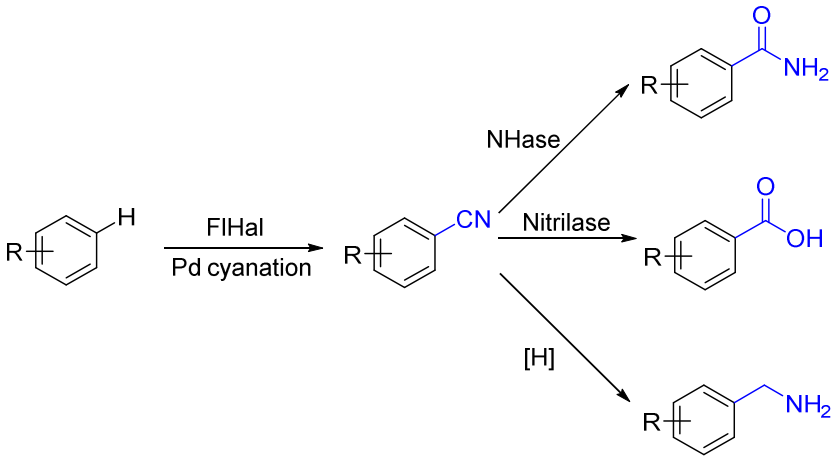
Scheme 82: Proposed use of integrated FlHal-SMC reactions for the screening of FlHal activity in a high-throughput manner. Ideally the fluorescent adducts A to C would possess distinct spectrophotometric properties to allow the determination of FlHal regioselectivity through fluorescence spectroscopy – thereby allowing FlHal regioselectivity to be engineered with random mutagenesis.

To broaden the synthetic utility of combining FlHals and Pd-catalysed reactions, similar integrated reactions were also assembled for the regioselective cyanation of aromatic substrates (**Scheme 83A**). Again by exploiting the regioselective and regio-divergent nature of the FlHals, one-pot cyanation of the indole C7, C6 and C5 positions was achieved by integration with Pd-catalysed cyanation reactions (**Scheme 83B** and **C**). Unlike the FlHal-SMC cascade, this could not be conducted using PDMS compartmentalisation as the hydrophobic palladacycle catalyst required for this chemistry would not be compatible. However, the utility of aryl nitrile compounds was demonstrated by the subsequent NHase-catalysed hydration of the resultant nitriles – affording a regioselective C-H amidation overall. The C-H cyanation and amidation reactions of indoles reported to date are currently very limited, with only the C2 and C3 positions accessible in either case (**Figure 45** and **Scheme 61**),[241,242] suggesting the integrated FlHal-cyanation cascade is of value.



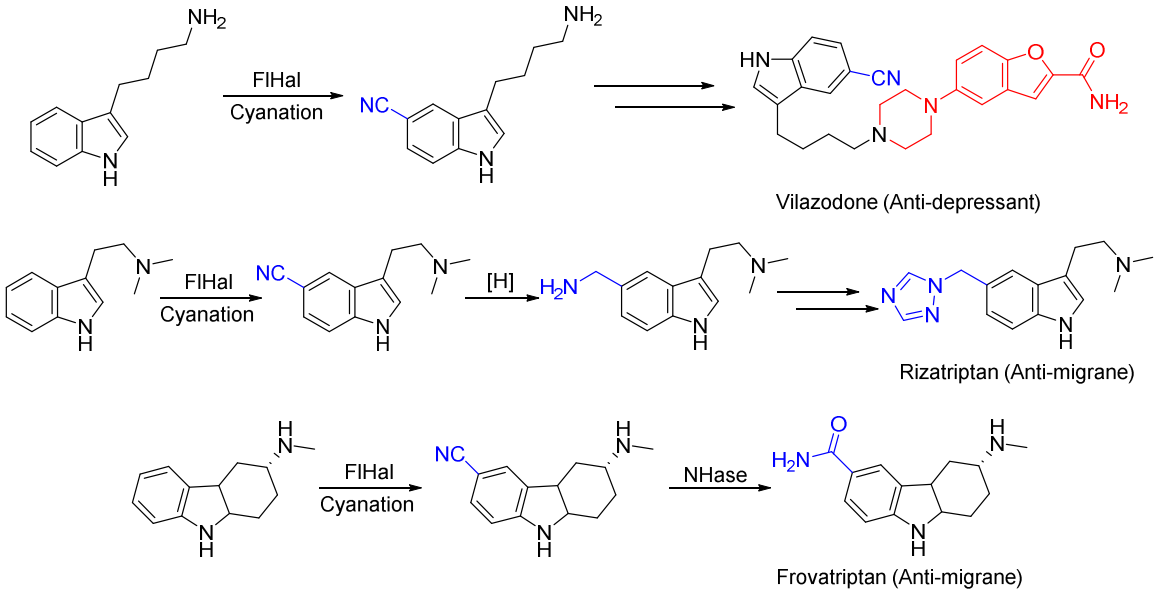
Scheme 83: (**A**) Overview of the regioselective C-H cyanation reaction through combination of the FlHal bromination with Pd-catalysed cyanation. (**B**) Regiodivergent C-H cyanation achieved through use of FlHals with different regioselectivity. (**C**) Cyanated and amidated products obtained using integrated FlHal-cyanation reactions.

Integration of the FlHal-cyanation reaction with the nitrile hydration was not particularly efficient, because of the fairly narrow substrate scope of the NHase employed and its intolerance to the organic co-solvent required for effective cyanation. However future work with different preparations of the NHase biocatalyst (such as an immobilised CLEA which may be more stable with respect to the organic co-solvent) or a Co-centred NHase (which are reported to more effectively hydrate larger aromatic nitriles) may provide more efficient amidation cascades. The versatile reactivity of the aryl nitrile group however allows significant further prospects for the integration of other transformations – allowing a wider array of functional groups to be regioselectivley installed (**Scheme 84**).



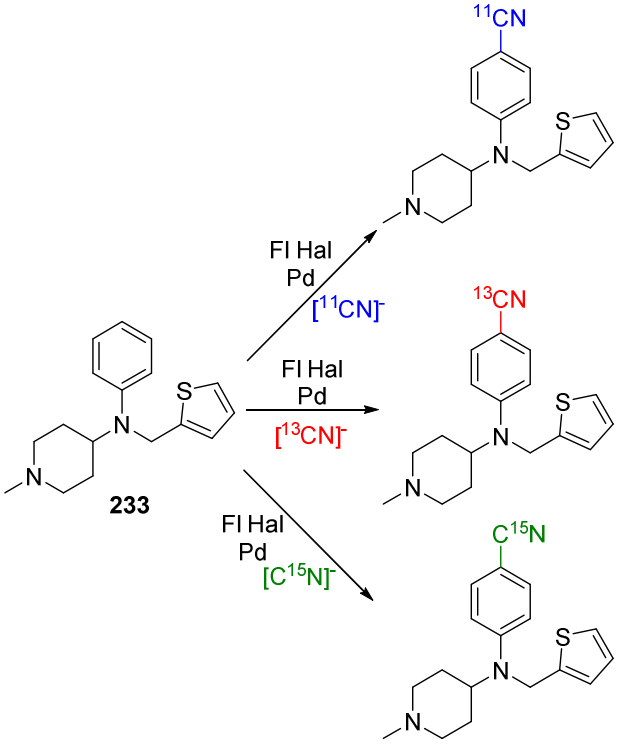
Scheme 84: Proposed reactions which could be integrated with the regioselective cyanation protocol to deliver greater diversity of products.

For example integration of a nitrilase enzyme with the FIHal-cyanation cascade could allow the direct installation of carboxylic acid moieties onto aromatic substrates. Through the identification of mild conditions for the reduction of the aryl nitrile, a formal C-H amino-methylation could also be developed (**Scheme 84**). Both the carboxylic acid and amino-methyl groups could also act as substrates for further bond-formation reactions, allowing significant molecular complexity to be introduced with the integrated FIHal-cyanation reaction at the centre of setting regioselectivity in these syntheses. **Scheme 85** details some proposed synthetic approaches to compounds of medicinal importance which may be synthesised using the integrated FIHal-cyanation reaction followed by further functionalisation. Preliminary work in our group has suggested that the successful realisation of these syntheses would require engineering of the FIHals to afford enzyme variants which are capable of halogenating the required position of the proposed starting materials (**Scheme 85**), although work is currently underway using the FIHal-SMC screening approach (**Scheme 82**) to achieve this.



Scheme 85: Proposed synthesis of pharmaceuticals through use of the integrated FIHal-cyanation reactions with subsequent transformations of the aryl nitrile.

Another interesting future prospect for these FlHal-Pd cascades would be their use for the late-stage modification of complex biologically-active molecules. Such methods are sought after since they can allow the efficient generation of natural product analogues. This has been achieved to some degree by the group of Lewis, using engineered FlHal variants for the late-stage halogenation of complex molecules, followed by isolation of the intermediate aryl halide and further Pd-catalysed functionalisations.[211] This requirement to isolate the intermediate aryl bromide somewhat limits the utility of this process from the green aspect, however. Use of the methods discussed in this thesis with the engineered FlHals reported by Lewis could allow more efficient late-stage modification of complex biologically-active molecules. Of particular interest would be the cyanation of large bioactive molecules in this way, since use of a $^{-}$CN source enriched with non-natural C or N isotopes would allow the facile labelling of such molecules (**Scheme 86**) – commonly used to determine the interaction of such compounds with biological systems.



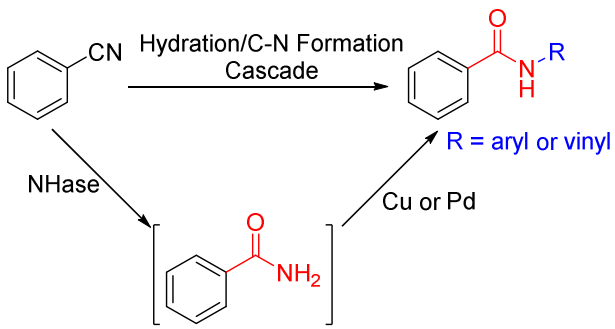
Scheme 86: Proposed use of the FlHal-cyanation cascade for the labelling of complex biological molecules with carbon and nitrogen isotopes which may allow facile detection of these compounds and metabolites in biological systems using the antihistamine thenalidine (**233**) as an example.[211]

Such integrated cyanation reactions could possibly also be employed for the high-throughput screening of FlHal activity, in a similar vein to the FlHal-SMC screening approach proposed above. Since the C≡N group is vibrationally distinct from most functional groups endogenous to biological systems, vibrational spectroscopy could potentially be used to easily detect the products from integrated FlHal-cyanation reactions without the need for time-consuming chromatography processes. Therefore the cyanation of FlHal biotransformations could be used to allow the detection of FlHal activity in an expedient manner, and therefore allow the high-throughput screening of FlHal variants derived from directed evolution – potentially allowing the engineering of

FlHals towards enzymes more amenable to application as biocatalysts. Using the cyanation reaction is potentially a more general approach than the FlHal-SMC screen proposed above (**Scheme 82**) since only detection of the CN moiety is required, rather than a change in the electronic properties of an extended part of the molecule (needed to induce a change in fluorescence properties).

As a means of improving the overall efficiency of combining FlHals with Pd-catalysed cross-couplings, it may be desirable to develop a biocatalytic iodination reaction. Many of the tryptophan FlHals studied to date have been found to be capable only of chlorination and bromination, with the presence of iodide inhibiting halogenation activity of these enzymes.[76,78] The recent discovery of Bmp5, a phenolic FlHal capable of bromination and iodination but not chlorination, however suggests that a biocatalytic iodination reaction may be achievable – either by development of Bmp5 as a biocatalyst or by using the structural features of Bmp5 thought to allow iodination to guide development of other FlHals. Whilst this ambitious goal is likely to require significant work, the more facile nature of Pd^0 oxidative addition across C-I bonds compared to C-Br bonds could allow the assembly of FlHal-Pd cascades which do not require significant heating – thereby allowing truly concurrent cascade reactions with a lower energy demand and potentially lower loadings of Pd catalysts.

In addition to the methods combining FlHals and Pd catalysis, a number of attempts were made in this thesis to integrate the biocatalytic hydration of nitriles with metal-catalysed couplings for the arylation and vinylation of the resultant amides (**Scheme 87**). Although a robust NHase-catalysed nitrile hydration was afforded, the integration of this reaction with various metal-catalysed transformations was unsuccessful. It has been reported numerous times that the low acidity of the amide N-H and poor nucleophilicity of the amide N make these compounds poor substrates for metal-catalysed reactions since they necessitate coordination of the amide N to metal catalysts.[291] Some conditions were found which allowed successful derivatisation of primary amides, either using a high concentration of amide substrate, strong bases or electrophilic coupling partners, however integration with the NHases was limited either by the concentration of intermediate amide afforded by the NHase or deactivation of the strong bases by residual water. A number of attempts were made to conduct the nitrile hydration in microaqueous environments, in an attempt to reduce the deactivating effect of water, although this was unsuccessful.



Scheme 87: Attempted integration of NHase-catalysed amide formation with metal-catalysed C-N bond formation reactions.

Whilst future work on improving the productivity of the NHase biocatalysts may afford effective integrated reactions (**Scheme 87**), the value of this transformation must first be carefully considered. One major goal of this combination was to allow the synthesis of secondary amides and enamides without the need to use stoichiometric activating agents in order to improve the atom economy of such transformations. The use of stoichiometric organometallic coupling partners (i.e. boronic acids in the CLE reaction) or iodonium salts (whereby an aryl iodide is jettisoned in each coupling) is therefore illogical with this goal in mind. The value of this transformation therefore lies in the chemo-selectivity, and potential stereoselectivity, of the NHase.

Although the combination of NHases with transition metal-catalysed reactions was not entirely successful, this thesis demonstrates the utility of successful integrated catalysis approaches in synthesis. Most striking are the transformations afforded by the combination of FIHals with palladium-catalysed cross-coupling reactions. The ability of the FIHals to selectively install a reactive bromide moiety, when combined with Pd-catalysed reactions, allowed regioselective C-C bond formations at different positions of similar aromatic substrates. The versatility of palladium-catalysed reactions allowed a plethora of transformations to be conducted, with the selectivity of the FIHals key to regio-control. Such control is particularly important for the functionalisation of aromatic substrates, since traditional methods often offer poor regioselectivity and afford mixtures of isomeric products. To allow these transformations to be conducted without the need to purify or isolate intermediates, compartmentalisation of the biocatalysts and chemocatalysts was required to prevent their mutual deactivation. Despite the need to overcome these compatability issues, the transformations afforded demonstrate the role which biocatalysis can have in imparting selectivity into synthesis – an important aspect to achieving economical and environmentally benign chemical synthesis.

7. Experimental

7.1 General Reagents, Equipment and Methods

All chemicals were purchased from Sigma-Aldrich, Arcos Organics, Fluorochem, Fischer Scientific UK or Alfa Aesar and used without further purification unless otherwise stated. Anhydrous solvents were either purchased in Sure-Seal® bottles from Sigma-Aldrich or dried over 3 Å molecular sieves for a minimum of 24 hrs. 3 Å and 4 Å molecular sieves were dried by heating with $MgSO_4$ at 300 ^oC for 12 hrs. Petroleum ether (PE) refers to the 40 ^oC – 60 ^oC fraction.

^1H and ^{13}C NMR spectra were recorded on either a Bruker Avance I 400 MHz or Bruker Avance III 800 MHz spectrometer and are reported relative to residual solvent.^{19}F NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer. IR spectra were recorded on a Bruker Alpha-P IR spectrometer as a neat oil or powder. Mass Spectra were recorded on an Agilent Q-TOF LC-MS or ThermoFischer LCQ Orbitrap XL. Optical rotations were measured at 589 nm using an Optical Activity AA1000 polarimeter. Melting points are unadjusted and were recorded on a Gallenkamp MPD350 melting point apparatus.

For the reporting of infrared data, the following abbreviations are used to describe peak shape and intensity; br. (broad), w (weak), m (medium), s (strong). The multiplicity of peaks in NMR spectra are described using the following abbreviations; s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), m (multiplet), app. (apparent).

7.1.1 HPLC Methods

All analytical HPLC methods were run on an Agilent 1266 Infinity® HPLC system. All preparative HPLC methods were run on a Varian ProStar® HPLC system.

7.1.1.1 Analytical HPLC Method 1

5 µL of supernatant was injected onto a Security Guard Ultra® trapping column (2 µm, 2 mm x 2.1 mm) prior to separation using a Phenomenex Kinetex® C18 analytical column (5 µm packing, 100 mm x 4.6 mm). Gradient starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 2 min prior to development to 75 % H_2O/ACN over 5 min. 95 % H_2O/ACN was then held for 2 min prior to equilibration back to 5 % H_2O/ACN. Flow rate was maintained at 1 mL min^{-1} and column temperature at 30 ^oC throughout. UV absorbance was detected at 254, 280, 310, 223 and 233 nm throughout.

7.1.1.2 Analytical HPLC Method 2

5 µL of supernatant was injected onto a Security Guard Ultra® trapping column (2 µm, 2 mm x 2.1 mm) prior to separation using a Phenomenex Kinetex® C18 analytical column (5 µm packing, 100 mm x 4.6 mm). Gradient starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 2 min prior to development to 95 % H_2O/ACN over 10 min. 95 % H_2O/ACN was then held for 2 min prior to equilibration back to 5 % H_2O/ACN. Flow rate was maintained at 1 mL min^{-1} and column

temperature at 30 ^{o}C throughout. UV absorbance was detected at 254, 280, 310, 223 and 233 nm throughout.

7.1.1.3 Analytical HPLC Method 3

5 µL of supernatant was injected onto a Security Guard Ultra® trapping column (2 µm, 2 mm x 2.1 mm) prior to separation using a Phenomenex Kinetex® C18 analytical column (5 µm packing, 150 mm x 4.6 mm). Gradient starting conditions of 30 % H_2O/ACN (+0.05 % TFA) were held for 7 min followed by washing at 95 % H_2O/ACN for 2 min prior to equilibration back to 30 % H_2O/ACN. Flow rate was maintained at 1 mL min^{-1} and column temperature at 30 ^{o}C throughout. UV absorbance was detected at 254, 280, 310, 223 and 233 nm throughout.

7.1.1.4 Analytical HPLC Method 4

5 µL of supernatant was injected onto a Security Guard Ultra® trapping column (2 µm, 2 mm x 2.1 mm) prior to separation using a Phenomenex Kinetex® C18 analytical column (5 µm packing, 150 mm x 4.6 mm). Gradient starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 2 min prior to development to 40 % H_2O/ACN over 1 min. 40 % H_2O/ACN was then held for 7 min prior to development to 95 % H_2O/ACN over 1 min. 95 % H_2O/ACN was then held for 2 min prior to equilibration back to 5 % H_2O/ACN. Flow rate was maintained at 1 mL min^{-1} and column temperature at 30 ^{o}C throughout. UV absorbance was detected at 254, 280, 310, 223 and 233 nm throughout.

7.1.1.5 Analytical HPLC Method 5

5 µL of supernatant was injected onto a Security Guard Ultra® trapping column (2 µm, 2 mm x 2.1 mm) prior to separation using a Phenomenex Kinetex® C18 analytical column (5 µm packing, 150 mm x 4.6 mm). Gradient starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 10 min prior to development to 95 % H_2O/ACN over 2 min. 95 % H_2O/ACN was then held for 2 min prior to equilibration back to 5 % H_2O/ACN. Flow rate was maintained at 1 mL min^{-1} and column temperature at 30 ^{o}C throughout. UV absorbance was detected at 254, 280, 310, 223, 233 and 225 nm throughout.

7.1.1.6 Semi-Preparative HPLC Method 1

400 µL of solution containing crude was injected onto a Phenomenex Gemini® C18 HPLC column (5 µm packing, 250 mm x 10 mm). Starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 3 min prior to development to 95 % H_2O/ACN over 23 min. 95 % H_2O/ACN was then held for 5 min to equilibration back to 5 % H_2O/ACN over 3 min. Flow rate was maintained at 5 mL min^{-1} throughout. UV absorbance was detected at 254 and 280 nm throughout.

7.1.1.7 Semi-Preparative HPLC Method 2

400 µL of solution containing crude was injected onto a Phenomenex Gemini® C18 HPLC column (5 µm packing, 250 mm x 10 mm). Starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 3 min prior to development to 35 % H_2O/ACN over 23 min. 95 % H_2O/ACN was then held for 5

min prior to equilibration back to 5 % H_2O/ACN over 3 min. Flow rate was maintained at 5 mL min^{-1} throughout. UV absorbance was detected at 280 and 325 nm throughout.

7.1.1.8 Semi-Preparative HPLC Method 3

400 µL of solution containing crude was injected onto a Phenomenex Kinetex® C18 HPLC column (5 µm packing, 150 mm x 10 mm). Starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 3 min prior to development to 95 % H_2O/ACN over 17 min. 95 % H_2O/ACN was then held for 5 min to equilibration back to 5 % H_2O/ACN over 2 min. Flow rate was maintained at 5 mL min^{-1} throughout. UV absorbance was detected at 254 and 280 nm throughout.

7.1.1.9 Semi-Preparative HPLC Method 4

400 µL of solution containing crude was injected onto a Phenomenex Kinetex® C18 HPLC column (5 µm packing, 150 mm x 10 mm). Starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 3 min prior to development to 60 % H_2O/ACN over 6 min and held for 5 min before development to 95 % H2O/ACN over 3 min. 95 % H_2O/ACN was then held for 3 min prior to equilibration back to 5 % H2O/ACN over 2 min. Flow rate was maintained at 5 mL min^{-1} throughout. UV absorbance was detected at 254 and 280 nm throughout.

7.1.1.10 Semi-Preparative HPLC Method 5

400 µL of solution containing crude was injected onto a Phenomenex Kinetex® C18 HPLC column (5 µm packing, 150 mm x 10 mm). Starting conditions of 30 % H_2O/ACN (+0.05 % TFA) were held for 20 min prior to development to 95 % H_2O/ACN over 5 min. 95 % H_2O/ACN was then held for 5 min before equilibration back to 5 % H_2O/ACN over 3 min. Flow rate was maintained at 5 mL min^{-1}. UV absorbance was detected at 254 and 280 nm throughout.

7.1.1.11 Semi-Preparative HPLC Method 6

400 µL of solution containing crude was injected onto a Phenomenex Kinetex® C18 HPLC column (5 µm packing, 150 mm x 10 mm). Starting conditions of 10 % H_2O/ACN (+0.05 % TFA) were held for 5 min prior to development to 40 % H_2O/ACN over 2 min. 40 % H_2O/ACN was then held for 15 min prior to development to 95 % H_2O/ACN over 2 min. 95 % H_2O/ACN was then held for 5 min prior to equilibration back to 5 % H_2O/ACN over 2 min. Flow rate was maintained at 5 mL min^{-1} throughout. UV absorbance was detected at 280 and 310 nm throughout.

7.1.1.12 General Preparative C18 HPLC Method

400 µL of solution containing crude was injected onto a Phenomenex Luna® C18 HPLC column (10 µm packing, 250 mm x 21.2 mm). Starting conditions of 5 % H_2O/ACN (+0.05 % TFA) were held for 3 min prior to development to 95 % H_2O/ACN over 23 min. 95 % H_2O/ACN was then held for 3 min prior to equilibration back to 5 % H_2O/ACN over 3 min. Flow rate was maintained at 10 mL min^{-1} throughout. UV absorbance was detected at 254 and 280 nm throughout.

7.1.1.13 General Preparative HILIC HPLC Method

400 µL of solution containing crude was injected onto a Phenomenex Luna® C18 HILIC column (5 µm packing, 250 mm x 21.2 mm). Starting conditions of 10 % ACN/NH_4HCO_2 buffer (5 mM in H_2O, pH 3.3) were held for 10 min prior to development to 50 % ACN/NH_4HCO_2 buffer over 3 min. The column was then equilibrated back to 10 % ACN/NH_4HCO_2 buffer over 2 min. Flow rate was maintained at 15 mL min^{-1}. UV absorbance was detected at 254 and 280 nm throughout.

7.1.2 General LC-MS Method

3 µL of supernatant was injected onto a Phenomenex Kinetex® C18 UHPLC column (2.6 µ packing, 100 mm x 2.1 mm). Starting conditions of 5 % H_2O/ACN (+0.1 % FA) were held for 2 min prior to development to 95 % H_2O/ACN over 5 min. 95 % H_2O/ACN was then held for 3 min prior to equilibration back to 5 % H_2O/ACN. Eluent was injected directly into a ThermoFischer LCQ Orbitrap XL and mass spectra recorded from *m/z* 100 to 1000 throughout. Flow rate was maintained at 0.3 mL min^{-1} and column temperature at 30 ^oC throughout.

7.1.3 General GC-MS Method

GC-MS analysis was performed on an Agilent 7890B GC equipped with an Agilent Technologies 5977A MSD using a DB-WAX column (0.25 µm film, 30 m x 0.32 mm).

The injector temperature was set at 240 ^oC and 1 µL of extract injected with a 20:1 split ratio. The column oven was initially set at 35 ^oC and held for 5 min. Temperature was then ramped to 110 ^oC at 20 ^oC min^{-1} then held for 5 min before ramp to 220 ^oC at 20 ^oC min^{-1}. After holding at 220 ^oC for 5 min, the oven was cooled back to 35 ^oC. Ion source temperature of the mass spectrometer was set to 230 ^oC and spectra recorded from *m/z* 50 to 250 throughout after a 3.5 min delay. Helium was used as carrier gas and a flow rate of 2 mL min^{-1} maintained throughout.

7.1.4 Competent Cell Preparation

7.1.4.1 Competent Cell Preparation Buffer

25 mM MES and 60 mM $CaCl_2$ in H_2O was adjusted to pH 5.8 with KOH. $MgCl_2$ and $MnCl_2$ were then added to final concentrations of 5 mM each and the solution sterilised by autoclaving at 121 ^oC for 20 min.

7.1.4.2 Competent Cell Storage Buffer

25 mM MES and 60 mM $CaCl_2$ in H_2O containing 10 % glycerol was adjusted to pH 5.8 with KOH. $MgCl_2$ and $MnCl_2$ were then added to final concentrations of 5 mM each and the solution sterilised by autoclaving at 121 ^oC for 20 min.

7.1.4.3 Method for Preparation of Competent *E. coli* cells

A 10 mL overnight culture of the appropriate *E. coli* strain was diluted 1:50 into 100 mL of fresh LB medium containing any antibiotics required for selection. Once OD_{600} = 0.4, the culture was cooled on ice for 15 min prior to harvesting of the cells by centrifugation (1500 rpm, 10 min, 4 ^oC). The pellet was then resuspended in 100 mL of competent cell preparation buffer and cooled on ice for

1 hr. The cells were then harvested by centrifugation (1500 rpm, 10 min, 4 ^{o}C) and resuspended into 10 mL competent cell storage buffer. Aliquots of 100 µL were then stored at -80 ^{o}C prior to use.

7.1.5 Transformation of *E. coli*

A 100 µL aliquot of competent cells prepared above was defrosted on ice prior to addition of purified plasmid DNA (2 µL). After incubation on ice for 1 hr, cells were briefly heated at 42 ^{o}C for 45 sec before cooling on ice for 5 min. LB broth (900 µL) was then added and the culture incubated at 37 ^{o}C for 1 hr. Cells were then harvested by centrifugation (4000 rpm, 5 min) and 500 µL of supernatant removed. The cells were then resuspended in the remaining supernatant and 200 µL spread onto LBA plates supplemented with the appropriate antibiotics for selection until nearly dry. The resultant plates were then incubated at 37 ^{o}C overnight.

7.1.6 SDS-PAGE Electrophoresis

7.1.6.1 X10 SDS Running Buffer

30 g Tris-base, 144 g glycine, 10 g SDS made up to 1 L with distilled H_2O. Buffer was diluted 1:10 with distilled H_2O prior to use.

7.1.6.2 X4 Loading Dye

0.5 % β-mercaptoethanol, 0.04 % bromophenol blue, 8 % SDS and 240 mM Tris HCl (pH 6.8) in distilled H_2O containing 40 % v/v glycerol.

7.1.6.3 SDS-PAGE Electrophoresis Method

To 15 µL of protein-containing sample was added 5 µL of X4 loading dye. Samples were heated at 95 ^{o}C for 10 min prior to centrifugation (13000 rpm, 2 min) and loading onto the appropriate SDS-PAGE gels. 8-16 % gradient SDS gels were purchased from Amersham (ECL 8-16%) and used in accordance with the manufacturer's instructions. 16 % SDS gels were prepared according to standard protocols and run at 140 V until complete. After running, gels were stained with 25 mL Instant Blue^{TM} (Expedeon) overnight followed by destaining with distilled H_2O.

7.1.7 Protein Purification Buffers

7.1.7.1 10 mM imidazole KPi buffer

50 mM potassium phosphate buffer, 500 mM NaCl and 10 mM imidazole in distilled H_2O. pH was adjusted to 7.4 prior to use and the solution chilled at 4 ^{o}C.

7.1.7.2 80 mM imidazole KPi buffer

50 mM potassium phosphate buffer, 500 mM NaCl and 80 mM imidazole in distilled H_2O. pH was adjusted to 7.4 prior to use and the solution chilled at 4 ^{o}C.

7.1.7.3 300 mM imidazole KPi buffer

50 mM potassium phosphate buffer, 500 mM NaCl and 300 mM imidazole in distilled H_2O. pH was adjusted to 7.4 prior to use and the solution chilled at 4 ^{o}C.

7.1.7.4 10 mM Tris buffer

50 mM Tris HCl buffer, 100 mM NaCl, 40 mM butanoic acid and 10 mM imidazole in distilled H_2O. pH was adjusted to 7.5 prior to use and the solution chilled at 4 ^oC.

7.1.7.5 30 mM Tris buffer

50 mM Tris HCl buffer, 100 mM NaCl, 40 mM butanoic acid and 30 mM imidazole in distilled H_2O. pH was adjusted to 7.5 prior to use and the solution chilled at 4 ^oC.

7.1.7.6 300 mM Tris buffer

50 mM Tris HCl buffer, 100 mM NaCl, 40 mM butanoic acid and 300 mM imidazole in distilled H_2O. pH was adjusted to 7.5 prior to use and the solution chilled at 4 ^oC.

7.1.8 Growth Media

All media components were purchased from either Formedium or Fischer Chemicals and prepared as described below.

7.1.8.1 LB Broth

10 g tryptone, 5 g yeast extract and 10 g NaCl made up to 1 L with distilled water. The media was sterilised by autoclaving 121 ^oC for 20 min.

7.1.8.2 LB Agar

10 g tryptone, 5 g yeast extract, 10 g NaCl and 15 g agar made up to 1 L with distilled water. The media was sterilised by autoclaving 121 ^oC for 20 min.

7.1.8.3 Terrific Broth (TB)

53 g tryptone, 10 g yeast extract and 18 mL glycerol made up to 4 L with distilled water. The media was sterilised by autoclaving at 121 ^oC for 20 min. X10 TB salts were added in a 1:10 ratio prior to inoculation.

7.1.8.4 X10 TB Salts

125.4 g K_2HPO_4, 23.1 g KH_2PO_4 made up to 1 L with distilled water. pH was adjusted to 7.2 and the solution filter sterilised prior to use.

7.1.8.5 2YT Broth

64 g tryptone, 40 g yeast extract and 20 g NaCl made up to 4 L with distilled water. The media was sterilised by autoclaving at 121 ^oC for 20 min.

7.2 Chapter 2 Experimental

All compounds detailed in this section have been characterised as part of an accepted article in a peer-reviewed journal. Copies of the original spectral data for these compounds is therefore available as part of the Supporting Information of this publication. The relevant publication is cited where appropriate.

7.2.1 Enzyme Expression and Purification

The gene encoding the thermophilic tryptophan halogenase from *Streptomyces violaceusniger* SPC6 (Th_Hal) was synthesised by GenScript (Piscataway, USA) and cloned into pET28b (+) by Binuraj Menon as previously described.[175] The gene encoding the thermophilic flavin reductase from *Bacillus subilis* WU-S2B (Th_Fre) was synthesised by GenScript (Piscataway, USA) and cloned into a pACYCDUET-1 vector by Binuraj Menon as previously described.[175] Cloning and expression of the mesophilic FlHals used for comparison is described in the next experimental chapter. The composition of each protein purification buffer is detailed above (7.1.7).

7.2.1.1 Thermophilic Tryptophan Halogenase (Th_Hal) Expression

The construct encoding the Th_Hal gene was briefly transformed into *E. coli* BL21 (DE3) using kanamycin (50 µg/mL) for selection. Single colonies were then picked and used to inoculate 10 mL of LB medium containing kanamycin (50 µg/mL) prior to incubation with shaking at 37 ^{o}C until OD_{600} = 0.6. This culture was then diluted 1:100 into 100 mL of fresh LB containing kanamycin (50 µg/mL) and incubated at 37 ^{o}C overnight. The culture was then further diluted 1:100 into 10 x 400 mL 2YT medium containing kanamycin (50 µg/mL) and incubated at 30 ^{o}C with shaking at 180 rpm until OD_{600} = 0.6. Protein expression was then induced by addition of IPTG (0.1 mM final) and incubation at 20 ^{o}C with shaking at 180 rpm continued for 14 hrs. The cells were then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C) and stored at –20 ^{o}C before use.

7.2.1.2 Thermophilic Flavin Reductase (Th_Fre) Expression

The construct encoding Th_Fre was briefly transformed into *E. coli* BL21 (DE3) using chloramphenicol (35 µg/mL) for selection. Single colonies were then picked and used to inoculate 10 mL of LB medium containing chloramphenicol (35 µg/mL) prior to incubation with shaking at 37 ^{o}C until OD_{600} = 0.6. This culture was then diluted 1:100 into 100 mL of fresh LB containing chloramphenicol (35 µg/mL) and incubated at 37 ^{o}C overnight. The culture was then further diluted 1:100 into 10 x 400 mL 2YT medium containing chloramphenicol (35 µg/mL) and incubated at 30 ^{o}C with shaking at 180 rpm until OD_{600} = 0.6. Protein expression was then induced by addition of IPTG (0.1 mM final) and incubation at 18 ^{o}C with shaking at 180 rpm continued for 14 hrs. The cells were then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C) and purified immediately.

7.2.1.3 Protein Purification

Pelleted cells were re-suspended in 10 mM imidazole KPi buffer (ca 60 mL total volume or 4 L culture). Cells were then lysed by sonication (10 min, 50 % pulse, 70 % power, 700 W) and the lysate clarified by centrifugation (18000 rpm, 45 min, 4 ^{o}C). After filtration through a 0.45 µm cellulose syringe filter, lysate was loaded onto Ni-NTA resin (Quaigen), which had been pre-

equilibrated with 10 mM imidazole KPi buffer, under gravity. After washing of the resin with 3 column volumes (CV) of 80 mM imidazole KPi buffer, protein was eluted using 5 CV of 300 mM imdazole KPi buffer. Eluent was then exchanged into 100 mM KPi buffer (pH 7.2) using either spin concentration or dialysis. After concentration to ca 5 mL, protein concentration was determined using the A280 feature of a ThermoScientific 2000 Nanodrop and purity assessed by SDS-PAGE.

7.2.2 Halogenase Biotransformations

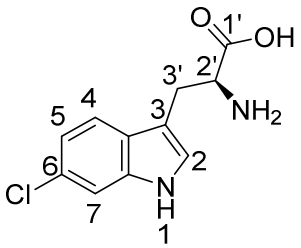
7.2.2.1 General Method for Analytical Scale Th_Hal Biotransformations

To a solution of substrate (0.5 mM), Th_Hal (2.5 µM), Th_Fre (1 µM), FAD (1 µM) and $MgCl_2$ (10 mM) in 10 mM potassium phosphate buffer (pH 7.2) was added NADH (2.5 mM) to a total volume of 200 µL. After incubation at either 30 ^{o}C or 45 ^{o}C with shaking at 800 rpm for 30 min, protein was precipitated by heating (95 ^{o}C, 10 min) prior to removal by centrifugation (13000 rpm, 15 min) and analysis of the supernatant by analytical HPLC method 1.

7.2.2.2 General Method for Preparative Scale Th_Hal Biotransformations

Substrate (2.0 mM) was dissolved in 10 mM potassium phosphate buffer (pH 7.2) with sonication prior to addition of $MgCl_2$ (10 mM), FAD (1 µM), Th_Hal (50 µM), Th_Fre (5 µM) and NADH to a total volume of 40 mL in a loosely-capped falcon tube. After incubation at room temperature for 24 hrs, protein was precipitated by heating (70 ^{o}C, 15 min) and removed by centrifugation (10000 rpm, 30 min) followed by filtration through a 0.2 µM cellulose syringe filter. The filtrate was then applied to a C18 Bond Elute® (Agilent, 20 g mesh) which had been activated with 3 column lengths (CL) of methanol and equilibrated with 3 CL H_2O . After washing with water (3 CL), organics were eluted with MeOH (3 CL) and solvent removed *in vacuo* prior to purification.

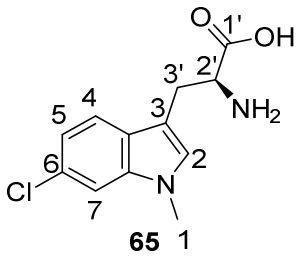
7.2.2.3 6-chlorotryptophan (57)[156,175]



57

Prepared from L-tryptophan (**31**) using two 40 mL biotransformations according to the general method above (7.2.2.2) which were combined prior to purification. Purification by preparative C18 HPLC afforded the title compound as an off-white solid (12.4 mg, 62.5 %); ^{1}H NMR (400 MHz, MeOD) δ 8.48 (1H, s, H1), 7.63 (1H d, J = 8.5 Hz, H4), 7.33 (1H d, J = 1.8 Hz, H7), 7.18 (1H, s, H2), 6.99 (1H dd, J = 8.5, 1.8 Hz, H5), 3.80 (1H, m, H2'), 3.43 (1H, m, H3'), 3.12 (1H, m, H3'); ^{13}C NMR (101 MHz, MeOD) δ 174.3, 138.7, 128.7, 127.2, 126.2, 120.6, 120.5, 112.2, 110.0, 56.6, 28.2; *m/z* (ESI) 237.1 ([^{35}M-H]$^{-}$, 75 %), 239.1 ([^{37}M-H]$^{-}$, 25); HRMS *m/z* calcd for $C_{11}H_{10}N_2O_2{}^{35}Cl$ 237.0431, found: 237.0417.

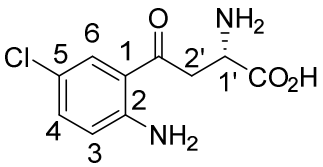
7.2.2.4 1-methyl-6-chlorotryptophan (65)[156,175]



65

Prepared from 1-methyl-L-tryptophan (**58**) using two 40 mL biotransformations according to the general method above (7.2.2.2) which were combined prior to purification. Purification by preparative C18 HPLC afforded the title compound as an off-white solid (4.8 mg, 12 %); ^{1}H NMR (400 MHz, D_2O) δ 7.60 (1H, d, *J* = 8.4 Hz, H4), 7.51 (1H, s, H7), 7.16 (1H, s, H2), 7.14 (1H, d, *J* = 8.6 Hz, H5), 4.06 (1H, s, H1'), 3.72 (3H, s, H1), 3.34 (2H, m, H2'); *m/z* (ESI) 251.1 ([^{35}M-H]$^-$, 75 %), 253.1 ([^{37}M-H]$^-$, 25), 363.1 ([^{35}M+TFA-H]$^-$, 75), 365.1 ([^{37}M+TFA-H]$^-$, 25); HRMS *m/z* calcd for $C_{12}H_{14}N_2O_2{}^{35}Cl$ 253.0738, found: 253.0744.

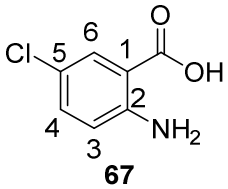
7.2.2.5 5-chloro-L-kynurenine (66)[155,175]



60

Prepared from L-kynurenine (**60**) using two 40 mL biotransformations according to the general method above (7.2.2.2) which were combined prior to purification. Purification by preparative C18 HPLC afforded the title compound as a yellow solid (7.9 mg, 21 %); ^{1}H NMR (400 MHz, MeOD) δ 7.68 (1H, d, *J* = 2.4 Hz, H6), 7.22 (1H, dd, *J* = 9.0, 2.4 Hz, H4), 6.75 (1H, d, *J* = 9.0 Hz, H3), 4.32 (1H, dd, *J* = 4.6, 4.3 Hz, H1'), 3.65 (2H, m, H2'); *m/z* (ESI) 241.1 ([^{35}M-H]$^-$, 75 %), 243.1 ([^{37}M-H]$^+$, 25); HRMS *m/z* calcd for $C_{10}H_{12}{}^{35}ClN_2O_3$ 243.0531 found: 243.0536.

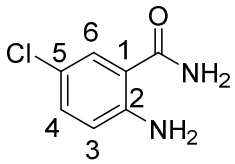
7.2.2.6 5-chloroanthranilic acid (67)[155,175]



67

Prepared from anthranilic acid (**61**) using two 40 mL biotransformations according to the general method above (7.2.2.2) which were combined prior to purification. Purification by preparative C18 HPLC afforded the title compound as a white solid (2.2 mg, 8 %); ^{1}H NMR (400 MHz, D_2O) δ 7.89 (1H, d, *J* = 2.4 Hz, H6), 7.42 (1H, dd, *J* = 8.6, 2.4 Hz, H4), 7.01 (1H, d, *J* = 8.6 Hz, H3); *m/z* (ESI) 125.9 ([^{35}M-CO_2]$^-$, 75 %), 127.9 ([^{37}M-CO_2]$^-$, 25), 169.9 ([^{35}M-H]$^-$, 75), 171.9 ([^{37}M-H]$^-$, 25); HRMS *m/z* calcd for $C_7H_5{}^{35}ClNO_2$ 170.0014 found: 170.0025.

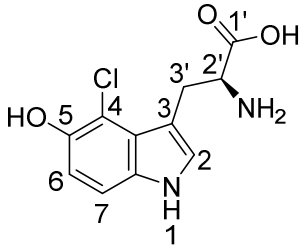
7.2.2.7 5-chloroanthranilamide (68)[155,175]



68

Prepared from anthranilamide (**62**) using two 40 mL biotransformations according to the general method above (7.2.2.2) which were combined prior to purification. Purification by preparative C18 HPLC afforded the title compound as a white solid (5.1 mg, 19 %); ^{1}H NMR (400 MHz, MeOD) δ 7.53 (1H, d, *J* = 2.4 Hz, H6), 7.15 (1H, dd, *J* = 8.8, 2.4 Hz, H4), 6.72 (1H, d, *J* = 8.8 Hz, H3); *m/z* (ESI) 193.1 ([^{35}M+Na]$^{+}$, 75 %), 195.1 ([^{37}M+Na]$^{+}$, 25); HRMS *m/z* calcd for $C_7H_8{}^{35}ClN_2O$ 171.0320 found: 171.0323.

7.2.2.8 4-chloro-5-hydroxy-L-tryptophan (69)[175,292]



69

Prepared from 5-hydroxy-L-tryptophan (**59**) using a single 40 mL biotransformation according to the general method above (7.2.2.2). After removal of protein by heating (70 ^{o}C, 15 min) and centrifugation (10000 rpm, 30 min), the reaction was lyophilised before being re-dissolved in H_2O (3 mL) and purified by preparative HILIC HPLC to afford the title compound as a brown solid (3.2 mg, 16 %); ^{1}H NMR (400 MHz, MeOD) δ 7.18 (1H, d, *J* = 8.7 Hz, H6), 7.15 (1H, s, H2), 6.83 (1H, d, *J* = 8.7 Hz, H7), 4.28 (1H, dd, *J* = 10.6, 4.5 Hz, H1'), 3.93 (1H, m, H2'), 3.13 (1H, m, H2'); *m/z* (ESI) 255.1 ([^{35}M+H]$^{+}$, 75 %), 257.1 ([^{37}M+H]$^{+}$, 25); HRMS *m/z* calcd for $C_{11}H_{12}{}^{35}ClN_2O_3$ 255.0531 found: 255.0537.

7.2.3 Th_Hal Docking Methods

The resolved apo crystal structure of Th_Hal (PDB 5LV9) was used as receptor for substrate docking using AutoDock 4.2.[172] Polar hydrogen atoms and Gasteiger charges were added to the protein structure using AutoDock Tools to create the PDBQT file. The substrate tryptophan (**31**) was MM2 minimised with Chem3D 16.0 and then set with rotatable bonds around the amino acid side-chain and converted to a PDBQT file with Gasteiger charges using AutoDock Tools. The grid box was centred based on the coordinates of tryptophan (**31**) bound to PrnA (38.23, -7.951, 223.972, *xyz*) with 0.375 Å spacing and dimensions of 34 x 28 x 24 (*x* x *y* x *z*). Number of binding modes was set to 15 and all other parameters left as default.

7.3 Chapter 3 Experimental

All compounds detailed in this section have been characterised as part of accepted articles in peer-reviewed journals. Copies of the original spectral data for these compounds is therefore available as part of the Supporting Information of these publications. The relevant publications are cited where appropriate.

7.3.1 Enzyme Expression and Purification

Cloning of the halogenases RebH, PyrH and SttH into pET 28a (+) and the flavin reductase Fre into pET45b (+) has been previously reported by our group.[156,166,175] RadH cloned into pET 28b (+) as previously described by our group was used for protein expression.[193] These constructs were provided by Sarah Shepherd or Binuraj Menon. Alcohol dehydrogenase from *Rhodoccus* sp. (ADH) was cloned into pET21b (+) by Sarah Shepherd as previously described.[160] GDH2 cloned into pET 21b (+) was provided by the Scrutton Group (University of Manchester).

For protein expression, the above RebH, PyrH and SttH plasmids were separately transformed into chemically-competent *E. coli* Arctic Express (DE3) using kanamycin (50 µg/mL) and gentamycin (20 µg/mL) for selection. The RadH plasmid was transformed into chemically-competent *E. coli* Rosetta 2 (DE3) cells for overexpression using kanamycin (50 µg/mL) and chloramphenicol (30 µg/mL) for selection. The above Fre and GDH plasmids were separately transformed into chemically competent *E. coli* BL21 (DE3) using ampicillin (50 µg/mL) for selection prior to protein expression. The ADH plasmid was transformed into *E. coli* BL21 (DE3) containing pGro7 (Takara) as previously described for protein expression.[160] The composition of each protein purification buffer is detailed above (7.1.7).

7.3.1.1 RebH, SttH and PyrH Expression

Single colonies of transformant from the plates prepared above were picked and used to inoculate 10 mL of LB medium containing kanamycin (50 µg/mL) and gentamycin (20 µg/mL). After incubation with shaking at 37 ^{o}C overnight, this culture was diluted 1:100 into 100 mL fresh LB containing kanamycin (50 µg/mL) and gentamycin (20 µg/mL). After further incubation at 37 ^{o}C overnight, the culture was diluted 1:100 into 10 x 440 mL TB medium containing kanamycin (50 µg/mL) and incubated at 37 ^{o}C with shaking at 180 rpm until OD_{600} = 0.6 – 0.8. After incubation at 4 ^{o}C for 30 min as cold shock, protein expression was induced with IPTG (0.1 mM final concentration) and cultures incubated for a further 20 hrs at 15 ^{o}C with shaking at 180 rpm. Cells were then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C) and stored at -20 ^{o}C prior to purification.

7.3.1.2 RadH Expression

Single colonies of transformant from the plates prepared above were picked and used to inoculate 10 mL of LB medium containing kanamycin (50 µg/mL) and chloramphenicol (30 µg/mL). After incubation with shaking at 37 ^{o}C overnight, this culture was diluted 1:100 into 100 mL fresh LB containing kanamycin (50 µg/mL) and chloramphenicol (30 µg/mL). After further incubation at 37 ^{o}C overnight, the culture was diluted 1:100 into 10 x 400 mL 2YT broth containing kanamycin

(50 µg/mL) and incubated at 37 ^{o}C with shaking at 180 rpm until OD_{600} = 0.4 – 0.6. Protein expression was then induced with IPTG (0.1 mM final concentration) and incubation continued at 18 ^{o}C with shaking at 180 rpm for 20 hrs. Cells were then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C) and stored at -20 ^{o}C prior to purification.

7.3.1.3 Fre and GDH Expression

Single colonies of transformant from the plates prepared above were picked and used to inoculate 10 mL of LB medium containing ampicillin (50 µg/mL). After incubation at 37 ^{o}C with shaking overnight, this culture was diluted 1:100 into 100 mL fresh LB containing ampicillin (50 µg/mL) and incubated at 37 ^{o}C with shaking overnight. This culture was then diluted 1:100 into 10 x 400 mL 2YT broth and incubated at 37 ^{o}C with shaking at 180 rpm until OD_{600} = 0.4 – 0.6 was reached. After induction with IPTG (1.0 mM final concentration), cultures were incubated at 18 ^{o}C with shaking at 180 rpm overnight and then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C). Cells expressing GDH were stored at -20 ^{o}C prior to purification whilst Fre was purified immediately.

7.3.1.4 ADH Expression

Alcohol dehydrogenase (ADH) was expressed as previously described by Frese *et. al.*[160]

7.3.1.5 General Protein Purification

Pelleted cells were re-suspended into 10 mM imidazole KPi buffer containing 4 protease inhibitor tablets (ca. 60 mL total volume for cells from 4 L culture). Cells were then lysed by sonication (10 min, 50 % pulse, 70 % power, 700 W) and the lysate clarified by centrifugation (18000 rpm, 45 min, 4 ^{o}C). Lysate was then loaded onto Ni-NTA resin (Qiagen), pre-equilibrated with 10 mM imidazole KPi buffer, under gravity. The resin was then washed with 3 column volumes (CV) of 80 mM imidazole KPi buffer prior to elution with 300 mM imidazole KPi buffer (5CV). Eluent was then buffer exchanged into 100 mM KPi buffer (pH 7.2) using either spin concentration or dialysis. After spin concentration to ca 1.5 mL, protein concentration was determined using the A280 feature of a ThermoScientific 2000 Nanodrop. Clarified cell lysate, flow through, wash and eluent fractions were then analysed by SDS-PAGE.

7.3.1.6 SDS-PAGE Analysis

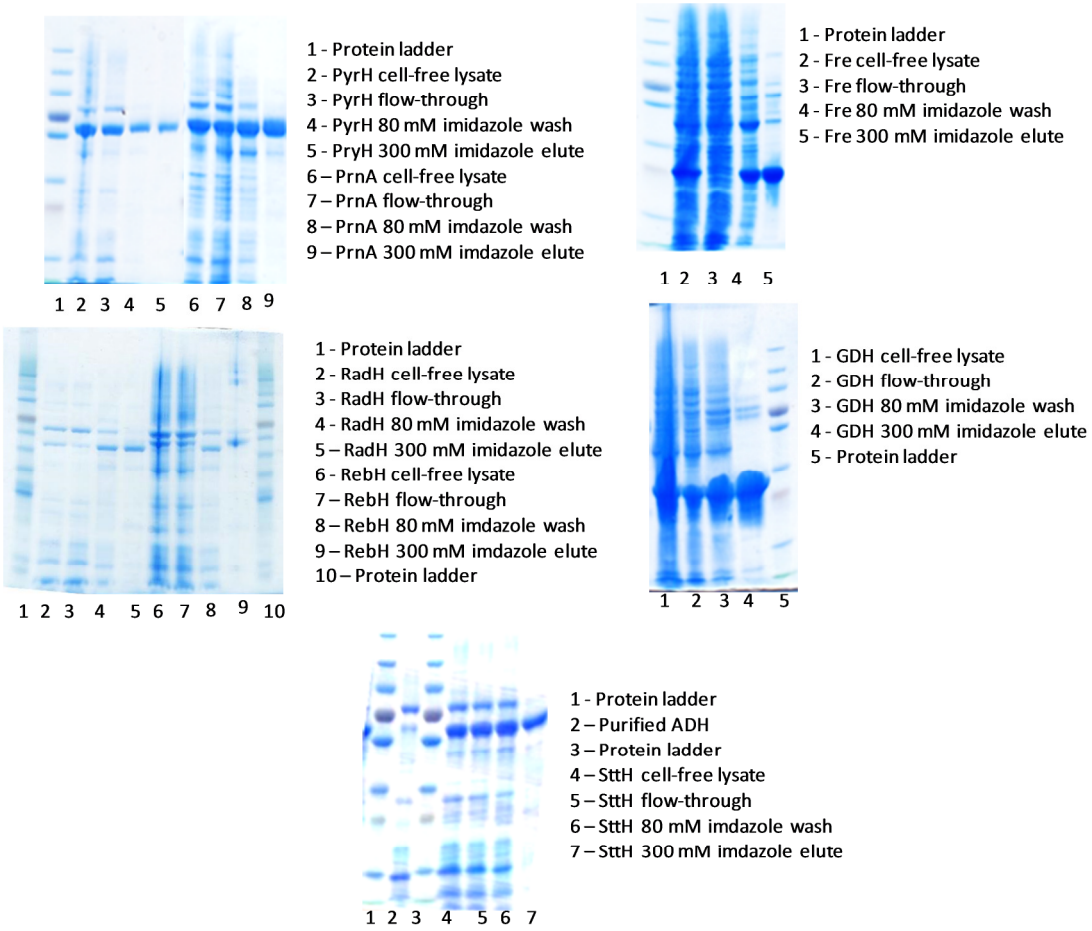


Figure S 1: SDS-PAGE images of each of the proteins purified as part of this chapter. Protein mwt: 58.1 kDa (PyrH); 61.1 kDa (PrnA); 26.1 kDa (Fre); 56.5 kDa (RadH); 60.3 kDa (RebH); 28085 (GDH); 58.6 kDa (SttH). The ladder used is PageRuler^{TM} Prestained Protein Ladder from ThermoScientific^{TM}.

7.3.1.7 ADH Purification

Alcohol hydrogenase (ADH) was purified by Sarah Shepherdusing the method described by Frese *et. al.*[160]

7.3.1.8 ADH and Fre Activity Determination

Activity units of ADH and Fre, U (µmol/min), were determined spectrophotometrically using the previously reported NAD^+/NADH reduction/oxidation method.[160]

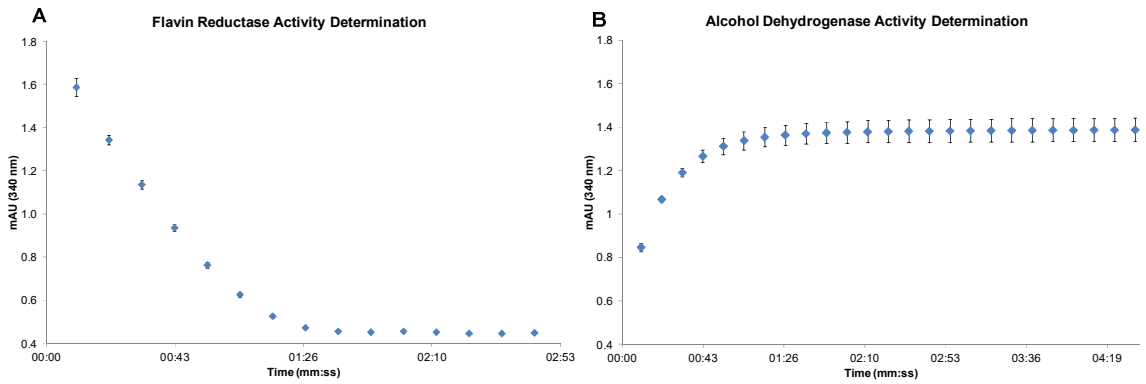


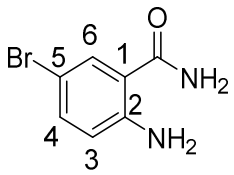
Figure S 2: Typical timecourses of (**A**) NADH oxidation by Fre and (**B**) NAD^+ reduction by ADH determined by measuring UV absorbance at 340 nm used to quantify activity of Fre and ADH batches.

7.3.1.9 CLEA Preparation

The method for preparing halogenase CLEAs is based upon previous reports.[197] The pET 28a (+) halogenase constructs were transformed into chemically-competent *E. coli* BL21 (DE3) cells containing pGro7 (Takara) using kanamycin (50 µg/mL) and chloramphenicol (20 µg/mL) for selection. Single colonies were picked and used to inoculate 10 mL of LB medium, containing the same concentrations of antibiotics, and incubated at 37 °C with shaking until OD_{600} = 0.6. This culture was then diluted into 1:100 into 100 mL fresh LB containing the same concentrations of antibiotics and incubated at 37 °C with shaking overnight. The overnight culture was then diluted 1:100 into 6 x 1 L LB medium containing kanamycin (50 µg/mL) and chloramphenicol (20 µg/mL) and grown at 37 °C with shaking at 180 rpm until OD_{600} = 0.6. Halogenase and chaperone expression was then induced using IPTG (0.1 mM final) and L-arabinose (1 mg/mL final). Cultures were then incubated at 18 °C with shaking at 180 rpm overnight prior to harvesting by centrifugation (4000 rpm, 20 min, 4 °C). The resultant cell pellets were either used immediately or stored at -20 °C until use. Cell pellets from a total of 1.5 L of culture were re-suspended into 30 mL 100 mM sodium phosphate buffer (pH 7.4) and lysed by sonication (12 min, 50 % pulse, 70 % power, 700 W). Lysate was then clarified by centrifugation (10000 rpm, 30 min, 4 °C) prior to addition of purified Fre (2.5 U/mL) and ADH (1 U/mL). After thorough mixing by inversion, protein was precipitated by addition of ammonium sulphate (16.2 g, 95 % saturation) and agitated at 4 °C for 1 hr. Glutaraldehyde was then added as a 50 % wt/v solution in water (0.5 % w/v final concentration) prior to agitation at 4 °C for 2.5 hrs. The resultant aggregate was then collected by centrifugation (10000 rpm, 30 min, 4 °C) and washed with 40 mL 100 mM sodium phosphate buffer (pH 7.4) thrice. The CLEA was then either directly re-suspended into reaction buffer or 100 mM sodium phosphate buffer (pH 7.4) for storage at 4 °C until required.

7.3.2 Synthesis of Substrates, Standards and Ligands

7.3.2.1 2-amino-5-bromo benzamide (84)[191,193]



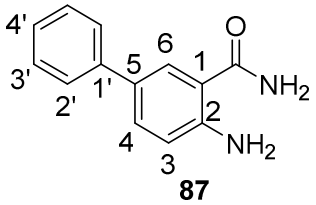
84

2-amino-5-bromobenzoic acid (**85**, 215 mg, 1.0 mmol, 1.0 eq) and 1,1'-carbonyl diimidazole (178 mg, 1.1 mmol, 1.0 eq) were added to an oven-dried flask containing a stirrer bar. After backfill with N_2, anhydrous THF (10 mL) was added and the colourless solution stirred at room temperature for 6 hrs. Saturated ammonium hydroxide solution (10 mL) was added and stirring continued overnight. THF was then removed under a stream of N_2 and the resultant suspension extracted with EtOAc (3 x 10 mL). The combined organics were then dried over $MgSO_4$ and solvent removed *in vacuo*. The crude was then purified by flash chromatography (SiO_2; Et_2O) to afford the title compound as a white solid (157 mg, 74 %); mp 182 – 185 ^oC; ν_{max}/cm^{-1} 3394 (br., N-H_2), 3349 (br., N-H_2), 3287 (br., CON-H_2), 3156 (br., CON-H_2), 2956 (w, C-H), 2921 (w, C-H), 2852 (w, C-H), 1738 (s, C=O), 1579 (m, C=C); ^{1}H NMR (400 MHz, DMSO) δ 7.84 (1H, br. s, CONH_2), 7.69 (1H, d, *J* = 2.4 Hz, H6), 7.25 (1H, dd, *J* = 8.8, 2.4 Hz, H4), 7.19 (1H, br. s, CONH_2), 6.72 (2H, br. s, NH_2), 6.65 (1H, d, J = 8.8 Hz, H3); ^{1}H NMR (400 MHz, MeOD) δ 7.53 (1H, d, *J* = 2.5 Hz, H6), 7.15 (1H, dd, *J* = 8.8, 2.5 Hz, H4), 6.72 (1H, d, *J* = 8.8 Hz, H3); ^{13}C NMR (101 MHz, DMSO) δ 167.0 (C=O), 149.4 (ArC), 134.3 (C4), 130.8 (C6), 118.5 (ArC), 115.2 (ArC), 104.7 (ArC); m/z (ESI) 213.0 ([M^{79}Br-H]$^-$, 50 %), 215.0 ([M^{81}Br-H]$^-$, 50); HRMS *m/z* calcd for $C_7H_7ON_2Br_1Na_1$ 236.9639, found: 236.9635; λ_{max} (H_2O/CH_3CN) 228, 248 and 328 nm.

7.3.2.2 General Method for Preparation of 2-amino-5-aryl benzamides

A flask containing 2-amino-5-bromobenzamide (**84**, 0.4 mmol, 1.0 eq), aryl boronic acid (1.0 mmol, 2.6 eq) and potassium carbonate (3.6 mmol, 9.0 eq) was fitted with a stirrer bar and backfilled with N_2. Under a positive pressure of N_2, tetrakis(triphenylphosphine)palladium^0 (40 µmol, 10 mol %) was added prior to addition of degassed toluene (7.5 mL), water (5 mL) and ethanol (2.5 mL) successively. The resultant suspension was then stirred vigorously at reflux under N_2 overnight. Upon cooling, the reaction was diluted with CH_2Cl_2 (25 mL) and washed with saturated NH_4Cl solution (25 mL). The aqueous phase was then extracted with further CH_2Cl_2 (2 x 25 mL) and the combined organics washed with saturated $NaHCO_3$ solution (25 mL) then water (25 mL). The combined organics were then dried over $MgSO_4$ prior to removal of solvent *in vacuo* and purification by semi-preparative HPLC method 1.

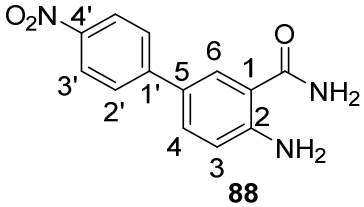
7.3.2.3 2-amino-5-phenyl benzamide (87)[193,293]



87

Prepared and purified according to the general method above (7.3.2.2) using phenyl boronic acid. The title compound was afforded as an off-white solid (72.8 mg, 88 %); mp 159 – 161 °C; ν_{max}/cm^{-1} 3399 (br., N-H$_2$), 3299 (br., N-H$_2$), 3173 (br., CON-H$_2$), 3051 (w, C-H), 3022 (w, C-H), 1676 (s, C=O), 1509 (m, C=C); ^{1}H NMR (400 MHz, DMSO) δ 7.96 (1H, br. s, CONH$_2$), 7.86 (1H, d, *J* = 2.1 Hz, H6), 7.65 (2H, dd, *J* = 8.3, 1.1 Hz, H2'), 7.50 (1H, dd, *J* = 8.5 Hz, H4), 7.40 (2H, t, *J* = 7.7 Hz, H3'), 7.25 (1H, t, *J* = 7.3 Hz, H4'), 7.14 (1H, br. s, CONH$_2$), 6.78 (1H, d, *J* = 8.6 Hz, H3), 6.70 (2H, br. s, NH$_2$); ^{13}C NMR (101 MHz, DMSO) δ 171.22 (C=O), 150.1 (C-N), 131.2 (ArC), 128.6 (ArCH), 126.9 (ArCH), 126.3 (ArCH), 126.0 (ArCH), 125.6 (ArCH), 117.1 (ArCH), 114.2 (ArC); *m/z* (ESI) 213.0 ([M+H]$^+$, 20 %), 196.0 ([M-NH$_2$]$^+$, 100), 168.0 ([M-CONH$_2$]$^+$); HRMS *m/z* calcd for $C_{13}H_{10}NO$ [M-NH$_2$]$^+$ 196.0757, found: 196.0754; λ_{max} (H$_2$O/CH$_3$CN) 278 and 336 nm.

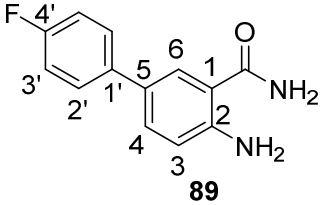
7.3.2.4 2-amino-4'-nitro-5-phenyl benzamide (88)[193]



88

Prepared and purified according to the general method above (7.3.2.2) using 4-nitrophenyl boronic acid. The title compound was afforded as an orange solid (55 mg, 91 %); ν_{max}/cm^{-1} 3452 (br., N-H$_2$), 3329 (br., N-H$_2$), 3182 (br., CON-H$_2$), 2921 (w, C-H), 2851 (w, C-H), 1652 (s, C=O), 1650, 1620 (m, C=C), 1587, 1330 (NO); ^{1}H NMR (400 MHz, CD$_3$CN) δ 7.68 (2H, d, *J* = 9.0 Hz, H3'), 7.49 (1H, br s, CONH$_2$), 7.48 (1H, d, *J* = 2.2 Hz, H6), 7.39 (2H, d, *J* = 9.0 Hz, H2'), 7.11 (1H, dd, *J* = 8.8, 2.2 Hz, H4), 6.67 (1H, br s, CONH$_2$), 6.26 (1H, d, *J* = 8.8 Hz, H3); ^{13}C (101 MHz, DMSO) δ 170.9 (C=O), 151.1, 146.6, 146.6, 132.1 (C4), 130.6 (C5), 127.9 (C2'), 125.9 (C3'), 124.0, 123.2 (C6), 117.1, 113.8 (C3); *m/z* (ESI) 256.0 ([M-H]$^-$); HRMS *m/z* calcd for $C_{13}H_{12}N_3O_3$ 258.0873, found: 258.0888.; λ_{max} (H$_2$O/CH$_3$CN) 322 and 404 nm.

7.3.2.5 2-amino-4'-fluoro-5-phenyl benzamide (89)[193]

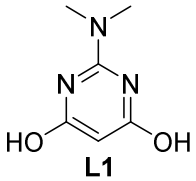


89

Prepared and purified according to the general method above (7.3.2.2) using 4-fluorophenyl boronic acid. The title compound was afforded as an off-white solid (46 mg, 85 %); ν_{max}/cm^{-1} 3469 (br., NH$_2$), 3349 (br., NH$_2$), 3210 (br., CONH$_2$), 2954 (w, C-H), 1744 (s, C=O), 1674 (m, C=C), 1603

(m, C=C); ^{1}H NMR (400 MHz, MeOD) δ 7.82 (1H, d, *J* = 2.2 Hz, 1H, H6), 7.57 (2H, dd, *J* = 8.8, 5.4 Hz, H3'), 7.52 (1H, dd, *J* = 8.6, 2.2 Hz, H4), 7.09 (2H, t, *J* = 8.8 Hz, H2'), 6.94 (1H, d, *J* = 8.6 Hz, 1H, H3); ^{13}C NMR (101 MHz, MeOD) 173.69 (C=O), 163.55 (d, *J* = 245.4 Hz, C4'), 146.35 (C2), 137.88 (C1'), 132.16 (C4), 132.10 (C5), 132.13 (d, *J* = 8 Hz, C2'), 128.13 (C6), 120.51 (C1), 118.58 (C3), 116.46 (d, *J* = 21.2 Hz, C3'); ^{19}F NMR (376 MHz, CD_3CN) δ -76.77 (s); *m/z* (ESI) 253.1 ($[M+Na]^+$, 80 %), 214.1 ($[M-NH_2]^+$, 100); HRMS *m/z* calcd for $C_{13}H_{11}N_2OFNa$ ($[M+Na]^+$) 253.0753, found: 253.0763; λ_{max} (H_2O/CH_3CN) 274 and 326 nm.

7.3.2.6 2-(dimethylamino)pyrimidine-4,6-diol (L1)[178]



L1

1,1-dimethyl guanidine sulphate (5.1 g, 18.75 mmol, 1.0 eq) was added to a stirred solution of sodium ethoxide (1.46 g, 21.4 mmol, 1.1 eq) in ethanol (11 mL). Diethyl malonate (3 mL, 18.75 mmol, 1.0 eq) dissolved in a second solution of sodium ethoxide (2.42 g, 35.6 mmol, 1.9 eq) in ethanol (19 mL) was then added to the guanidine solution prior to reflux of the combined solution for 5 hrs. After cooling and removal of solvent *in vacuo*, the resultant residue was dissolved in water (10 mL) and pH adjusted to 6 using acetic acid. Collection of the white solid from the resultant suspension by vacuum filtration afforded the title compound as a white solid (1.89 g, 65 %); mp 292 °C dec.; ν_{max}/cm^{-1} 3325 (br, O-H), 2953 (w, C-H); ^{1}H NMR (400 MHz, DMSO) δ 10.48 (2H, br. s, OH), 4.66 (1H, s, ArH), 3.00 (6H, s, CH_3); ^{13}C NMR (101 MHz, DMSO) δ 78.0 (ArCH), 75.8 (C-O), 39.3 (C-N), 36.8 (CH_3); *m/z* (ESI) 156.2 ($[M+H]^+$, 100 %), 178.2 ($[M+Na]^+$, 70); HRMS *m/z* calcd for $C_6H_{10}N_3O_2$ 156.0768, found: 156.0776.

7.3.2.7 Pyrimidine and Guanidine Palladium Complexes

Stock solutions of the pyrimidine and guanidine palladium complexes at 10 mM were prepared as per pervious reports.[178,181] Pyrimidine or guanidine ligand (0.1 mmol) was dissolved in 100 mM NaOH (2 mL) and H_2O (1 mL) with rapid stirring. $Pd(OAc)_2$ (11.2 mg, 50 µmol) was then added prior to stirring of the suspension at 65 °C for 45 min. The solution was then cooled and diluted to 5 mL to give a 10 mM catalyst stock.

7.3.3 Analytical Scale Halogenase Biotransformations

7.3.3.1 PyrH-Catalysed Bromination

To a solution containing NaBr (100 mM), glucose (20 mM), FAD (1.0 µM), Fre (2.0 µm), GDH (12 µM), PyrH (20 µm) and anthranilamide (**62**) or 3-indolepropionic acid (**106**) (2.0 mM) in 10 mM potassium phosphate buffer (pH 7.2) was added NADH (100 µM) to a total volume of 200 µL. After incubation at 20 °C with shaking at 300 rpm overnight, reactions were quenched by heating at 95 °C for 10 min prior to centrifugation (13000 rpm, 10 min) and analysis of the supernatant by analytical HPLC method 2.

7.3.3.2 RebH-Catalysed Bromination

To a solution containing NaBr (100 mM), glucose (20 mM), FAD (100 μM), Fre (5.0 μm), GDH (10 μM), RebH (50 μm) and tryptophol (**104**) (2.0 mM) in 10 mM potassium phosphate buffer (pH 7.2) containing 4 % isopropanol was added NADH (100 μM) to a total volume of 200 μL. After incubation at 20 °C with shaking at 300 rpm overnight, reactions were quenched by addition of 200 μL of methanol followed by heating at 95 °C for 10 min prior to centrifugation (13000 rpm, 10 min) and analysis of the supernatant by analytical HPLC method 2.

7.3.3.3 SttH-Catalysed Bromination

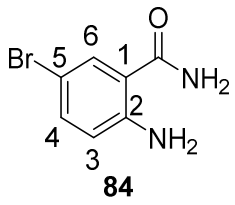
To a solution containing NaBr (100 mM), glucose (20 mM), FAD (7.5 μM), Fre (1.0 μm), GDH (6.0 μM), SttH (10 μm) and 3-indolepropionic acid (**106**) (2.0 mM) in 10 mM potassium phosphate buffer (pH 7.2) was added NADH (100 μM) to a total volume of 200 μL. After incubation at 20 °C with shaking at 300 rpm overnight, reactions were quenched by heating at 95 °C for 10 min prior to centrifugation (13000 rpm, 10 min) and analysis of the supernatant by analytical HPLC method 2.

7.3.3.4 RadH-Catalysed Bromination

To a solution containing NaBr (10 mM), FAD (1.0 μM), RadH (25 μM), Fre (4 μM) and 6-hydroxyisoquinoline (**109**) (0.5 mM) in 10 mM potassium phosphate buffer (pH 7.2) containing 1 % ethanol was added NADH (2.5 mM) to a total volume of 200 μL. Reactions were then incubated at 30 °C with shaking at 300 rpm overnight prior to quenching by heating at 95 °C for 10 min. Precipitated protein was then removed using centrifugation (13000 rpm, 10 min) and the supernatant analysed by analytical HPLC method 2.

7.3.4 Preparative Scale Halogenase Biotransformations

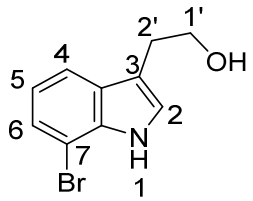
7.3.4.1 2-amino-5-bromo benzamide (84)[191,193]



84

Biotransformation was prepared in a loosely-capped falcon tube as per the general PyrH method stated above (7.3.3.1) using anthranilamide (**62**) to a total volume of 10 mL. After incubation at 20 °C overnight, protein was removed by boiling (95 °C, 10 min) and centrifugation (10000 rpm, 30 min, 4 °C). The supernatant was then lyophilised and re-dissolved in methanol. Purification by semi-preparative HPLC method 1 afforded the title compound as a white solid (3.9 mg, 91 %); ^{1}H NMR (400 MHz, MeOD) δ 7.53 (1H, d, *J* = 2.5 Hz, H6), 7.15 (1H, dd, *J* = 8.8, 2.5 Hz, H4), 6.72 (1H, d, *J* = 8.8 Hz, H3); *m/z* (ESI) 213.0 ([M^{79}Br-H]$^{-}$, 50 %), 215.0 ([M^{81}Br-H]$^{-}$, 50 %); HRMS *m/z* calcd for $C_7H_7ON_2{}^{79}BrNa$ 236.9639, found: 236.9635.

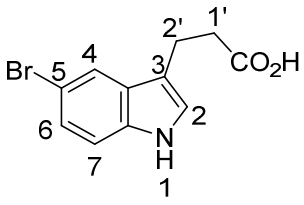
7.3.4.2 7-bromotryptophol (105)[193,236]



105

Biotransformation was prepared in a loosely-capped falcon tube as per the general RebH method stated above (7.3.3.2) using tryptophol (**104**) to a total volume of 20 mL. After incubation at 20 °C overnight, protein was removed by boiling (95 °C, 10 min) and centrifugation (10000 rpm, 30 min, 4 °C). The supernatant was then lyophilised and re-dissolved in methanol. Purification by semi-preparative HPLC method 3 afforded the title compound as a white solid (4.4 mg, 92 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 (1H, d, *J* = 8.0 Hz, H6), 7.36 (1H, d, *J* = 8.0 Hz, H4), 7.17 (1H, d, *J* = 2.0 Hz, H2), 7.02 (1H, t, *J* = 8 Hz, H5), 3.91 (2H, t, *J* = 7.4 Hz, H1'), 3.02 (2H, t, *J* = 7.4 Hz, H2'); *m/z* (ESI) 239.0 ([M^{79}Br+H]$^+$, 50 %), 241.0 ([M^{81}Br+H]$^+$, 50 %); HRMS *m/z* calcd for $C_{10}H_9{}^{79}BrNO$ 237.9873, found: 237.9855.

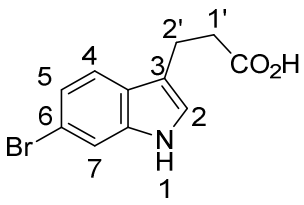
7.3.4.3 5-bromo-3-indolepropionic acid (107)[193,237]



107

Biotransformation was prepared in a loosely-capped falcon tube as per the general PyrH method stated above (7.3.3.1) using indole-3-propionic acid (**106**) to a total volume of 10 mL. After incubation at 20 °C overnight, protein was removed by boiling (95 °C, 10 min) and centrifugation (10000 rpm, 30 min, 4 °C). The supernatant was then lyophilised and re-dissolved in methanol. Purification by semi-preparative HPLC method 3 afforded the title compound as a white solid (4.6 mg, 87 %); ^{1}H NMR (400 MHz, MeOD) δ 7.68 (1H, d, *J* = 1.8 Hz, H4), 7.24 (1H, d, *J* = 8.6 Hz, H7), 7.16 (1H, dd, *J* = 8.6, 1.8 Hz, H6), 7.09 (1H, s, H2), 3.01 (2H, t, *J* = 7.6, 1'H), 2.61 (2H, t, *J* = 7.6 Hz, 2'H); *m/z* (ESI) 266.0 ([M^{79}Br-H]$^-$, 50 %), 268.0 ([M^{81}Br-H]$^-$, 50); HRMS *m/z* calcd for $C_{11}H_{11}{}^{79}BrNO_2$ 267.9968, found: 267.9964.

7.3.4.4 6-bromo-3-indolepropionic acid (108)[193,294]

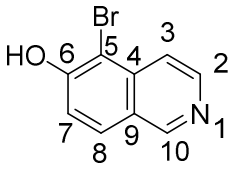


108

Biotransformation was prepared in a loosely-capped falcon tube as per the general SttH method stated above (7.3.3.3) using indole-3-propionic acid (**106**) to a total volume of 10 mL. After incubation at 20 °C overnight, protein was removed by boiling (95 °C, 10 min) and centrifugation

(10000 rpm, 30 min, 4 °C). The supernatant was then lyophilised and re-dissolved in methanol. Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (3.8 mg, 72 %); ^{1}H NMR (400 MHz, Acetone) δ 10.19 (1H, s, NH), 7.59 (1H, s, H7), 7.57 (1H, d, *J* = 8.4 Hz, H5), 7.22 (1H, s, H2), 7.17 (1H, d, *J* = 8.4 Hz, H4), 3.05 (2H, t, *J* = 7.6, H1'), 2.70 (2H, t, *J* = 7.6 Hz, 2'H); *m/z* (ESI) 266.0 ([M^{79}Br-H]$^{-}$, 50 %), 268.0 ([M^{81}Br-H]$^{-}$, 50); HRMS *m/z* calcd for $C_{11}H_{11}{}^{79}BrNO_2$ 267.9968, found: 267.9966. 5-bromo-3-indolepropionic acid (**107**) was also afforded as a minor by-product (0.6 mg, 12 %).

7.3.4.5 5-bromo-6-hydroxyisoquinoline (110)[193,295]



110

Biotransformation was prepared in a loosely-capped falcon tube as per the general RadH method stated above (7.3.3.4) using 6-hydroxyisoquinoline (**109**) to a total volume of 30 mL. After incubation at 30 °C with shaking at 180 rpm overnight, protein was removed by boiling (95 °C, 10 min) and centriguation (10000 rpm, 30 min, 4 °C). The supernatant was then lyophilised and re-dissolved in methanol and water (1:1). Purification by semi-preparative HPLC method 2 afforded the title compound as a yellow solid (2.4 mg, 72 %); ^{1}H NMR (400 MHz, DMSO) δ 9.44 (1H, s, H10), 8.56 (1H, d, *J* = 6.4 Hz, H2), 8.25 (1H, d, *J* = 8.9 Hz, H7), 8.07 (1H, d, *J* = 6.4 Hz, H3), 7.57 (1H, d, *J* = 8.9 Hz, H8); *m/z* (ESI) 224.2 ([M^{79}Br+H]$^{+}$, 50 %), 226.2 ([M^{81}Br+H]$^{+}$, 50); HRMS *m/z* calcd for $C_9H_7ON^{79}Br$ 223.9706, found: 223.9710.

7.3.5 Suzuki-Miyarua Coupling (SMC) Optimisation Methods

7.3.5.1 Analytical Scale SMC for Conditions Screening (Pyrimidine and Guanidine Complexes)

To a solution of 2-amino-5-bromobenzamide (**84**) (2 mM), CsF (10 mM) and phenyl boronic acid (10 mM) in 10 mM potassium phosphate buffer in a glass crimp-top vial was added the appropriate palladium-pyrimidine or –guanidine complex as a 10 mM stock solution (0.2 mM final) to a final volume of 200 µL. Vials were then crimped closed prior to heating at the specified temperature for the specified time. Reactions were then partitioned into Et_2O (200 µL) and the organic phase removed *in vacuo*. The resultant residue was then re-dissolved in MeOH (200 µL) and clarified by centrifugation (13000 rpm, 10 min) prior to analysis by analytical HPLC method 1. The above method was also used to determine the tolerance of these conditions to the halogenase biotransformation components, with additive added at the specified concentration prior to addition of palladium catalyst.

7.3.5.2 Analytical Scale SMC for Tolerance Screening (Phosphine Complexes)

To a solution of 2-amino-5-bromobenzamide (**84**) (2 mM) and K_3PO_4 (2.4 mM) in 10 mM potassium phosphate buffer in a glass crimp-top vial was added phenyl boronic acid at the specific concentration. The vial was then crimped closed and degassed under a stream of N_2 prior to

addition of tppts and Na_2PdCl_4 as stock solutions in deoxygenated water to a total volume of 200 µL. Reactions were then heated at 80 ^oC overnight. Upon cooling, reactions with partitioned into Et_2O (200 µL) and organics removed *in vacuo*. The resultant residue was then re-dissolved in MeOH (200 µL) and particulates removed using centrifugation (13000, 10 min) prior to analysis by analytical HPLC method 2. The above method was also used to determine the tolerance of these conditions to the halogenase biotransformation components, with additive added at the specified concentration prior to degassing.

7.3.5.3 Analytical Scale SMC using Pure Protein

A halogenase-catalysed biotransformation was prepared as per the general methods described above (7.3.3) and allowed to proceed overnight before being filtered through a 10 kDa MWCO filter (Vivaspin 20). 200 µL aliquots were then transferred to crimp-top glass vials prior to addition of boronic acid and base as stock solutions in water at the concentrations specified. The vials were then sealed closed and degassed under a steam of N_2. Tppts and Na_2PdCl_4 were then added as stock solutions in deoxygenated water and the reactions heated at 80 ^oC for the specified time. Upon cooling, reactions were partitioned into Et_2O, EtOAc or CH_2Cl_2 and organics removed *in vacuo*. The resultant residue was then dissolved in MeOH (200 µL) and particulates removed by centrifugation (13000 rpm, 10 min) prior to analysis by analytical HPLC method 2.

7.3.5.4 Analytical Scale SMC using CLEAs

A halogenase CLEA-catalysed biotransformation was prepared as described below (7.3.6.1) and allowed to proceed overnight. After removal of CLEA (10000 rpm, 30 min), supernatant was transferred to a flask, sealed with a rubber septa and degassed under a stream of N_2. To individual glass reaction vials fitted with rubber septa and stirrer bars was then added boronic acid (15 mM) and K_3PO_4 (30 mM) prior to backfill with N_2. 1 mL of deoxygenated biotransformation supernatant was then added to each vial, followed by tppts and Na_2PdCl_4 at the concentrations specified as stock solutions in deoxygenated water. After heating at 80 ^oC overnight, reactions were cooled and extracted into EtOAc (1 mL) and organics removed *in vacuo*. The resultant residue was then dissolved in MeOH (1 mL) and clarified by centrifugation (13000 rpm, 10 min) prior to analysis by analytical HPLC method 2 or LC-MS.

7.3.6 General Integrated Halogenase-SMC Methods

7.3.6.1 General Method for Bromination using CLEAs

Substrate (3.0 mM), NaBr (30 mM), NADH (100 µM) and FAD (10 µM) were dissolved in 15 mM sodium phosphate buffer (pH 7.2) containing 5 % v/v isopropanol to a total volume of 30 mL. The specified CLEA was then suspended into the above reaction buffer and the resultant suspension incubated at room temperature with orbital shaking overnight. To allow isolation and purification of brominated intermediate, the CLEA could then be removed using centrifugation (10000 rpm, 30 min, 4 ^oC) and the supernatant processed as described above (7.3.4).

7.3.6.2 General Method for Arylation of 2-amino benzamide using Pure Protein

Anthranilamide (**62**) (2.0 mM), NaBr (100 mM), FAD (1 μM), glucose (20 mM), PyrH (20 μM), Fre (2 μM), GDH (12 μM) and NADH (100 μM) were dissolved in 10 mM potassium phosphate buffer (pH 7.2) to a total volume of 30 mL and the mixture incubated with shaking at room temperature overnight. The biotransformation was then filtered through a 10 kDa MWCO filter (Vivaspin 20) prior to addition of K_3PO_4 (1.2 eq) and boronic acid (30 eq) and freeze-thaw degassing. Following backfill with N_2, tppts (2.0 mM) and Na_2PdCl_4 (1.0 mM) were added as deoxygenated stock solutions in water and the reaction heated with stirring at 80 ^oC for 24 hrs under N_2. After cooling, reaction was basified with 4M NaOH prior to extraction into EtOAc (3 x 30 mL) and removal of solvent *in vacuo* before purification.

7.3.6.3 General Method for Arylation of 6-hydroxy isoqinoline (109) using Pure Protein

6-hydroxyisoquinoline (**109**) (0.5 mM), NaBr (10 mM), FAD (1 μM), RadH (25 μM), Fre (4 μM) and NADH (2.5 mM) were dissolved in 10 mM potassium phosphate buffer (pH 7.2) containing 1 % v/v EtOH to a total volume of 40 mL and the mixture incubated with shaking overnight at 30 ^OC. The biotransformation was then filtered through a 10 kDa MWCO filter (Vivaspin 20) prior to addition of CsF (4.8 eq) and boronic acid (30 eq). Following freeze-thaw degassing and backfill with N_2, tppts (0.5 mM) and Na_2PdCl_4 (0.25 mM) were added as deoxygenated stock solutions in water and the reaction heated at 80 ^oC with stirring for 24 hrs under N_2. Upon cooling, pH was adjusted to pH 6 using 1M HCl and the reaction extracted with CH_2Cl_2 before removal of solvent *in vacuo* and purification.

7.3.6.4 General Method for Arylation of indole-3-propionic acid using Pure Protein

A biocatalytic bromination of indole-3-propionic acid (**106**) using either PyrH or SttH was assembled as per the general methods above (7.3.4.1 and 7.3.4.3) to a total volume of 30 mL and incubated overnight at room temperature. The reaction was then filtered through a 10 kDa MWCO filter (Vivaspin 20) prior to addition of K_3PO_4 (30 mg, 0.14 mmol, 2.4 eq) and phenyl boroic acid (219 mg, 0.9 mmol, 30 eq) followed by freeze-thaw degassing. After backfill with N_2, tppts (2.0 mM) and Na_2PdCl_4 (1.0 mM) were added as deoxygenated stock solutions in water and the reaction heated at 80 ^oC overnight under N_2. After cooling, the reaction was washed with CH_2Cl_2 (10 mL) before adjusting to pH 1 with 1 M HCl and extraction of product into EtOAc (3 x 10 mL). Solvent was then removed *in vacuo* prior to purification.

7.3.6.5 General Method for Arylation using CLEAs

A halogenase CLEA-catalysed bromination was prepared as above and allowed to proceed at room temperature overnight. After removal of the CLEA by centrifugation (10000 rpm, 30 min, 4 ^oC), base (10 eq) and boronic acid (5 eq) were added to the supernatant and the resulting solution degassed with a stream of N_2. Tppts (20 mol %) and Na_2PdCl_4 (10 mol %) were then added as deoxygenated stock solutions in water before heating at 80 ^oC overnight with stirring under N_2. After cooling, the reaction was partitioned into either EtOAc or CH_2Cl_2 and concentrated *in vacuo* prior to purification.

7.3.6.6 General Method for Arylation using Pure Protein and PDMS Thimbles

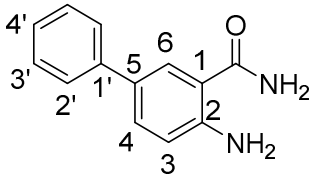
Anthranilamide (**62**) (2.0 mM), NaBr (100 mM), FAD (1 μM), glucose (20 mM), PyrH (20 μM), Fre (2 μM) and GDH (12 μM) and NADH (100 μM) were dissolved in 10 mM potassium phosphate buffer (pH 7.2) to a total volume of 30 mL. A PDMS thimble containing **L1**$_2$.Pd(OAc)$_2$ (10 mol %), CsF (10 eq) and PhB(OH)$_2$ (5 eq) in water (final volume 1 mL) was then added to the Erlenmeyer flask containing the biotransformation. After incubation at room temperature overnight, the reaction was heated to 80 ^{o}C for 8 hrs. Upon cooling the inner and outer compartments were combined and basified with 4N NaOH, whilst the thimble was soaked in EtOAc (30 mL). The aqueous portion was then extracted with further EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo* and purification.

7.3.6.7 General Method for Arylation using CLEAs and PDMS Thimbles

Substrate (5.0 mM), NaBr (30 mM), FAD (10 μM), and NADH (100 μM) were dissolved in 100 mM sodium phosphate buffer containing 5 % v/v isopropanol to a total volume of 30 mL. Halogenase CLEA from 3 L of culture was then resuspended into this reaction buffer and a PDMS thimble containing **L1**$_2$.Pd(OAc)$_2$ (2-10 mol %), CsF (10 eq) and ArB(OH)$_2$ (5 eq) in water to a total volume of 3 mL was added to the suspension. After incubation at room temperature with orbital shaking overnight, the reaction was heated to 80 ^{o}C for a further 24 hrs and products isolated and purified as previously described. The CLEA can also be removed using centrifugation (10000 rpm, 30 min, 4 ^{o}C) after overnight incubation at room temperature, before addition of the PDMS thimble containing SMC components, to allow recycling of the biocatalyst.

7.3.7 Regioselective Arylation using Pure Protein

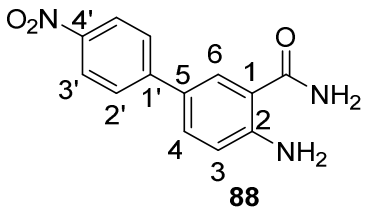
7.3.7.1 2-amino-5-phenyl benzamide (87)[193,293]



87

Prepared from anthranilamide (**62**) using PyrH and phenyl boronic acid according to the general method above (7.3.6.2). Purification using semi-preparative HPLC method 3 afforded the title compound as a white solid (10.0 mg, 79 %); ν_{max}/cm^{-1} 3433 (br., N-H), 3379 (br., N-H), 3303 (br., N-H), 3177 (br., N-H), 3033 (w, C-H), 2957 (w, C-H), 1640 (s, C=O), 1620 (m, C=C), 1567 (m, C=O); ^{1}H NMR (800 MHz, MeOH) δ 7.81 (1H, d, *J* = 2.2 Hz, H6), 7.59 (2H, d, *J* = 7.8 Hz, H2'), 7.50 (1H, dd, *J* = 8.5, 2.2 Hz, H4), 7.38 (2H, t, *J* = 7.6 Hz, H3'), 7.24 (1H, t, *J* = 7.6 Hz, H4'), 6.83 (1H, d, *J* = 8.5 Hz, H3); ^{13}C NMR (101 MHz, MeOD) δ 174.6 (C=O), 150.5 (C2), 141.9 (C1), 132.2 (C4), 130.3 (C5), 129.7 (C3'), 128.0 (C2), 127.3 (C4'), 127.1 (C6), 118.8 (C3), 116.0 (C1'); *m/z* (ESI) 235.1 ([M+Na]$^{+}$, 100 %), 213.1 ([M+H]$^{+}$, 45), 196.1 ([M-NH$_3$]$^{+}$, 50); HRMS *m/z* calcd for $C_{13}H_{12}N_2ONa$ 235.0847, found: 235.0847; λ_{max} (H_2O/CH_3CN) 278 and 336 nm.

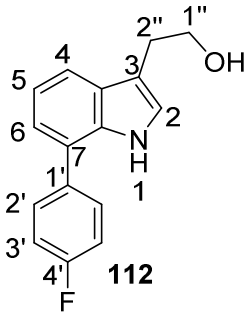
7.3.7.2 2-amino-5-(4'-nitrophenyl) benzamide (88)[193]



88

Prepared from anthranilamide (**62**) using PyrH and 4-nitrophenyl boronic acid according to the general method above (7.3.6.2). Purification by preparative TLC (2:3 EtOAc:Hexane) afforded the title compound as a yellow solid (7.4 mg, 48 %); ^{1}H NMR (400 MHz, CD_3CN) δ 7.67 (2H, app. d, H3'), 7.49 (1H, br. s, NH), 7.48 (1H, d, *J* = 2.4 Hz, H6), 7.39 (2H, app. d, H2'), 7.11 (1H, dd, *J* = 8.8, 2.4 Hz, H4), 6.67 (1H, br. s, NH), 6.26 (1H, d, *J* = 8.8 Hz); ^{13}C NMR (101 MHz, CD_3CN) δ 172.1 (C=O), 151.8 (ArC), 147.8 (ArC), 147.1 (ArC), 132.2 (ArCH), 128.8 (ArCH), 127.3 (ArCH), 126.0 (ArC), 125.1 (ArC), 118.3 (ArCH), 115.0 (ArC); *m/z* (ESI) 256.0 ($[M-H]^-$), 280 ($[M+Na]^+$); HRMS *m/z* calcd for $C_{13}H_{12}N_3O_3$ 258.0873, found: 258.0888.

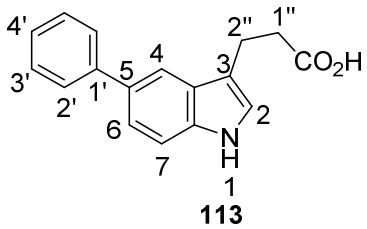
7.3.7.3 4'-fluoro-5-phenyl tryptophol (112)[193]



112

A RebH-catalysed biotransformation was assembled as per the general method stated above (7.3.3.2) to a total volume of 30 mL. After incubation at room temperature overnight, the reaction was filtered through a 10 kDa MWCO filter prior to addition of K_3PO_4 (15.2 mg, 70 μmol, 1.2 eq) and 4-fluorophenyl boronic acid (140 mg, 0.9 mmol, 30 eq) and stirring at 50 ^{O}C for 15 min to ensure dissolution of the boronic acid. The solution was then freeze-thaw degassed and backfilled with N_2 prior to addition of tppts (2.0 mM) and Na_2PdCl_4 (1.0 mM) as deoxygenated solutions in water. After heating under N_2 at 80 ^{O}C overnight, the reaction was cooled and partitioned into CH_2Cl_2 (3 x 30 mL) before removal of solvent *in vacuo*. Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (5.5 mg, 72 %); ^{1}H NMR (800 MHz, CD_3CN) δ 9.06 (1H, s, NH), 7.81 (2H, m, H3'), 7.57 (1H, dd, *J* = 8.0, 1.0 Hz, H6), 7.38 (1H, d, *J* = 8.0, H4), 7.11 – 7.07 (3H, M, H2+2'), 7.03 (1H, ddd, *J* = 8.0, 7.0, 1.0 Hz, H5), 3.74 (2H, t, *J* = 7.0 Hz, H1''), 2.92 (2H, t, *J* = 7.0, H2''); ^{13}C NMR (201 MHz, CD_3CN) δ 165.4 (d, $^{1}J_{CF}$ = 246.6 Hz, 4'C), 137.3 (d, $^{4}J_{CF}$ = 8.1 Hz, 1'C), 128.6 (ArC), 125.9 (C4), 123.7 (C2), 122.3 (C6), 119.5 (d, $^{2}J_{CF}$ = 25.2 Hz, C3'), 118.4 (ArC), 115.3 (d, $^{3}J_{CF}$ = 20.1 Hz, C2'), 114.1 (ArC), 113.1 (ArC), 112.1 (ArC), 63.1 (C1''), 29.5 (C2''); *m/z* (ESI) 254.3 ($[M-H]^-$); HRMS *m/z* calcd for $C_{16}H_{13}FNO$ ($[M-H]^-$) 254.0987, found: 254.0992; λ_{max} (H_2O/CH_3CN) 242 nm.

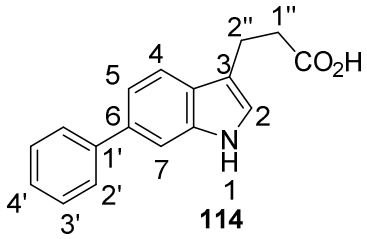
7.3.7.4 5-phenyl indole-3-propionic acid (113)[193]



113

Prepared from indole-3-propanoic acid (**106**) using PyrH as per the general method reported above (7.3.6.4). Purification by semi-preparative HPLC method 4 afforded the title compound as a brown solid (9.2 mg, 59 %); ^{1}H NMR (400 MHz, Acetone) δ 7.65 (1H, d, *J* = 1.9 Hz, H2), 7.47 (1H, d, *J* = 1.7 Hz, H4), 7.44 (2H, d, *J* = 8.6 Hz, H2'), 7.24 (1H, d, *J* = 8.6, H7), 7.15 – 7.07 (3H, m, H4'+3'), 7.04 (1H, dd, *J* = 8.4, 1.8 Hz, H6), 2.92 (2H, t, *J* = 7.6 Hz, H1''), 2.57 (2H, td, *J* = 7.5, 1.9 Hz, H2''); *m/z* (ESI) 266.2 ([M+H]$^+$, 100); HRMS *m/z* calcd for $C_{17}H_{14}NO_2$ ([M-H]$^-$) 264.1030, found: 264.1007; λ_{max} (H_2O/CH_3CN) 286, 324 nm.

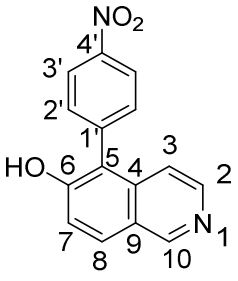
7.3.7.5 6-phenyl indole-3-propionic acid (114)[193]



114

Prepared from indole-3-propanoic acid (**106**) using SttH as per the general method reported above (7.3.7.4). Purification by preparative TLC (10:90:1, MeOH:CH_3Cl:AcOH) afforded the title compound as a brown solid (8.5 mg, 54 %); ^{1}H NMR (400 MHz, Ethanol) δ 7.70-7.69 (3H, m, H2'+2), 7.64 (1H, d, *J* = 8.3 Hz, H4), 7.45 (2H, t, *J* = 8 Hz, H3'), 7.34 (1H, t, *J* = 8 Hz, H4'), 7.30 (1H, m, H5), 6.49 (1H, m, H7), 4.36 (2H, d, *J* = 6.9 Hz, H1''), 2.76 (2H, t, *J* = 6.9 Hz, 2''); *m/z* (ESI) 266.2 ([M+H]$^+$, 100); HRMS *m/z* calcd for $C_{17}H_{14}NO_2$ ([M-H]$^-$) 264.1030, found: 264.1009; λ_{max} (H_2O/CH_3CN) 256, 312 nm.

7.3.7.6 4'-nitro-5-phenyl-6-hydroxy isoquinoline (116)[193]

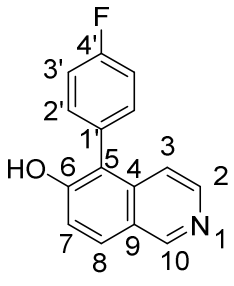


116

Prepared from 6-hydroxy isoquinoline (**109**) using RadH and 4-nitrophenyl boronic acid according to the general method above (7.3.6.3). Purification using semi-preparative HPLC method 2 afforded the title compound as a yellow solid (3.4 mg, 64 %); ^{1}H NMR (800 MHz, CD_3CN) δ 9.34 (1H, s, H1), 8.40 (2H, d, *J* = 9.1, H3'), 8.34 (1H, d, *J* = 9.1, H2), 8.22 (1H, d, *J* = 6.8 Hz, H7), 7.75

(1H, d, *J* = 9.1 Hz, H3), 7.64 (2H, d, *J* = 9.1 Hz, H2'), 7.61 (1H, d, *J* = 6.8 Hz, H8); ^{13}C NMR (201 MHz, CD_3CN) δ 162.2 (C6), 148.8 (ArC), 146.9 (ArC), 146.8 (ArC), 141.6 (ArC), 140.3 (ArC), 133.8 (ArC), 133.6 (ArC), 133.2 (C3'), 124.8 (C2'), 123.9 (ArC), 123.3 (C5), 121.1 (C3), 120.6 (C8); *m/z* (ESI) 265.1 ([M-H]$^-$); HRMS *m/z* calcd for $C_{15}H_{11}N_2O_3$ 267.0774, found: 267.0760; λ_{max} (H_2O/CH_3CN) 228, 252, 304, 330 nm.

7.3.7.7 4'-fluoro-5-phenyl-6-hydroxy isoquinoline (117)[193]



117

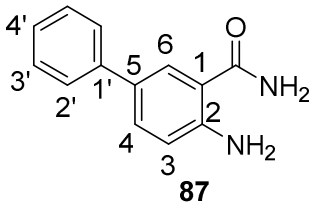
Prepared from 6-hydroxy isoquinoline (**109**) using RadH and 4-fluorophenyl boronic acid according to the general method above (7.3.6.3). Purification using semi-preparative HPLC method 2 afforded the title compound as a yellow solid (3.7 mg, 78 %); ^{1}H NMR (400 MHz, CD_3CN) δ 9.33 (1H, s, H1), 8.31 (1H, d, *J* = 9.1 Hz, H7), 8.19 (1H, d, *J* = 6.9 Hz, H2), 7.74 (1H, d, *J* = 9.1 Hz, H8), 7.65 (1H, d, *J* = 6.9 Hz, H3), 7.41 – 7.27 (4H, m, H2' & H3'); ^{13}C NMR (101 MHz, CD_3CN) δ 163.7 (d, $^1J_{CF}$ = 245.4 Hz, C4'), 162.6 (C6), 146.1 (C1), 141.2 (ArC), 133.9 (d, $^3J_{CF}$ = 8.4 Hz, C2'), 133.4 (ArCH), 132.2 (ArCH), 130.0 (d, $^4J_{CF}$ = 3.2 Hz, C1'), 124.3 (ArCH), 123.2 (ArC), 121.8 (ArCH), 116.8 (d, $^2J_{CF}$ = 21.8 Hz, C3'); ^{19}F NMR (470 MHz, CD_3CN) -115.5; *m/z* (ESI) 240.4 ([M+H]$^+$); HRMS *m/z* calcd for $C_{15}H_{11}FNO$ 240.0819, found: 240.0814; λ_{max} (H_2O/CH_3CN) 225, 258, 334 nm.

7.3.7.8 Arylation with Biocatalyst Recycling

A PyrH biotransformation of anthranilamide (**62**) (2.0 mM) was prepared to a total volume of 15 mL as per the general method detailed above (7.3.3.1) and incubated at room temperature for 3 hrs. The reaction was then concentrated to 500 µL using a 10 kDa MWCO filter (Vivaspin 20) and 100 µL of filtrate taken for analysis by HPLC. The concentrate was then diluted with 10 mM potassium phosphate buffer containing further anthranilamide (2.0 mM), NADH (100 µM), FAD (1 µM), NaBr (100 mM) and glucose (20 mM) to 15 mL total volume. The process of concentration and dilution was repeated after 3 hrs and the reaction left at room temperature overnight. The combined filtrates were then subjected to the SMC conditions detailed above and purified by semi-preparative HPLC method 1 to afford 2-amino-5-phenyl benzamide (**87**) as a white solid (10.3 mg, 54 %). Spectra were consistent with those previously reported herein.

7.3.8 Regioselective Arylation of Benzamides and Indoles using CLEAs

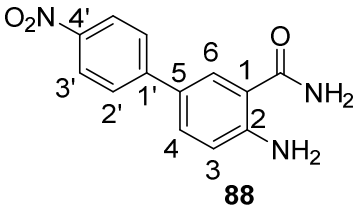
7.3.8.1 2-amino-5-phenyl benzamide (87)[293]



87

Prepared from anthranilamide (**62**) using PyrH CLEAs, K_3PO_4 and phenyl boronic acid according to the general method above (7.3.6.5). After cooling, reaction was basified using 4N NaOH and extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as a white solid (16.0 mg, 84 %). Spectra were consistent with those reported herein.

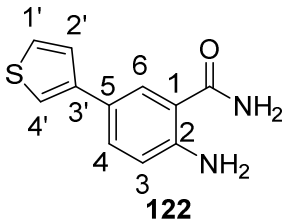
7.3.8.2 2-amino-4'-nitro-5-phenyl benzamide (88)[193]



88

Prepared from anthranilamide (**62**) using PyrH CLEAs, K_3PO_4 and 4-nitrophenyl boronic acid according to the general method above (7.3.6.5). After cooling, reaction was basified using 4N NaOH and extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as a yellow solid (18.8 mg, 81 %). Spectra were consistent with those reported herein.

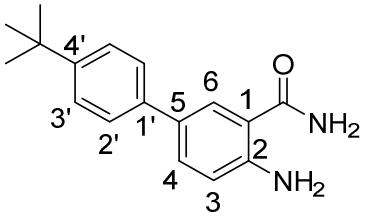
7.3.8.3 2-amino-5-(thiophen-3-yl) benzamide (122)[193]



122

Prepared from anthranilamide (**62**) using PyrH CLEAs, K_3PO_4 and 3-thieny boronic acid according to the general method above (7.3.6.5). After cooling, the reaction was extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as a yellow solid (13.0 mg, 71 %); ^{1}H NMR (400 MHz, CD_3CN) δ 7.27 (1H, d, *J* = 2.0 Hz, H6), 7.06 (1H, dd, *J* = 8.4, 2.0 Hz, H4), 6.97 (3H, m, H1', H2' & H4'), 6.30 (1H, d, *J* = 8.4 Hz, H3). ^{13}C NMR (101 MHz, CD_3CN) δ 160.6 (C=O), 159.0 (ArC), 131.6 (C4), 127.4 (ArCH), 127.3 (ArC), 126.9 (ArCH), 126.9 (ArCH), 124.8 (ArC), 119.1 (ArCH), 117.3 (C6); *m/z* (ESI) 202.0 ([M-NH_2]$^+$, 100 %), 219.0 ([M+H]$^+$, 5), 241.0 ([M+Na]$^+$, 20); HRMS *m/z* calcd for $C_{11}H_8NOS$ ([M-NH_2]$^+$) 202.0321, found: 202.0323; λ_{max} (H_2O/CH_3CN) 220, 276, 348 nm.

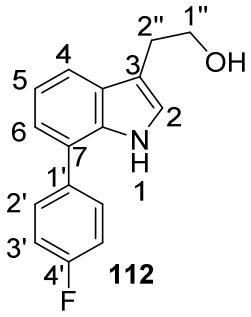
7.3.8.4 2-amino-4'-(tert-butyl)-5-phenyl benzamide (123)[193]



123

Prepared from anthranilamide (**62**) using PyrH CLEAs, K_3PO_4 and 4-tertbutylphenyl boronic acid according to the general method above (7.3.7.5). After cooling, reaction was basified using 4N NaOH and extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as a white solid (14 mg, 65 %); ^{1}H NMR (400 MHz, DMSO) δ 7.97 (1H, br s, CON$\underline{H_2}$), 7.84 (1H, d, *J* = 2.0 Hz, H6), 7.56 (2H, d, *J* = 8.0 Hz, H2'), 7.49 (1H, dd, *J* = 8.4, 2.0 Hz, H4), 7.41 (2H, d, *J* = 8.0 Hz, H3'), 7.18 (1H, br s, CON$\underline{H_2}$), 6.81 (1H, d, *J* = 8.4 Hz, H3), 1.30 (9H, s, C$\underline{Me_3}$); ^{13}C NMR (201 MHz, DMSO) δ 171.1 ($\underline{C}$=O), 158.2 (ArC), 158.1 (ArC), 148.5 (ArC), 137.2 (ArC), 130.0 (C4), 126.7 (ArCH), 125.5 (ArCH), 125.4 (ArCH), 117.5 (C3), 34.1 ($\underline{C}Me_3$), 31.2 (C$\underline{Me_3}$); *m/z* (ESI) 252.1 ([M-NH_2]$^+$, 100 %) 269.1 ([M+H]$^+$, 30); HRMS *m/z* calcd for $C_{17}H_{18}NO$ ([M-NH_2]$^+$) 252.1383, found: 252.1385; λ_{max} (H_2O/MeCN) 210, 283, 347 nm.

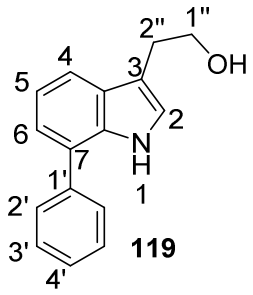
7.3.8.5 4'-fluoro-5-phenyl tryptophol (112)[193]



112

Prepared from tryptophol (**104**) using RebH CLEAs, K_2CO_3 and 4-fluorophenyl boronic acid according to the general method above (7.3.6.5). After cooling, reaction was extracted into CH_2Cl_2 (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (16.3 mg, 71 %). Spectra were consistent with those reported herein.

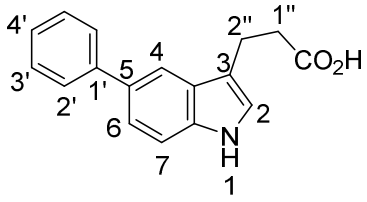
7.3.8.6 7-phenyl tryptophol (119)[193]



119

Prepared from tryptophol (**104**) using RebH CLEAs, K_2CO_3 and phenyl boronic acid according to the general method above (7.3.6.5). After cooling, reaction was extracted into CH_2Cl_2 (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (11.5 mg, 54 %); ^{1}H NMR (400 MHz, CD_3CN) δ 9.15 (1H, s, H1), 7.65 (2H, m, H2'), 7.59 (1H, d, *J* = 8 Hz, H6), 7.52 (2H, m, H3'), 7.42 (1H, m, H4'), 7.19 – 7.13 (2H, m, 2ArH), 7.12 (1H, m, H2), 3.78 (2H, t, *J* = 7.0 Hz, H1''), 2.96 (2H, m, H2''); ^{13}C NMR (101 MHz, MeOD) δ 141.0, 135.4, 129.9, 129.8, 129.4 128.1, 127.0, 124.2, 122.4, 120.2, 118.7, 113.2, 63.7 (C1''), 29.8 (C2''); *m/z* (ESI) 238.1 ($[M+H]^+$); HRMS *m/z* calcd for $C_{16}H_{16}NO$ ($[M+H]^+$) 238.1226, found: 238.1224; λ_{max} (H_2O/CH_3CN) 261, 352 nm.

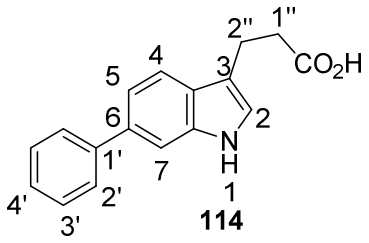
7.3.8.7 5-phenyl indole-3-propionic acid (113)[193]



113

Prepared from indole-3-propionic acid (**106**) using PyrH CLEAs, K_2CO_3 and phenyl boronic acid according to the general method above (7.3.6.5). After cooling, pH was adjusted to pH 1 using 1M HCl and extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (16.2 mg, 68 %); ^{1}H NMR (400 MHz, MeOD) δ 7.76 (1H, m, H4), 7.65 (2H, m, H2'), 7.42 – 7.38 (4H, 4 x ArH), 7.26 (1H, m, H4'), 7.09 (1H, s, H2), 3.11 (2H, t, *J* = 7.5 Hz, H1''), 2.70 (2H, td, *J* = 7.6, 2.6 Hz, H2''); ^{13}C NMR (101 MHz, MeOD) δ 178.1 (CO), 130.3, 130.2, 128.8, 128.7, 128.0, 127.6, 124.3, 122.6, 120.1, 118.1, 113.1, 111.1, 36.7 (C1''), 22.4 (C2''); *m/z* (ESI) 266.2 ($[M+H]^+$, 100); HRMS *m/z* calcd for $C_{17}H_{14}NO_2$ ($[M-H]^-$) 264.1030, found: 264.1032; λ_{max} (H_2O/CH_3CN) 286, 324 nm.

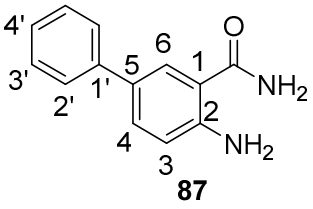
7.3.8.8 6-phenyl indole-3-propionic acid (114)[193]



114

A 30 mL bromination of indole-3-propionic acid (**106**) using SttH CLEAs was assembled as per the general method above (7.3.6.5) and allowed to proceed at room temperature overnight with shaking. CLEA was then removed using centrifugation (10000 rpm, 30 min, 4 ^{o}C) before re-suspension into fresh reaction buffer containing indole-3-propionic acid and further incubation at room temperature overnight with shaking. This process was then repeated once more before combination of the bromination supernatants and addition of phenyl boronic acid (5 eq) and K_2CO_3 (3 eq). The reaction was then freeze-thaw degassed and backfilled with N_2 before addition of tppts (20 mol %) and Na_2PdCl_4 (10 mol %) as deoxygenated stock solutions in water and heating at 80 ^{o}C overnight. After cooling, pH of the reaction was adjusted to pH 1 using 1M HCl and extracted into EtOAc (3 x 30 mL) before concentration of the combined organics *in vacuo*. Purification by semi-preparative HPLC method 5 afforded the title compound as an off-whtie solid (53.7 mg, 75 %); ^{1}H NMR (400 MHz, d_7-ethanol) δ 7.70-7.69 (3H, m, H2'+2), 7.64 (1H, d, *J* = 8.3 Hz, H4), 7.45 (2H, t, *J* = 8 Hz, H3'), 7.34 (1H, t, *J* = 8 Hz, H4'), 7.30 (1H, m, H5), 6.49 (1H, m, H7), 4.36 (2H, d, *J* = 6.9 Hz, H1''), 2.76 (2H, t, *J* = 6.9 Hz, 2''); ^{13}C NMR (101 MHz, MeOD) δ 177.4 (CO), 129.6, 129.6, 128.1, 128.1, 127.4, 127.0, 123.7, 122.0, 119.5, 117.5, 112.5, 110.5, 36.1 (C1''), 21.7 (C2''); *m/z* (ESI) 264.1 ([M-H]$^-$); HRMS *m/z* calcd for $C_{17}H_{14}NO_2$ ([M-H]$^-$) 264.1030, found: 264.1019; λ_{max} (H_2O/CH_3CN) 256, 312 nm.

7.3.8.9 Large Scale Arylation of 2-amino benzamide (62)[293]



87

A biotransformation was prepared using anthranilamide (**62**) (5.0 mM) and PyrH CLEAs from 3 L of culture in a total reaction volume of 200 mL. The biotransformation was conducted at room temperature of a total of 8 days, with addition of isopropanol (5 % v/v) every 48 hrs. After removing CLEA by centrifugation (10000 rpm, 30 min, 4 ^{o}C), phenyl boronic acid (5 eq) and K_3PO_4 (10 eq) were added to the supernatant prior to degassing under a stream of nitrogen. After backfill with N_2, tppts (0.2 mM) and Na_2PdCl_4 (0.1 mM) were then added under a positive pressure of nitrogen prior to heating at 80 ^{o}C for 24 hrs. After workup and purification as described above, purification by flash chromatography (SiO_2; Et_2O) afforded **87** as a white solid (110 mg, 52 %). Spectra were consistent with those reported herein.

7.3.9 PDMS Compartmentalisation

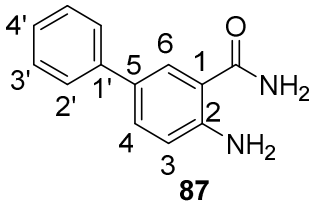
7.3.9.1 PDMS Thimble Preparation

Thimbles were prepared using a method based on those of Bowden and Gröger.[36,37,39] PDMS elastomer and curing agents were purchased as a Sylgard 184 kit from Sigma Aldrich and used as received. Elastomer and curing agents were combined and thoroughly mixed using a spatula before degassing under a static vacuum for 1 hr. Glass vials (ca. 15 mm x 34 mm) were plasma cleaned (15 min) before being placed under a static vacuum in a desiccator containing 1 mL of trichloro(1H,1H,2H,2H-perfluorooctyl)silane as silanizer for 2 hrs. Coated vials were then dipped in the degassed PDMS mixture and partially cured upside down by heating at 70 ^{o}C for 1 hr. Vials were then dipped into the PDMS mixture again and cured overnight at 70 ^{o}C to allow cross-linking. A razor blade was then used to cut around the top of the vials before being swelled by soaking in hexane for 30 min. After carefully separating them from the vial, complete PDMS thimbles were soaked in CH_2Cl_2 twice for 2 hrs. Vials were then dried by heating at 70 ^{o}C overnight before use.

7.3.9.2 PDMS Flux Determination

To a beaker containing either 5-bromo anthranilamide (**84**) (2.5 mM), 7-bromo tryptophol (**105**) (2.5 mM), NADH (2 mM), FAD (22 µM) or glucose (20 mM) in 10 mM potassium phosphate buffer to a total volume of 10 mL was added a thimble prepared above containing 1 mL of 10 mM potassium phosphate buffer. The breaker containing thimble was then incubated at room temperature with orbital shaking and 100 µL aliquots taken from both inside and outside the thimble at the timepoints specified. In the case of **L1**$_2$.Pd(OAc)$_2$, catalyst solution was placed inside the thimble (10 mM) to a total volume of 1 mL in H_2O and placed into a beaker containing 10 mM potassium phosphate buffer to total volume of 10 mL and 100 µL aliquots taken as stated above. All aliquots were then transferred to a UV-clear 96 well plate and UV spectra of each well measured. Concentration of each component at each timepoint was then determined using calibration curves of each component.

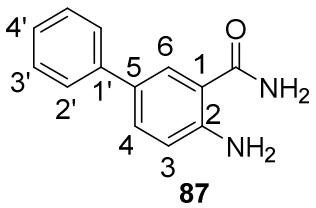
7.3.9.3 2-amino-5-phenyl benzamide (87)[193,293]



87

Prepared from anthranilamide (**62**) (2.0 mM) using purified PyrH, CsF (20 mM), phenyl boronic acid (10 mM) and **L1$_2$**.Pd(OAc)$_2$ (0.2 mM) according to the general method above (7.3.6.6). Purification by semi-preparative HPLC method 3 afforded the title compound as an off-white solid (9.4 mg, 74 %). Spectra were consistent with those reported herein.

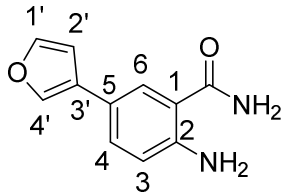
7.3.9.4 2-amino-5-phenyl benzamide (87)[193,293]



87

Prepared from anthranilamide (**62**) (5.0 mM) using PyrH CLEAs, CsF (50 mM), phenyl boronic acid (25 mM) and **L1$_2$**.Pd(OAc)$_2$ (0.1 mM) according to the general method above (7.3.6.7). Purification by semi-preparative HPLC method 3 afforded the title compound as an off-white solid (11.3 mg, 59 %). Spectra were consistent with those reported herein.

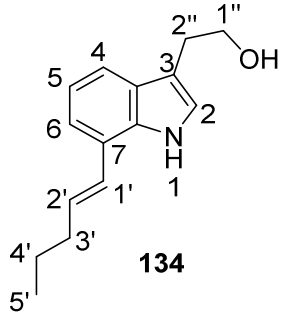
7.3.9.5 2-amino-5-(furan-3-yl) benzamide (128)[193]



128

Prepared from anthranilamide (**62**) (5.0 mM) using PyrH CLEAs, CsF (50 mM), 3-furanylboronic acid (25 mM) and **L1$_2$**.Pd(OAc)$_2$ (0.5 mM) according to the general method above (7.3.6.7). Purification by semi-preparative HPLC method 3 afforded the title compound as an orange solid (19.2 mg, 64 %); ^{1}H NMR (400 MHz, MeOD) δ 8.00 (1H, d, *J* = 2.0 Hz, H6), 7.95 (1H, s, H2'), 7.68 (1H, dd, *J* = 8.4, 2.0 Hz, H4), 7.58 (1H, m, H4'), 7.17 (1H, d, *J* = 8.4 Hz, H3), 6.87 (1H, m, H5'); ^{13}C NMR (101 MHz, MeOD) δ 174.4 (C1''), 149.9 (ArC), 144.6 (C4'), 138.6 (C2'), 131.2 (C4), 127.4 (ArC), 126.7 (ArC), 122.0 (C3), 118.8 (ArC), 115.8 (ArC), 109.4 (C5'); *m/z* (ESI) 186.1 ([M-NH$_3$]$^+$, 80 %), ([M+H]$^+$, 60 %), 225.1 ([M+Na]$^+$, 100); HRMS *m/z* calcd for $C_{11}H_{10}O_2N_2$ 202.0742, found: 202.0740; λ_{max} (H_2O/CH_3CN) 229, 292 nm.

7.3.9.6 7-(pent-1-en-1-yl) tryptophol (134)[193]

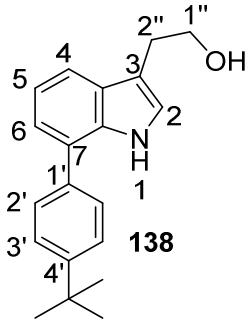


134

Prepared from tryptophol (**104**) (5.0 mM) using RebH CLEAs, CsF (50 mM), E-pent-1-en-1-yl boronic acid (25 mM) and **L1$_2$**.Pd(OAc)$_2$ (0.5 mM) according to the general method above (7.3.6.7). Purification by semi-preparative HPLC method 4 afforded the title compound as a white solid (19.1 mg, 61 %); ^{1}H NMR (400 MHz, MeOD) δ 7.41 (1H, dd, *J* = 7.4, 0.8 Hz, H6), 7.17 (1H, d, *J* = 7.4 Hz, H4), 7.08 (1H, s, H2), 6.97 (1H, t, *J* = 7.6 Hz, H5), 6.76 (1H, d, *J* = 16 Hz, H1'), 6.32 (1H, dt, *J* = 16, 7.2 Hz, H2'), 3.80 (2H, t, *J* = 7.2 Hz, H1''), 2.96 (2H, t, *J* = 7.2 Hz, H2''), 2.28 (2H, qd, *J* =

7.2, 1.6 Hz, H3'), 1.54 (2H, sextet, *J* = 7.6 Hz, H4'), 1.01 (3H, t, *J* = 7.6 Hz, H5'); ^{13}C NMR (101 MHz, MeOD) δ 135.6 (ArC), 131.9 (ArC), 129.5 (ArC), 127.4 (ArC), 123.7 (ArC), 123.1 (ArC), 119.9 (ArC), 119.2 (ArC), 118.1 (ArC), 113.0 (ArC), 63.7, 61.2, 38.8, 36.7, 29.8, 23.8, 14.1; *m/z* (ESI) 230.1 ([M+H]$^+$); HRMS *m/z* calcd for $C_{15}H_{20}NO$ 230.1539, found: 230.1549; λ_{max} (H_2O/CH_3CN) 238, 310 nm.

7.3.9.7 4'-tertbutyl-7-phenyl tryptophol (138)[193]



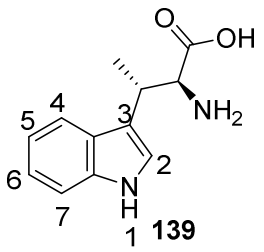
Prepared from tryptophol (**104**) (5.0 mM) using RebH CLEAs, CsF (50 mM), 4-tertbutylphenyl boronic acid (25 mM) and **L1**$_2$.Pd(OAc)$_2$ (0.5 mM) according to the general method above (7.3.6.7). Purification by flash chromatography (SiO_2:Et_2O) afforded the title compound as a white solid (26.8 mg, 61 %); ^{1}H NMR (400 MHz, MeOD) δ 7.59 – 7.49 (5H, m, 5ArH), 7.15 – 7.04 (3H, m, 3ArH), 3.83 (2H, t, *J* = 7.4 Hz, H1''), 3.01 (d, *J* = 7.4 Hz, H2''), 1.38 (9H, s, C<u>Me$_3$</u>); ^{13}C NMR (101 MHz, MeOD) δ 151.09 (ArC), 138.02 (ArC), 135.45 (ArC), 129.7 (ArC), 129.0 (Ar<u>C</u>H), 126.9 (ArC), 126.8 (Ar<u>C</u>H), 124.1 (Ar<u>C</u>H), 122.3 (Ar<u>C</u>H), 120.2 (Ar<u>C</u>H), 118.5 (Ar<u>C</u>H), 113.1 (ArC), 63.7 (C1''), 35.4 (<u>C</u>Me$_3$), 31.8 (C<u>Me$_3$</u>), 29.9 (C2''); *m/z* (ESI) 294.2 ([M+H]$^+$); HRMS calcd for $C_{20}H_{24}NO$ 294.1852, found: 294.1911; λ_{max} (H_2O/CH_3CN) 224, 285 nm.

<u>7.3.10 Preparation of β-methyl Tryptophan Derivatives</u>

7.3.10.1 Preparation of *E. coli* Lysate Containing Tryptophan Synthase

As previously described by our group, a plasmid containing the gene encoding the two-unit tryptophan synthase from *Salmonella enterica* Typhimurium in pSTB7 (ATCC 37845) was used for preparation of the βL116V tryptophan synthase mutant by Daniel Francis.[203] This construct was then transformed into *E. coli* BL21 (DE3) using ampicillin (100 μg/mL) for selection. Single colonies of transformant were picked and used to inoculate 10 mL of LB containing ampicillin (100 μg/mL) and grown at 37 °C overnight. This culture was then diluted 1:100 into 400 mL of fresh LB containing ampicillin (100 μg/mL) and incubated at 37 °C for 24 hrs. Cells were then harvested by centrifugation (4000 rpm, 15 min, 4 °C) and washed with 100 mL 100 mM potassium phosphate buffer (pH 7.8) before suspension in 20 mL lysis buffer containing pyridoxal 5'-phosphate (PLP) in 100 mM potassium phosphate (pH 7.8). Cells were then lysed by sonication (10 min, 50 % pulse, 70 % power, 700 W) and clarified by centrifugation (10000 rpm, 45 min, 4 °C). Resultant lysate was then either used immediately or stored at -80 °C until further use.

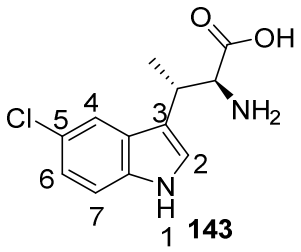
7.3.10.2 (2*S*,3*S*)-β-methyl Tryptophan (139)[203]



139

To a solution containing indole (2.0 mM) and L-threonine (6.0 mM) in 100 mM potassium phosphate buffer (pH 7.8) to total volume 100 mL was added two dialysis bags (12 – 14 kDa MWCO) containing 5 mL of βL116V mutant tryptophan synthase lysate prepared as above (7.3.10.1) each. The biotransformation was then incubated at 37 °C for 72 hrs. Dialysis bags were then soaked in water (3 x 25 mL) and the extract combined with the remaining reaction buffer. After washing with EtOAc (3 x 100 mL), the aqueous phase was concentrated to approximately 30 mL and applied to a C18 Bond Elute® (Agilent, 20 g mesh) which had previously been activated with 3 column lengths (CL) of methanol and equilibrated with 3 CL H_2O . After washing with water (300 mL), organics were eluted with MeOH (3 CL). Removal of solvent *in vacuo* afforded the title compound as an off-white solid (23 mg, 51 %); ^{1}H NMR (400 MHz, MeOD) δ 7.71 (1H, d, *J* = 8.0 Hz, H4), 7.35 (1H, d, *J* = 8.0 Hz, H7), 7.19 (1H, s, H2), 7.10 (1H, td, *J* = 8.0, 0.8 Hz, H6), 7.02 (1H, td, *J* = 8.0, 0.8 Hz, H5), 3.78 (1H, d, *J* = 7.4 Hz, Hα), 3.56 (1H, p, *J* = 7.4 Hz, Hβ), 1.54 (3H, d, *J* = 7.4 Hz, βMe); *m/z* (ESI) 219 ($[M+H]^+$ 100 %); HRMS *m/z* calcd for $C_{12}H_{15}N_2O_2$ 219.1128 found: 219.1114. Enantiopurity was determined by Daniel Francis as previously described.

7.3.10.3 (2*S*,3*S*)-5-chloro-β-methyltryptophan (143)[203]

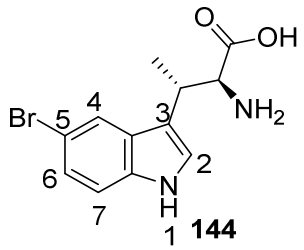


143

Prepared from (2*S*,3*S*)-β-methyl tryptophan (**139**) using PyrH CLEAs from a total of 3 L of culture according to the general method described above (7.3.6.1), with $MgCl_2$ (30 mM) in place of NaBr, and allowed to proceed at room temperature overnight with orbital shaking. CLEA was then removed by centrifugation (10000 rpm, 30 min, 4 °C) and washed with 30 mL 100 mM potassium phosphate buffer (pH 7.2) twice. Combined supernatants were then concentrated *in vacuo* to ca 30 mL and loaded onto a C18 Bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of MeOH and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted with MeOH (3 CL) and organics dried *in vacuo*. Purification by HPLC method 3 afforded the title compound as a white solid (14 mg, 62 %); ^{1}H NMR (400 MHz, MeOD) δ 7.62 (1H, d, *J* = 2 Hz, H4), 7.36 (1H, d, *J* = 8.6 Hz, H7), 7.29 (1H, s, H2), 7.10 (1H, dd, *J* = 8.6, 2.0 Hz, H6), 4.16 (1H, d, *J* = 5.7 Hz, Hα), 3.80 – 3.71 (1H, m, Hβ), 1.56 (3H, d, *J* = 7.4 Hz, βMe); ^{13}C NMR (101 MHz, MeOD) δ 171.3 (C=O), 136.7 (ArC), 128.6 (ArC), 126.1 (ArC), 126.0 (ArC), 123.1 (C6), 118.9 (C4), 113.9 (ArC), 113.6 (ArC), 58.91 (Hα), 33.8 (Hβ), 18.5 (βMe); *m/z* (ESI) 253.1 ($[M^{35}Cl+H]^+$, 100 %),

255.1 ([$M^{37}Cl+H]^+$, 30); HRMS *m/z* calcd for $C_{12}H_{14}{}^{35}ClN_2O_2$ 253.0738 found: 253.0734; $[\alpha]_D{}^{25}$ +19.61 (c = 1.02, 0.1 M HCl).

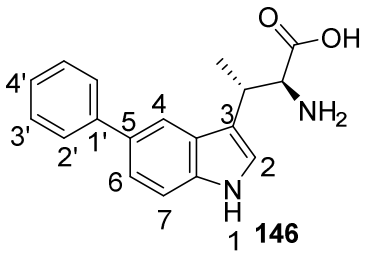
7.3.10.4 (2*S*,3*S*)-5-bromo-β-methyltryptophan (144)[203]



144

Prepared from (2*S*,3*S*)-β-methyl tryptophan (**139**) using PyrH CLEAs from a total of 3 L of culture and NaBr (30 mM) according to the general method described above (7.3.6.1) and allowed to proceed at room temperature overnight with orbital shaking. CLEA was then removed by centrifugation (10000 rpm, 30 min, 4 ^oC) and washed with 30 mL 100 mM potassium phosphate buffer (pH 7.2) twice. Combined supernatants were then concentrated *in vacuo* to ca 30 mL and loaded onto a C18 Bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of MeOH and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted with MeOH (3 CL) and organics dried *in vacuo*. Purification by HPLC method 3 afforded the title compound as a white solid (11.4 mg, 41 %); ^1H NMR (400 MHz, MeOD) δ 7.76 (1H, d, *J* = 2.0 Hz, H4), 7.34 (1H, d, *J* = 8.5 Hz, H7), 7.28 – 7.23 (2H, m, H2H6), 4.28 (1H, d, *J* = 4.4 Hz, Hα), 3.86 (1H, m, Hβ), 1.50 (3H, d, *J* = 7.4 Hz, βMe); ^{13}C NMR (201 MHz, MeOD) δ 171.3 (C=O), 137.0 (ArC), 129.3 (ArC), 125.9 (ArC), 125.7 (ArC), 122.1 (C4), 114.3 (ArCH), 113.7 (ArC), 113.4 (ArC), 59.0 (Hα), 33.8 (Hβ), 18.5 (βMe); *m/z* (ESI) 295.0 ([$M^{79}Br+H]^+$, 50 %), 297.0 ([$M^{81}Br+H]^+$, 50); HRMS *m/z* calcd for $C_{12}H_{12}BrN_2O_2$ 295.0088 found: 295.0078; $[\alpha]_D{}^{25}$ +14.85 (c = 2.02, 0.1 M HCl).

7.3.10.5 (2*S*,3*S*)-5-phenyl-β-methyltryptophan (146)[203]



146

A biotransformation was prepared using (2*S*,3*S*)-β-methyl tryptophan (**139**), PyrH CLEAs from a total of 3 L of culture and NaBr (30 mM) as reported above (7.3.10.4). After two days, CLEA was removed by centrifugation (10000 rpm, 30 min, 4 ^oC) before addition of phenyl boronic acid (55 mg, 0.45 mmol, 5 eq) and tripotassium phosphate (478 mg, 2.25 mmol, 25 eq) to the supernatant and degassing with a stream of N_2. sSPhos (7 mg, 13.5 μmol, 15 mol %) and Na_2PdCl_4 (1.3 mg, 4.5 μmol, 5 mol %) were then dissolved in deoxygenated water and heated to 45 ^oC for 15 min prior to transfer to the biotransformation supernatant. The reaction was then heated at 80 ^oC with stirring under N_2 overnight. After cooling, the reaction was then loaded onto a C18 bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of methanol and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted with MeOH (3 CL) and organics dried *in vacuo*.

Purification by HPLC method 3 afforded the title compound as an off-white solid 16.4 mg, 63 %); ^{1}H NMR (800 MHz, MeOD) δ 7.85 (1H, d, *J* = 1.5, H4), 7.67 – 7.65 (2H, m, 2ArH), 7.47 – 7.40 (4H, m, 4ArH), 7.29 – 7.27 (2H, m, H4'H2), 4.21 (1H, d, *J* = 5.6 Hz, Hα), 3.87 (1H, m, Hβ), 1.61 (3H, d, *J* = 7.4 Hz, βMe); ^{13}C NMR (201 MHz, MeOD) δ 171.4 (C=O), 144.0 (ArC), 137.9 (ArC), 134.1 (ArC), 129.7 (ArCH), 128.2 (ArCH), 128.0 (ArC), 127.3 (ArCH), 125.1 (ArCH), 122.7 (ArCH), 117.9 (C4), 113.0 (ArC), 59.1 (Hα), 33.9 (Hβ), 18.8 (βMe); *m/z* (ESI) 295.2 ([M+H]$^{+}$, 100 %); HRMS *m/z* calcd for $C_{18}H_{17}N_2O_2$ 293.1296 found: 293.1287.

7.3.10.6 L-Amino Acid Oxiase (LAAO) Assay

Type I LAAO from *Crotalus adamanteus* was obtained from Sigma Aldrich as dried venom and used as received. The LAAO powder was dissolved in 100 mM potassium phosphate buffer (pH 7.2) to final concentration 2 mg/mL before addition of substrate (0.5 mM) to a total volume of 200 μL and incubation at 37 °C with shaking at 800 rpm overnight. Protein was then precipitated by heating (95 °C, 10 min) and removed by centrifugation (13000 rpm, 10 min) prior to analysis of the supernatant by analytical HPLC method 1 and LC-MS.

7.3.10.7 L-Amino Acid Deaminase (LAAD) Preparation

Clones of *E. coli* BL21 (DE3) transformed with a pET28a construct containing LAAD from *Proteus mirabilis* were kindly provided by Fabio Parmeggiani of the Turner Group (University of Manchester).[209] These clones were used to inoculate 10 mL of LB medium containing kanamycin (50 μg/mL) and incubated at 37 °C overnight. This culture was then diluted 1:100 into 800 mL of LB medium containing kanamycin (50 μg/mL) and incubated at 37 °C with shaking at 180 rpm until OD_{600} = 0.6. After induction with IPTG (1 mM final) cultures were incubated at 30 °C with shaking at 180 rpm for a further 5 hrs before harvesting by centrifugation (4000 rpm, 20 min, 4 °C). After washing with 100 mM potassium phosphate buffer (pH 7.2), cells were either re-suspended into reaction buffer or stored at -20 °C prior to use.

7.3.10.8 L-Amino Acid Deaminase (LADD) Assay

In-tact *E. coli* BL21 (DE3) cells expressing LAAD prepared above (30 – 100 mg/mL) were suspended in reaction buffer containing substrate (0.5 mM) to a total volume of 200 μL in 100 mM potassium phosphate buffer (pH 8.0). After incubation at 30 °C with shaking at 1000 rpm overnight, MeOH (200 μL) was added and the suspension vigorously shaken for 10 min prior to clarification by centrifugation (13000 rpm, 20 min) and analysis of the supernatant by analytical HPLC method 1 and LC-MS.

7.3.11 Alkynylation Methods

7.3.11.1 General Method for Integrated Alkynylation Screening

A biocatalytic bromination of either anthranilamide (**62**) or tryptophol (**104**) using PyrH (anthranilamide) or RebH (tryptophol) CLEAs was performed as described above (7.3.6.1) and allowed to proceed at room temperature with orbital shaking overnight. After removal of the CLEA by centrifugation (10000 rpm, 30 min, 4 °C), alkynyl boron reagent (5 eq) and K_3PO_4 (10 eq) were

added prior to freeze-thaw degassing and backfill with N_2. Tppts (20 mol %) and Na_2PdCl_4 (10 mol %) were then added as deoxygenated stock solutions in water prior to heating at 80 ^oC overnight. Upon cooling, the reaction was partitioned into EtOAc (anthranilamide) or CH_2Cl_2 (tryptophol) and solvent removed *in vacuo*. The residue was then dissolved in MeOH and analysed by analytical HPLC method 2 and LC-MS.

7.3.11.2 General Method for Alkynylation using Pure Aryl Bromides

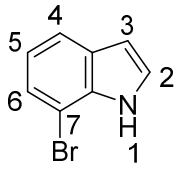
Either 2-amino-5-bromo anthranilamide (**84**) or 7-bromo tryptophol (**105**) were subjected to the SMC conditions specified above (7.3.2.2) using alkynyl boron species (3 eq) in place of aryl boronic acid. Upon workup, the crude was analysed by LC-MS, ^1H NMR, ^{13}C NMR and analytical HPLC method 2.

7.4 Chapter 4 Experimental

Original copies of the spectra recorded of novel compounds prepared in this chapter can be found in Appendix Two.

7.4.1 Substrate and Standards Synthesis

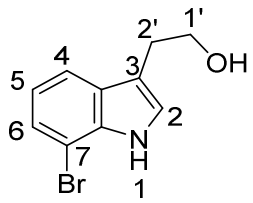
7.4.1.1 7-bromoindole (158)[235]



158

2-bromonitrobenzene (1.0 g, 5.0 mmol, 1.0 eq) was added to an oven-dried flask prior to backfill with N_2. Anhydrous THF (10 mL) was then added and the solution cooled to -42 ^oC with stirring. A 1 M solution of vinyl magnesium bromide in THF (15 mL, 15 mmol, 3.0 eq) was then added dropwise and stirring continued for 1 hr at -42 ^oC. The reaction was then slowly poured into saturated NH_4Cl (30 mL) and extracted into EtOAc (3 x 50 mL). The combined organic phase was then washed with brine (100 mL) and dried over $MgSO_4$ before concentration *in vacuo*. Purification by flash chromatography (19:1 PE:EtOAc) afforded the title compound as a white solid (289 mg, 30 %); ^1H NMR (400 MHz, $CDCl_3$) δ 8.27 (1H, br. s, NH), 7.60 (1H, d, *J* = 8.0 Hz, ArH), 7.37 (1H, d, *J* = 7.6 Hz, ArH), 7.19 (1H, dd, *J* = 3.2, 2.2 Hz, ArH), 7.02 (1H, dd, *J* = 8.0, 7.6 Hz, H5), 6.64 (1H, dd, *J* = 3.2, 2.2 Hz, ArH); ^{13}C NMR (101 MHz, $CDCl_3$) δ 134.7 (ArC), 129.1 (ArC), 124.8 (ArCH), 124.4 (ArCH), 121.1 (ArCH), 120.1 (ArCH), 104.8 (ArC), 103.9 (ArCH); *m/z* (ESI) 195.9 ($[M^{79}Br+H]^+$, 50 %), 197.9 ($[M^{81}Br+H]^+$, 50); HRMS *m/z* calcd for $C_8H_7{}^{79}BrN$ 195.9756 found: 195.9754.

7.4.1.2 7-bromotryptophol (105)[236]

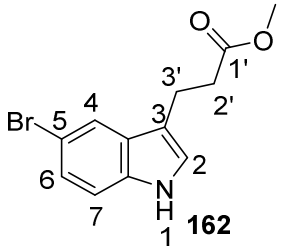


105

7-bromoindole (**158**, 196 mg, 1.0 mmol, 1.0 eq) was added to an oven-dried flask and backfilled with N_2 prior to addition of anhydrous Et_2O (20 mL) and cooling on ice. Oxalyl chloride (262 µL, 3.0 mmol, 3.0 eq) was then added and the resultant suspension stirred at room temperature for 1 hr prior to addition of anhydrous MeOH (217 µL, 5.0 mmol, 5.0 eq). After stirring at room temperature for 1 hr, solvent was removed *in vacuo*. After backfill with N_2, the residue was dissolved in anhydrous THF (5 mL) prior to cooling on ice and dropwise addition of $LiAlH_4$ as a 2 M solution in THF (2 mL, 4.0 mmol, 4 eq). The reaction was then warmed to room temperature and stirred overnight. After cooling on ice, the reaction was quenched by dropwise addition of H_2O (15 mL) followed by 1 M NaOH (5 mL). The reaction was then partitioned into EtOAc (3 x 30 mL) and combined organics dried over $MgSO_4$ prior to removal of solvent *in vacuo*. Purification by flash

chromatography (SiO_2;1:1 hexane:EtOAc) afforded the title compound as a brown oil (44 mg, 18 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (1H, br. s, NH), 7.57 (1H, d, *J* = 8.0, H4), 7.36 (1H, d, *J* = 8.0, H6), 7.16 (1H, d, *J* = 2.0 Hz, H2), 7.01 (t, *J* = 8.0, H5), 3.91 (2H, td, *J* = 6.5, 2.0 Hz, H2'), 3.02 (2H, t, *J* = 6.5 Hz, H1'); *m/z* (ESI) 240.0 ([^{79}M+H]$^+$, 50 %), 242.0 ([^{81}M+H]$^+$, 50); HRMS *m/z* calcd for $C_{10}H_{11}{}^{79}BrNO$ 240.0019 found: 240.0016.

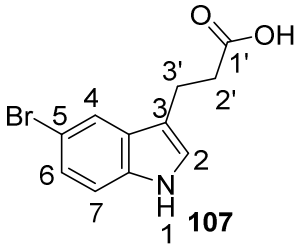
7.4.1.3 Methyl 5-bromoindole-3-propionate (162)[237]



162

To a stirred solution of 5-bromoindole (392 mg, 2.0 mmol, 1.0 eq) in CH_2Cl_2 (10 mL) was added methyl acrylate (360 µL, 4.0 mmol, 2.0 eq) followed by $ZrCl_4$ (236 mg, 1.0 mmol, 50 mol %). After stirring at room temperature for 3 hrs, water (20 mL) was added and the reaction extracted into EtOAc (3 x 20 mL). The combined organics were then washed with brine (50 mL) and dried over $MgSO_4$ prior to concentration *in vacuo*. Purification by flash chromatography (SiO_2; 7:3 hexane:EtOAc) afforded the title compound as an off-white solid (238 mg, 42 %); ν_{max}/cm^{-1} 3408 (N-H), 2961 (C-H), 2922 (C-H), 2849 (C-H), 1560 (s, C=O), 1480 (m, C=C), 1402 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.01 (1H, br. s, NH), 7.68 (1H, d, *J* = 2.2, H4), 7.23 (1H, dd, *J* = 8.2, 2.2 Hz, H6), 7.24 (1H, d, *J* = 8.2 Hz, H7), 7.04 (1H, d, *J* = 2.2 Hz, H2), 3.72 (3H, s, CH_3), 3.09 (2H, td, *J* = 7.2, 2.2 Hz, H3'), 2.74 (2H, t, *J* = 7.2 Hz, H2'); ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.8 (C1'), 135.0 (ArC), 129.1 (ArC), 125.0 (C6), 122.9 (C2), 121.5 (C4), 114.8 (ArC), 112.8 (ArC), 112.7 (C7), 51.8 ($\underline{C}H_3$), 34.8 (C2'), 20.9 (C3'); *m/z* (ESI) 252.0 ([^{79}M+H]$^+$, 50 %), 254.0 ([^{81}M+H]$^+$, 50); HRMS *m/z* calcd for $C_{12}H_{13}{}^{79}BrNO_2$ 282.0124 found: 282.0123.

7.4.1.4 5-bromoindole-3-propionoic acid (107)[237]

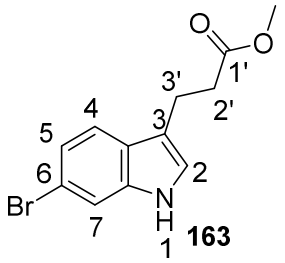


107

Methyl 5-bromoindole-3-propionoic acid (**162**) (160 mg, 0.57 mol, 1.0 eq) was dissolved in THF (2 mL) and H_2O (2 mL) prior to addition of $LiOH.H_2O$ (358 mg, 8.5 mmol, 15 eq) and stirring at room temperature overnight. The reaction was then washed with EtOAc (20 mL) and pH adjusted to pH 1 using 6 M HCl. The reaction was then extracted into EtOAc (3 x 30 mL) prior to washing of the combined organics with brine (100 mL) and drying over $MgSO_4$ before removal of solvent *in vacuo* to afford the title compound as a white solid (109 mg, 72 %); ν_{max}/cm^{-1} 3428 (N-H), 3038 (br., OH), 1697 (C=O), 1542 (C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 (1H, br. s, NH), 7.77 (1H, d, *J* = 1.6 Hz, H4), 7.34 – 7.28 (2H, m, 2ArH), 7.08 (1H, d, *J* = 2.4 Hz, H2), 3.11 (2H, td, *J* = 8.0, 2.4 Hz,

H3'), 2.80 (2H, t, *J* = 8.0 Hz, H2'); ^{13}C NMR (101 MHz, $CDCl_3$) δ 178.2 (C1'), 128.9 (ArC), 125.0 (ArCH), 122.8 (C2), 121.3 (C4), 114.4 (ArC), 112.7 (ArC), 112.6 (ArCH), 34.4 (C2'), 20.1 (C3'); *m/z* (ESI) 268.0 ([^{79}M+H]$^+$, 50 %), 270.0 ([^{81}M+H]$^+$, 50); HRMS *m/z* calcd for $C_{11}H_{11}{}^{79}BrNO_2$ 267.9968 found: 267.9964.

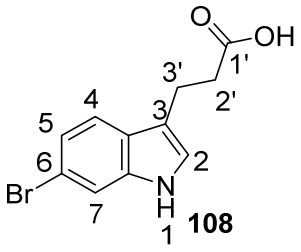
7.4.1.5 Methyl 6-bromoindole-3-propionate (163)



163

To a stirred solution of 6-bromoindole (196 mg, 1.0 mmol, 1.0 eq) in CH_2Cl_2 (5 mL) was added methyl acrylate (180 µL, 2.0 mmol, 2.0 eq) followed by $ZrCl_4$ (116 mg, 0.5 mmol, 50 mol %). After stirring at room temperature for 3 hrs, water (20 mL) was added and the reaction extracted into EtOAc (3 x 20 mL). The combined organics were then washed with brine (50 mL) and dried over $MgSO_4$ prior to concentration *in vacuo*. Purification by flash chromatography (SiO_2; 7:3 hexane:EtOAc) afforded the title compound as a brown solid (55.6 mg, 24 %); ν_{max}/cm^{-1} 3372 (N-H), 2962 (C-H), 2920 (C-H), 2852 (C-H), 1563 (s, C=O), 1432 (w, C=C), 1403 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.96 (1H, br. s, NH), 7.50 (1H, d, *J* = 1.6, H7), 7.45 (1H, d, *J* = 8.0 Hz, H4), 7.22 (1H, dd, *J* = 8.0, 1.6 Hz, H5), 6.99 (1H, d, *J* = 1.2 Hz, H2), 3.67 (3H, s, CH_3), 3.07 (2H, td, *J* = 7.4, 1.2 Hz, H3'), 2.70 (2H, t, *J* = 7.4 Hz, H2'); ^{13}C NMR (101 MHz, $CDCl_3$) δ 180.9 (C1'), 139.3 (ArC), 129.2 (ArC), 125.8 (ArC), 123.0 (ArCH), 122.7 (ArCH), 120.8 (ArCH), 114.3 (ArC), 114.0 (ArCH), 51.6 ($\underline{C}H_3$), 34.7 (C2'), 20.5 (C3'); *m/z* (ESI) 282.0 ([^{79}M+H]$^+$, 50 %), 284 ([^{81}M+H]$^+$, 50); HRMS *m/z* calcd for $C_{12}H_{13}{}^{79}BrNO_2$ 282.0124 found: 282.0122.

7.4.1.6 6-bromoindole-3-propionic acid (108)[294]

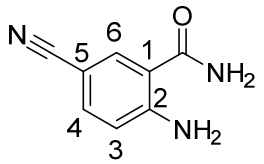


108

Methyl 6-bromoindole-3-propionic acid (**163**) (51 mg, 0.18 mol, 1.0 eq) was dissolved in THF (2 mL) and H_2O (2 mL) prior to addition of $LiOH.H_2O$ (111 mg, 2.66 mmol, 15 eq) and stirring at room temperature overnight. The reaction was then washed with EtOAc (20 mL) and pH adjusted to pH 1 using 6 M HCl. The reaction was then extracted into EtOAc (3 x 30 mL) prior to washing of the combined organics with brine (100 mL) and drying over $MgSO_4$ before removal of solvent *in vacuo* to afford the title compound as a white solid (32 mg, 67 %); ν_{max}/cm^{-1} 3427 (N-H), 3038 (br, O-H), 1699 (s, C=O), 1562 (C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 (1H, br. s, NH), 7.44 (d, *J* = 2.0 Hz, H7), 7.38 (1H, d, *J* = 8.4 Hz, H4), 7.15 (1H, dd, *J* = 8.4, 2.0 Hz, H5), 6.93 (1H, s, H2), 3.01 (2H, t, *J* = 7.4 Hz, H3'), 2.68 (2H, t, *J* = 7.4 Hz, H2'); ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.3 (C1'),

136.0 (ArC), 125.0 (ArC), 121.7 (C5), 121.0 (C2), 118.9 (C4), 114.7 (ArC), 113.9 (ArC), 113.0 (C7), 33.4 (C2'), 19.2 (C3'); *m/z* (ESI) 268.0 ($[^{79}M+H]^+$, 50 %), 270.0 ($[^{81}M+H]^+$, 50); HRMS *m/z* calcd for $C_{11}H_{11}BrNO_2$ 267.9968 found: 267.9965.

7.4.1.7 2-amino-5-cyano benzamide (155)



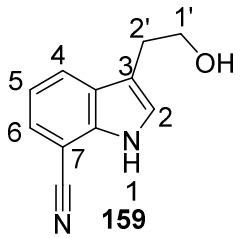
155

A flask was charged with 2-amino-5-bromo benzamide (**84**) (22 mg, 0.10 mmol, 1.0 eq), sodium carbonate (21 mg, 0.20 mmol, 2.0 eq), $K_4[Fe(CN)_6].3H_2O$ (17 mg, 40 µmol, 40 mol %), tppts (11 mg, 19 µmol, 19 mol %) and Na_2PdCl_4 (2.9 mg, 9.9 µmol, 9.9 mol %) prior to backfill with N_2 and addition of degassed H_2O (10 mL). The reaction was then stirred with heating at 85 °C overnight prior to cooling and extraction into EtOAc (3 x 10 mL). After drying the combined organics over $MgSO_4$, solvent was removed *in vacuo*. Purification by flash chromatography (SiO_2; 8:2 Et_2O:Hexane) afforded the title compound as a white solid (9.9 mg, 61 %); ν_{max}/cm^{-1} 3256 (br, CO<u>N-H</u>$_2$) 3192 (br, CO<u>N-H</u>$_2$), 2552 (<u>N-H</u>$_2$), 2501 (<u>N-H</u>$_2$), 2216 (m, C≡N) 1651 (s, C=O), 1611 (C=C), 1504 (C=C), 1410 (C=C); ^{1}H NMR (400 MHz, MeOD) δ 7.92 (1H, d, *J* = 1.8 Hz, H6), 7.42 (1H, dd, *J* = 8.8, 1.8 Hz, H4), 6.78 (d, *J* = 8.8 Hz, H3); ^{13}C NMR (101 MHz, MeOD) δ 172.6 (<u>C</u>=O), 155.0 (ArC), 136.1 (C4), 135.2 (C6), 120.8 (ArC), 118.2 (C3), 114.8 (<u>C</u>≡N) , 97.5 (C5); *m/z* (ESI) 145.0 ($[M-NH_2]^+$), 162.0 ($[M+H]^+$); HRMS *m/z* calcd for $C_8H_5N_2O$ ($[M-NH_2]^+$) 145.0396 found: 145.0389; λ_{max} (H_2O/CH_3CN) 232, 279, 328 nm.

7.4.1.8 General Method for the Cyanation of Indoles

A flask was charged with bromoindole (0.10 mmol, 1.0 eq), $K_4[Fe(CN)_6].3H_2O$ (21 mg, 50 µmol, 50 mol %) and ^{t}BuXPhos-Pd-G3 (**157**) (7.9 mg, 10 µmol, 10 mol %) prior to backfill with N_2. Deoxygenated THF (5 mL) was then added followed by deoxygenated H_2O (25 mL) and the resultant solution stirred vigorously at 80 °C overnight under N_2. After cooling, THF was removed under a stream of N_2 and the resultant suspension filtered and washed with 10 % MeOH in H_2O (10 mL). The filtrate was then applied to a C18 bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of methanol and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted using MeOH (3 CL) and solvent removed *in vacuo* prior to purification.

7.4.1.9 7-cyano trytophol (159)

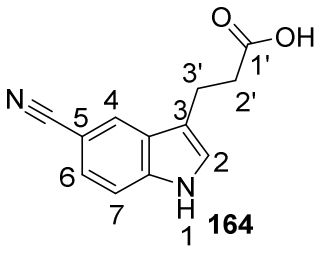


159

Prepared from 7-bromo tryptophol (**105**) according to the general method above (7.4.1.8). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (13.6

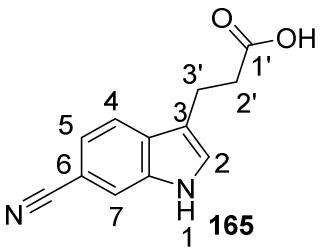
mg, 73 %); ν_{max}/cm^{-1} 3316 (br., O-H), 2935 (w, C-H), 2922 (w, C-H), 2893 (w, C-H), 2854 (w, C-H), 2225 (m, C≡N), 1449 (m, C=C); ^{1}H NMR (400 MHz, MeOD) δ 7.88 (1H, dd, *J* = 8.0, 1.0 Hz, H4), 7.48 (1H, dd, *J* = 7.6, 1.0 Hz, H6), 7.25 (1H, app. s, H2), 7.14 (1H, dd, *J* = 8.0, 7.6 Hz, H5), 3.81 (2H, t, *J* = 7.0 Hz, H1'), 2.98 (2H, td, *J* = 7.0, 1.0 Hz, H2'); ^{13}C NMR (101 MHz, MeOD) δ 137.7 (ArC), 130.1 (ArC), 127.2 (C4), 125.8 (C2), 125.1 (C6), 119.6 (C5), 118.4 (ArC), 114.7 (C≡N), 95.1 (C7), 63.4 (C1'), 29.4 (C2'); *m/z* (ESI) 187.1 ([M+H]$^{+}$); HRMS *m/z* calcd for $C_{11}H_{11}N_2O$ 187.0866 found: 187.0862; λ_{max} (H_2O/CH_3CN) 236, 314 nm.

7.4.1.10 5-cyano indole-3-propionic acid (164)



Prepared from 5-bromoindole-3-propionic acid (**107**) according to the general method above (7.4.1.8). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (15.4 mg, 72 %); ν_{max}/cm^{-1} 3367 (m, N-H), 3269 (br., O-H), 2918 (w, C-H), 2864 (w, C-H), 2214 (m, C≡N), 1708 (s, C=O), 1615 (m, C=C);^{1}H NMR (400 MHz, MeOD) δ 10.94 (1H, br. s, NH), 8.01 (1H, d, *J* = 0.8 Hz, H4), 7.47 (1H, d, *J* = 8.8, H7), 7.37 (1H, dd, *J* = 8.8, 0.8 Hz, H6), 7.24 (1H, d, *J* = 0.8 Hz, H2), 3.08 (2H, td, *J* = 7.4, 0.8 Hz, H3'), 2.68 (2H, t, *J* = 7.4 Hz, H2');^{13}C NMR (101 MHz, MeOD) δ 177.0 (C1'), 139.9 (ArC), 128.6 (ArC), 125.8 (ArCH), 125.3 (ArCH), 125.1 (ArCH), 122.0 (ArC), 116.5 (C≡N), 113.4 (C7), 102.2 (C5), 35.9 (C2'), 21.3 (C3'); *m/z* (ESI) 213.1 ([M-H]$^{-}$); HRMS *m/z* calcd for $C_{12}H_9N_2O_2$ 213.0670 found: 213.0670; λ_{max} (H_2O/CH_3CN) 242, 286 nm.

7.4.1.11 6-cyano indole-3-propionic acid (165)



Prepared from 6-bromoindole-3-propionic acid (**108**) according to the general method above (7.4.1.8). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (14.3 mg, 67 %); ν_{max}/cm^{-1} 3345 (m, N-H), 3124 (br., O-H), 2917 (w, C-H), 2850 (w, C-H), 2216 (m, C≡N), 1709 (s, C=O), 1664 (m, C=C), 1615 (m, C=C); ^{1}H NMR (400 MHz, MeOD) δ 7.73 (1H, d, *J* = 1.2 Hz, H7), 7.70 (1H, d, *J* = 8.2 Hz, H4), 7.35 (1H, s, H2), 7.28 (1H, dd, *J* = 8.2, 1.2 Hz, H5), 3.07 (2H, t, *J* = 7.4 Hz, H3'), 2.67 (2H, t, *J* = 7.4 Hz, H2'); ^{13}C NMR (101 MHz, MeOD) δ 177.0 (C1'), 136.8 (ArC), 131.7 (ArC), 127.8 (C2), 122.3 (C5), 121.8 (ArC), 120.4 (C4), 117.2 (C7), 116.3 (C≡N), 104.3 (C6), 35.9 (C2'), 21.3 (C3'); *m/z* (ESI) 215.1 ([M+H]$^{+}$); HRMS *m/z* calcd for $C_{12}H_{11}N_2O_2$ 215.0815 found: 215.0819; λ_{max} (H_2O/CH_3CN) 244, 288, 323 nm.

7.4.2 Cyanation Screening Methods

7.4.2.1 General Method for Analytical Scale Cyanation using Water-Soluble Ligands

Arylbromide (3.0 mM), base (6.0 mM) and the appropriate additive at the concentration specified were dissolved in H_2O and the solution degassed under a stream of N_2. Glass vials with rubber septa were charged with $K_4[Fe(CN)_6]$ at the concentration specified and backfilled with N_2 prior to addition of a 1 mL aliquot of the degassed arylbromide solution. The appropriate combination of Pd source and ligand were then added at the specified concentration as deoxygenated stock solutions in water. The reactions were then heated at 80 ^{o}C under N_2 with stirring overnight prior to cooling and extraction into EtOAc (1 mL). After removal of organic solvent *in vacuo*, the residue was dissolved in MeOH (1 mL) and particulates removed by centrifugation (13000 rpm, 10 min) prior to analysis of the supernatant by analytical HPLC method 2 or 4 and LC-MS. The same method was used to screen conditions for the cyanation of CLEA biotransformations except the supernatant of the biotransformation after removal of the CLEA by centrifugation (10000 rpm, 30 min, 4 ^{o}C) was used in place of the aryl bromide solution.

7.4.2.2 General Method for Analytical Scale Cyanation using Palladacycles

Substrate (3.0 mM), base (6.0 mM) and the appropriate additive at the concentration specified were dissolved in H_2O and the solution degassed under a stream of N_2. Glass vials with rubber septa were charged with $K_4[Fe(CN)_6]$ at the concentration specified and backfilled with N_2 prior to addition of a 1 mL aliquot of the degassed substrate solution. tBu-XPhos-Pd-G3 (**157**) or CM-Phos-Pd (**156**) were then added as stock solutions in deoxygenated THF, ACN or H_2O at the concentration specified prior to heating at 80 ^{o}C under N_2 with vigorous stirring overnight prior to cooling and extraction into EtOAc (1 mL). After removal of organic solvent *in vacuo*, the residue was dissolved in MeOH (1 mL) and particulates removed by centrifugation (13000 rpm, 10 min) prior to analysis of the supernatant by analytical HPLC method 2 or 4 and LC-MS. The same method was used to screen conditions for the cyanation of CLEA biotransformations except the supernatant of the biotransformation after removal of the CLEA by centrifugation (10000 rpm, 30 min, 4 ^{o}C) was used in place of the aryl bromide solution.

CM-Phos-Pd (**156**) was prepared immediately prior to use by heating of $Pd(OAc)_2$ (2.7 mg, 12 µmol, 1 eq) and CM-Phos (9.7 mg, 24 µmol, 2 eq) in Et_3N (100 µL) and CH_2Cl_2 (1 mL) at 40 ^{o}C with stirring for 10 min. Solvent was then removed under a stream of N_2 and the resultant complex re-suspended in either ACN,THF or H_2O to a final concentration of 3 mM.

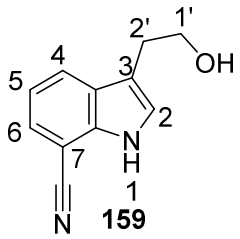
7.4.3 Integrated Cyanation Methods

7.4.3.1 General Method for Integrated Cyanation Using CLEAs

A solution of substrate (3 mM), NADH (100 µM), FAD (10 µM) and NaBr (30 mM) in 100 mM sodium phosphate buffer (pH 7.4) containing 5 v/v isopropanol to a total volume of 30 mL was used to re-suspend a halogenase CLEA from a total of 1.5 L culture volume. The resultant suspension was then incubated at room temperature with orbital shaking overnight prior to removal of the CLEA by centrifugation (10000 rpm, 30 min, 4 ^{o}C). $K_4[Fe(CN)_6].3H_2O$ (19 mg, 45 µmol, 50 mol %) was then added to the supernatant prior to freeze-thaw degassing and backfill with N_2. A

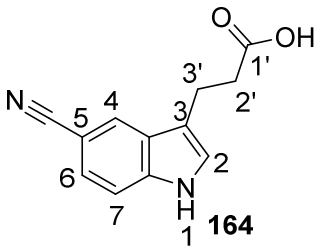
glass vial with a rubber septum was then backfilled with N_2 and charged with tBu-XPhos-Pd-G3 (**157**) (7.2 mg, 9 µmol, 10 mol %) followed by deoxygenated THF (2 mL). The tBu-XPhos-Pd-G3 (**157**) solution was then transferred to the degassed biotransformation supernatant followed by additional THF (4 mL) and heating of the resultant suspension at 80 ^{o}C with vigorous stirring overnight under N_2. Upon cooling, THF was removed under a stream of N_2 and the resulting suspension filtered with washing with 10 % MeOH in H_2O (10 mL). The filtrate was then applied to a C18 bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of methanol and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted using MeOH (3 CL) and solvent removed *in vacuo* prior to purification.

7.4.3.2 7-cyano trytophol (159)



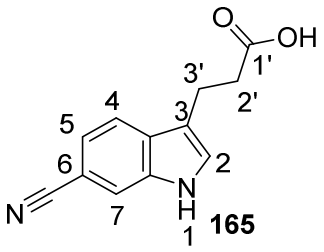
Prepared from tryptophol (**104**) using RebH CLEAs according to the general method above (7.4.3.1). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (10.9 mg, 65 %); ^{1}H NMR (400 MHz, MeOD) δ 7.88 (1H, dd, *J* = 8.0, 1.0 Hz, H4), 7.48 (1H, dd, *J* = 7.6, 1.0 Hz, H6), 7.25 (1H, app. s, H2), 7.14 (1H, dd, *J* =8.0, 7.6 Hz, H5), 3.81 (2H, t, *J* = 7.0 Hz, H1'), 2.98 (2H, td, *J* = 7.0, 1.0 Hz, H2'); *m/z* (ESI) 187.1 ([M+H]$^{+}$); HRMS *m/z* calcd for $C_{11}H_{11}N_2O$ 187.0866 found: 187.0864. Spectra were consistent with those previously reported herein.

7.4.3.3 5-cyano indole-3-propionic acid (164)



Prepared from indole-3-propionic acid (**106**) using PyrH CLEAs according to the general method above (7.4.3.1). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (13.7 mg, 71 %); ^{1}H NMR (400 MHz, MeOD) δ 10.94 (1H, br. s, NH), 8.01 (1H, d, *J* = 0.8 Hz, H4), 7.47 (1H, d, *J* = 8.8, H7), 7.37 (1H , dd, *J* = 8.8, 0.8 Hz, H6), 7.24 (1H, d, *J* = 0.8 Hz, H2), 3.08 (2H, td, *J* = 7.4, 0.8 Hz, H3'), 2.68 (t, *J* = 7.4 Hz, H2'); *m/z* (ESI) 215.1 ([M+H]$^{+}$); HRMS *m/z* calcd for $C_{12}H_{11}N_2O_2$ 215.0815 found: 215.0810. Spectra were consistent with those previously reported herein.

7.4.3.4 6-cyano indole-3-propionic acid (165)



Prepared from indole-3-propionic acid (**106**) using SttH CLEAs according to the general method above (7.4.3.1). Purification by semi-preparative HPLC method 6 afforded the title compound as a white solid (13.3 mg, 69 %); ^{1}H NMR (400 MHz, MeOD) δ 7.73 (1H, d, *J* = 1.2 Hz, H7), 7.70 (1H, d, *J* = 8.2 Hz, H4), 7.35 (1H, s, H2), 7.28 (1H, dd, *J* = 8.2, 1.2 Hz, H5), 3.07 (2H, t, *J* = 7.4 Hz, H3'), 2.67 (2H, t, *J* = 7.4 Hz, H2'); *m/z* (ESI) 215.1 ([M+H]$^{+}$); HRMS *m/z* calcd for $C_{12}H_{11}N_2O_2$ 215.0815 found: 215.0813. Spectra were consistent with those previously reported herein. 5-cyano indole-3-propionic acid (**164**) was also afforded as a minor by-product (1.0 mg, <5 %).

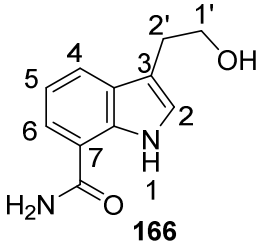
7.4.4 Integrated Amidation Methods

Expression and purification of the nitrile hydratase (NHase) used for these reactions is detailed in a later section (7.5.1). PDMS thimbles were prepared as described in section 7.3.9.

7.4.4.1 General Method for Analytical Scale NHase-Catalysed Hydration

To a solution of substrate (3.0 – 10 mM) in 100 mM potassium phosphate buffer (pH 8.0) was added NHase (8 – 65 μM) to a total volume of 200 μL. After incubation at room temperature with shaking at 800 rpm under laboratory lights for the specified time, reactions were quenched by boiling (95 °C, 10 min) and diluted 1:3 into MeOH to a total volume of 200 μL. After centrifugation (13000 rpm, 10 min), PhCN (1 mM) was added as internal standard and the supernatant analysed by analytical HPLC method 2 or 4. For analysis by HRMS, reactions were desalted using a ThermoFischer Pierce C18 Zip-Tip and eluted in MeOH (100 μL) according to the manufacturer's instructions.

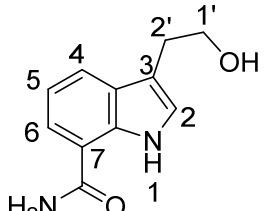
7.4.4.2 7-carboxamide tryptophol (166) – Standard Synthesis



To a solution of 7-cyano tryptophol (**159**) (3.0 mM) in 100 mM potassium phosphate buffer (pH 8.0) containing 5 % v/v MeOH was added Equi_NHase (65 μM) to a total volume of 25 mL. After incubation at room temperature with orbital shaking overnight, protein was precipitated by heating (95 °C, 10 min) and removed by centrifugation (10000 rpm, 30 min, 4 °C). The supernatant was then applied to a C18 bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of methanol and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted using

MeOH (3 CL) and solvent removed *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as an off-white solid (9.8 mg, 54 %); ν_{max}/cm^{-1} 3428 (m, N-H), 3310 (br., O-H), 2961 (w, C-H), 2920 (w, C-H), 2893 (w, C-H), 2497 (br., CON-H), 2326 (br., CON-H), 1622 (s, C=O), 1583 (m, C=C), 1554 (m, C=C); ^{1}H NMR (800 MHz, MeOD) δ 7.77 (1H, dd, *J* = 7.6, 1.0 Hz, H6), 7.62 (1H, dd, *J* = 7.5, 0.9 Hz, H4), 7.21 (1H, d, *J* = 1.0 Hz, H2), 7.08 (1H, t, *J* = 7.6 Hz, H5), 3.81 (2H, t, *J* = 7.2 Hz, H1'), 2.99 (2H, td, *J* = 7.2, 1.0 Hz, H2'); ^{13}C NMR (201 MHz, MeOD) δ 173.1 (C=O), 136.8 (ArC), 130.6 (ArC), 124.9 (C2), 123.9 (C6), 121.9 (C4), 119.0 (C5), 116.8 (ArC), 113.0 (ArC), 63.6 (C1'), 29.6 (C2'); *m/z* (ESI) 188.1 ($[M-NH_2]^+$, 100 %), 205.1 ($[M+H]^+$, 10); HRMS *m/z* calcd for $C_{11}H_{13}N_2O_2$ 205.0972 found: 205.0969; λ_{max} (H_2O/CH_3CN) 236, 318 nm.

7.4.4.3 7-carboxamide tryptophol (166) – Integrated Synthesis



166

An integrated cyanation of tryptophol (**104**) using RebH CLEAs was conducted as described above (7.4.3.1). After heating at 80 °C with vigorous stirring overnight, the reaction was cooled and THF removed under a stream of N_2. A PDMS thimble containing NHase (650 μM) in 100 mM potassium phosphate buffer (pH 8.0) to a total volume of 3 mL was then added to the crude cyanation reaction and reaction incubated at room temperature overnight. The reaction was then filtered and washed with 10 % MeOH in H_2O (10 mL). The filtrate was then applied to a C18 bond Elut® (Agilent, 20 g mesh) which had been activated with 3 CL of MeOH and equilibrated with 3 CL H_2O. After washing with H_2O (3 CL), product was eluted using MeOH (3 CL) and solvent removed *in vacuo*. Purification by semi-preparative HPLC method 3 afforded the title compound as an off-white solid (6.8 mg, 37 %); ν_{max}/cm^{-1} 3344 (br., O-H), 2961 (w, C-H), 2921 (w, C-H), 2893 (w, C-H), 2497 (br., CON-H), 2326 (br., CON-H), 1622 (s, C=O),1583 (m, C=C), 1555 (m, C=C); ^{1}H NMR (400 MHz, MeOD) δ 7.77 (1H, dd, *J* = 7.6, 1.0 Hz, H6), 7.63 (1H, dd, *J* = 7.6, 1.0 Hz, H4), 7.21 (1H, app. s, H2), 7.09 (1H, t, *J* = 7.6 Hz, H5), 3.81 (2H, t, *J* = 7.2 Hz, H1'), 2.99 (2H, td, *J* = 7.2, 0.8 Hz, H2'); *m/z* (ESI) 188.1 ($[M-NH_2]^+$, 100 %), 205.1 ($[M+H]^+$, 10); HRMS *m/z* calcd for $C_{11}H_{13}N_2O_2$ 205.0972 found: 205.0968; λ_{max} (H_2O/CH_3CN) 236, 318 nm.

7.4.5 Equi_NHase Docking Methods

Each substrate was MM2 minimised using Chem3D 16.0 and set with rotatable bonds around the ArC-CN bond and the ethyl chain (**159**, **164** and **165**) or ArC-Amide bond (**155** only) before addition of Gasteiger charges using AutoDock Tools to create the PDBQT file. Substrates were then docked into the Equi_NHase homology model using the parameters described in Chapter 7.5.

7.5 Chapter 5 Experimental

7.5.1 Nitrile Hydratase (NHase) Expression and Purification

Constructs containing the genes encoding the α- and β-subunits of the nitrile hydratase from *Rhodoccus equi* TG328-2 (Equi_NHase) in pET21a and the associated NHase activator in pET28a were kindly provided by the group of Uwe Bornscheuer (Greifswald University).[244] The composition of each protein purification buffer is detailed above (7.1.7).

7.5.1.1 Equi_NHase Expression

The pET28a construct containing the activator gene was transformed into *E. coli* BL21 (DE3) using kanamycin (50 μg/mL) for selection. Single colonies were used to inoculate 10 mL of LB containing kanamycin (50 μg/mL) and aliquots made chemically-competent using the protocol described above (7.1.4.3). The resulting cells were then transformed with the pET21a construct containing both the α -and β- NHase subunits using kanamycin (50 μg/mL) and ampicillin (100 μg/mL) for selection. Single colonies were then used to inoculate 10 mL of LB containing kanamycin (50 μg/mL) and ampicillin (100 μg/mL) and grown at 37 ^{o}C with shaking until OD_{600} = 0.6 at which point the culture was diluted 1:100 into 100 mL fresh LB medium containing kanamycin (50 μg/mL) and ampicillin (100 μg/mL) then incubated at 37 ^{o}C with shaking overnight. This culture was then diluted 1:100 into 10 x 400 mL 2YT medium containing kanamycin (50 μg/mL) and ampicillin (100 μg/mL) and incubated at 37 ^{o}C with shaking at 180 rpm until OD_{600} = 0.8. Incubation was then continued at 20 ^{o}C with shaking at 180 rpm until OD_{600} = 1.0 prior to induction with IPTG (0.1 mM) and incubation at 20 ^{o}C with shaking at 180 rpm for a further 22 hrs. Cells were then harvested by centrifugation (4000 rpm, 20 min, 4 ^{o}C) and either purified immediately or stored at -20 ^{o}C before use.

7.5.1.2 Equi_NHase Purification

Cell pellets were resuspended in 10 mM imidazole Tris buffer prior to lysis by sonication (15 min, 50 % pulse, 70 % power) and clarification by centrifugation (18000 rpm, 45 min, 4 ^{O}C). Lysate was then loaded onto Ni-NTA resin (Qiagen), pre-equilibrated with 10 mM imidazole Tris buffer, under gravity. The resin was then washed with 30 mM imidazole Tris buffer (1 CL) and protein eluted with 300 mM imidazole Tris buffer (5 CL). Eluent was then buffer exchanged into 50 mM Tris-HCl buffer (pH 7.4) containing butanoic acid (40 mM) using spin concentration and concentrated to ca 4 mL. Protein concentration was then determined using the A280 feature of a ThermoScientific 2000 Nanodrop and purity determined using SDS-PAGE.

7.5.2 NHase Biotransformations

All compounds characterised in this section have been previously characterised in the literature. Spectroscopic data is therefore compared to these publications, which are cited where appropriate.

7.5.2.1 General Method for Analytical Scale NHase Biotransformations

To a solution containing substrate (10 – 100 mM) in 100 mM potassium phosphate buffer (pH 8.0) containing 2 % v/v MeOH was added purified Equi_NHase (8 μM) to a total volume of 200 μL. After incubation at 20 ^{o}C with shaking at 800 rpm overnight, protein was precipitated by heating (95 ^{o}C,

10 min) and removed by centrifugation (13000 rpm, 15 min). Supernatant was then diluted to 5 mM final substrate concentration using MeOH and analysed by analytical HPLC method 2 or 5. For substrates **194** to **198**, reactions were extracted into EtOAc (200 µL) after removal of protein and butanoic acid (1 mM) added prior to analysis by GC-MS.

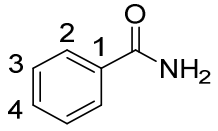
7.5.2.2 General Method 1 for Preparative Scale NHase Biotransformations

To a solution containing substrate (10 – 50 mM) in 100 mM potassium phosphate buffer (pH 8.0) containing 5 % v/v MeOH was added purified Equi_NHase (8 µM) to a total volume of 10 mL. After incubation at room temperature with shaking overnight, protein was precipitated by heating (70 ^{o}C, 15 min) and removed by centrifugation (10000 rpm, 15 min). Supernatant was then extracted with EtOAc (3 x 10 mL) and concentrated *in vacuo* prior to purification.

7.5.2.3 General Method 2 for Preparative Scale NHase Biotransformations

To a solution containing substrate (10 – 50 mM) in 100 mM potassium phosphate buffer (pH 8.0) containing 5 % v/v MeOH was added purified Equi_NHase (8 µM) to a total volume of 10 mL. After incubation at room temperature with shaking overnight, protein was precipitated by heating (70 ^{o}C, 15 min) and removed by centrifugation (10000 rpm, 15 min). Supernatant was then concentrated *in vacuo* and the resultant residue suspended in MeOH. After filtration, solvent was removed *in vacuo* and the residue dried under vacuum to afford pure amide.

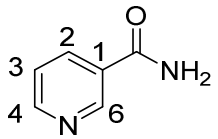
7.5.2.4 benzamide (168)[296]



168

Prepared from benzonitrile (**167**) (50 mM) according to general method 1 (7.5.2.2) reported above. Removal of organics *in vacuo* afforded the title compound as a white solid (39 mg, 64 %); ν_{max}/cm^{-1} 3362 (br., N-H), 3163 (br. N-H), 2849 (w, C-H), 2772 (w, C-H), 1622 (s, C=O), 1576 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (2H, app. d, *J* = 8.4 Hz, H2), 7.53 (1H, app. t, *J* = 7.6 Hz, H4), 7.45 (2H, app. t, *J* = 7.2 Hz, H3), 6.17 (2H, br. s, N<u>H</u>$_2$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.7 (<u>C</u>=O), 132.5, 129.0, 128.4, 127.8, 127.1; *m/z* (ESI) 105.0 ([M-NH_2]$^+$, 15 %), 122.1 ([M+H]$^+$, 100).

7.5.2.5 nicotinamide (181)[297]

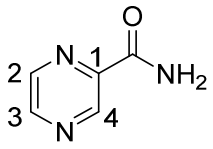


181

Prepared from nicotinonitrile (**185**) (50 mM) according to general method 2 (7.5.2.3) reported above to afford the title compound as a white solid (43 mg, 71 %); ν_{max}/cm^{-1} 3349 (br., N-H), 3147 (br., N-H), 2784 (w, C-H), 1672 (s, C=O); ^{1}H NMR (400 MHz, DMSO) δ 9.03 (1H, dd, *J* = 2.2, 1.2 Hz, H6), 8.70 (1H, dd, *J* = 3.6, 2.2 Hz, H4), 8.20 (1H, m, ArH), 8.18 (1H, br. s, N<u>H</u>$_2$), 7.61 (1H, br. s, N<u>H</u>$_2$)

7.49 (1H, m, ArH); ^{13}C NMR (101 MHz, DMSO) δ 166.4 (C=O), 151.9 (ArCH), 148.7 (C6), 135.2 (ArCH), 129.7 (C1), 123.4 (ArCH); *m/z* (ESI) 123.1 ([M+H]$^+$).

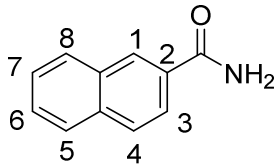
7.5.2.6 pyrazine-3-carboxamide (200) [297]



200

Prepared from pyrazine-2-carbonitrile (**186**) (50 mM) according to general method 2 (7.5.2.3) reported above. Purification by flash chromatography (SiO_2; 4:6 hexane:EtOAc) afforded the title compound as a white solid (38 mg, 62 %); v_{max}/cm^{-1} 3422 (br., N-H), 3136 (br. s, N-H), 2922 (w, C-H), 2751 (w, C-H), 1670 (s, C=O), 1582 (m, C=C); ^{1}H NMR (400 MHz, DMSO) δ 9.18 (1H, d, *J* = 1.2 Hz, H4), 8.86 (1H, d, *J* = 2.4 Hz, H2), 8.72 (1H, dd, *J* = 2.4, 1.2 Hz, H3), 8.28 (1H, br. s, NH$_2$), 7.88 (1H, br. s, NH$_2$);^{13}C NMR (101 MHz, DMSO) δ 165.1 (C=O), 147.4 (ArCH), 145.1 (ArC), 143.6 (ArCH), 143.4 (ArCH); *m/z* (ESI) 123.9 ([M+H]$^+$).

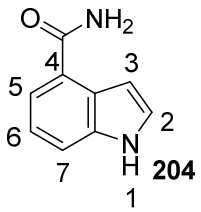
7.5.2.7 napthyl-2-carboxamide (205)[298]



205

Prepared from 2-napthonitrile (**191**) (25 mM) according to general method 1 (7.5.2.2) reported above. Purification by flash chromatography (SiO_2; 4:6 hexane:EtOAc) afforded the title compound as a white solid (10.3 mg, 24 %); ^{1}H NMR (400 MHz, DMSO) δ 8.48 (1H,s, ArH), 8.14 (1H, br. s, NH$_2$), 8.01 – 7.93 (4H, m, 4ArH), 7.63 – 7.56 (2H, m, 2ArH), 7.47 (1H, br. s, NH$_2$); ^{13}C NMR (101 MHz, DMSO) δ 168.0 (C=O), 134.2, 132.2, 131.6, 128.9, 127.8, 127.8, 127.6, 126.7, 124.4; *m/z* (ESI) 172.0 ([M+H]$^+$).

7.5.2.8 indole-4-carboxamide (204)[299]

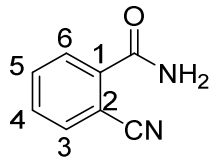


204

Prepared from indole-4-carbonitrile (**190**) (25 mM) according to general method 1 (7.5.2.2) reported above. Purification by flash chromatography (SiO_2; EtOAc) afforded the title compound as a white solid (40 mg, 38 %); ^{1}H NMR (400 MHz, MeOD) δ 7.57 (1H, d, *J* = 8.2 Hz, ArH), 7.48 (1H, dd, *J* = 7.4, 0.8 Hz, ArH), 7.37 (1H, d, *J* = 3.2 Hz, ArH), 7.17 (1H, app t., *J* = 7.6 Hz, H6), 6.89 (1H, dd, *J* = 3.2, 0.8 Hz, ArH); ^{13}C NMR (101 MHz, MeOD) δ 174.5 (C=O), 138.4, 127.5, 127.4, 126.7, 121.5,

120.5, 115.8, 102.6; *m/z* (ESI) 144.0 ([M-NH_2]$^+$, 20 %), 161.0 ([M+H]$^+$, 100), 183.0 ([M+Na]$^+$, 40); HRMS *m/z* calcd for $C_9H_8N_2NaO$ 183.0534 found: 183.0528.

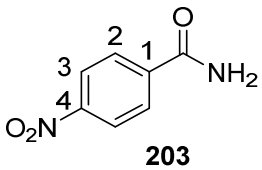
7.5.2.9 2-cyano benzamide (201)[300]



201

Prepared from phthalonitrile (**187**) (25 mM) according to general method 1 (7.5.2.2) reported above, except the biotransformation was quenched after 6 hrs. Purification by semi-preparative HPLC method 3 afforded the title compound as a white solid (8.4 mg, 23 %); ν_{max}/cm^{-1} 3353 (br., N-H), 3174 (br., N-H), 2960 (w, C-H), 2924 (w, C-H), 2228 (m, C≡N), 1652 (s, C=O), 1573 (m, C=C); ^{1}H NMR (400 MHz, DMSO) δ 8.18 (1H, br. s, N$\underline{H}_2$), 7.91 (1H, dd, *J* = 7.8, 1.4 Hz, ArH), 7.85 – 7.73 (3H, m, ArH and N$\underline{H}_2$), 7.67 (1H, td, *J* = 7.6, 1.4 Hz, ArH); ^{13}C NMR (101 MHz, DMSO-d_6) δ 169.1 ($\underline{C}$=O), 138.8 (ArC), 134.3 (ArCH), 132.8 (ArCH), 131.0 (ArCH), 128.2 (ArCH), 117.8 (C≡N), 110.3 (ArC); *m/z* (ESI) 147.9 ([M+H]$^+$).

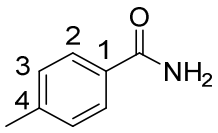
7.5.2.10 4-nitrobenzamide (203)[301]



203

Prepared from 4-nitrobenzonitrile (**189**) (25 mM) according to general method 1 (7.5.2.2) reported above. Purification by flash chromatography (SiO_2; 6:4 EtOAc:Hexane) afforded the title compound as a yellow solid (22.0 mg, 54 %); ν_{max}/cm^{-1} 3474 (br., N-H) 3171 (br. N-H), 1657 (s, C=O), 1598 (m, C=C), 1528 (m, C=C), 1316 (s, N=O), 1300 (s, N=O); ^{1}H NMR (400 MHz, DMSO) δ 8.35 – 8.23 (3H, m, 2ArH and N$\underline{H}_2$), 8.09 (2H, app. d, *J* = 8.8 Hz, 2ArH), 7.74 (2H, br. s, N$\underline{H}_2$); ^{13}C NMR (101 MHz, DMSO) δ 166.2 ($\underline{C}$=O), 149.0 (ArC), 139.9 (ArC), 128.9 (ArCH), 123.4 (ArCH); *m/z* (ESI) 165.0 ([M-H]$^+$); HRMS *m/z* calcd for $C_7H_5N_2O_3$ 165.0306 found: 165.0291.

7.5.2.11 4-methylbenzamide (202)[302]

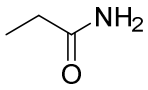


202

Prepared from 4-methylbenzonitrile (**188**) (25 mM) according to general method 1 (7.5.2.2) reported above. Purification by flash chromatography (SiO_2; 6:4 hexane:EtOAc) afforded the title compound as a white solid (20 mg, 59 %); ν_{max}/cm^{-1} 3341 (br., N-H), 3173 (br., N-H), 2917 (w, C-H), 2848 (w, C-H), 1660 (s, C=O), 1615 (m, C=C), 1566 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (2H, d, *J* = 8.2 Hz, 2ArH), 7.25 (2H, d, *J* = 8.2 Hz, 2ArH), 6.08 (1H, br. s, N$\underline{H}_2$), 5.83 (1H, br. s, N$\underline{H}_2$), 2.40 (3H, s, C$\underline{H}_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.5 ($\underline{C}$=O), 142.7 (ArC), 130.6 (ArC),

129.4 (ArCH), 127.5 (ArCH), 21.6 (<u>C</u>H_3); *m/z* (ESI) 135.9 ([M+H]$^+$, 10 %), 157.9 ([M+Na]$^+$, 100); HRMS *m/z* calcd for $C_8H_{10}NO$ 136.0757 found: 136.0753.

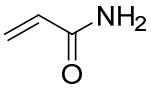
7.5.2.12 propionamide (206)[303]



206

Prepared from propiononitrile (**194**) (50 mM) according to general method 2 (7.5.2.3) reported above. Removal of solvent *in vacuo* afforded the title compound as a white solid (11.3 mg, 31 %); ν_{max}/cm^{-1} 3353 (br., N-H), 3382 (br., N-H), 2979 (w, C-H), 2941 (w, C-H), 1628 (s, C=O); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.34 (1H, br. s, N<u>H</u>$_2$), 5.88 (1H, br. s, N<u>H</u>$_2$), 2.20 (2H, q, *J* = 7.6 Hz, C<u>H</u>$_2$), 1.10 (3H, t, *J* = 7.6 Hz, C<u>H</u>$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.1 (<u>C</u>=O), 29.0 (<u>C</u>H_2), 9.7 (<u>C</u>H_3); *m/z* (ESI) 59.9 ([M-NH_2]$^+$, 50 %), 73.9 ([M+H]$^+$, 100).

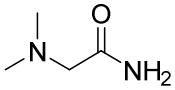
7.5.2.13 acrylamide (180)[304]



180

Prepared from acrylonitrile (**195**) (50 mM) according to general method 2 (7.5.2.3) reported above. Removal of solvent *in vacuo* afforded the title compound as a white solid (27.3 mg, 77 %); ν_{max}/cm^{-1} 3337 (br., N-H), 3103 (br., N-H), 2813 (w, C-H), 1667 (s, C=O), 1611 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.31 (1H, br. s, N<u>H</u>), 6.22 (1H, m, C<u>H</u>), 6.11 (1H, m, C<u>H</u>), 5.87 (1H, br. s, N<u>H</u>$_2$), 5.65 (1H, m, C<u>H</u>); ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.9 (<u>C</u>=O), 130.2, 127.6; *m/z* (EI) 55.0 ([M-NH_2]$^+$, 50 %), 71.0 ([M]$^+$, 100).

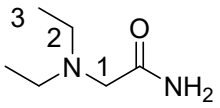
7.5.2.14 2-(dimethylamino)acetamide (207)[305]



207

Prepared from 2-(dimethylamino)acetonitrile (**197**) (50 mM) according to general method 1 (7.5.2.2) reported above. Removal of solvent *in vacuo* afforded the title compound as an off-white solid (34.2 mg, 67 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.01 (1H, br. s, N<u>H</u>), 5.71 (1H, br. s, N<u>H</u>), 2.94 (2H, s, C<u>H</u>$_2$), 2.30 (6H, s, 2C<u>H</u>$_3$); ^{13}C NMR (101 MHz, DMSO) δ 171.9 (<u>C</u>=O), 62.8 (<u>C</u>H_2), 45.5 (<u>C</u>H_3); *m/z* (ESI) 103.9 ([M+H]$^+$, 100 %), 124.9 ([M+Na]$^+$, 70).

7.5.2.15 2-(diethylamino)acetamide (208)[306]



208

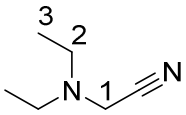
Prepared from 2-(diethylamino)acetonitrile (**198**) (50 mM) according to general method 1 (7.5.2.2) reported above. Removal of solvent *in vacuo* afforded the title compound as a yellow solid (46.8 mg, 71 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.00 (2H, s, H1), 2.55 (4H, q, *J* = 7.2 Hz, H2), 1.03 (6H, t, *J* = 7.2 Hz, H3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.7 (C=O), 57.5 (CH_2), 48.8 (CH_2), 12.5 (CH_3); *m/z* (ESI) 131.1 ($[M+H]^+$, 100 %), 153.1 ($[M+Na]^+$, 10).

7.5.3 Equi_NHase Docking Methods

The secondary structure of the Equi_NHase α- and β-subunit homology models (prepared using ExPASY) were aligned to the resolved crystal structure of the NHase from *Rhodococcus* erythropolis N-771 (PDB 3WVD) to create combined PDB file containing both subunits. Polar hydrogens and Gasteiger charges were then added using AutoDock Tools to create the PDBQT file. Charge on Fe was manually set to 3 as Autodock Tools does not calculate charges on metal ions. Substrates were set with rotatable bonds around the ArC-CN and ArC-$CONH_2$ bonds as appropriate and converted to PDBQT files with Gasteiger charges using AutoDock Tools. The grid box was centred around the substrate bound to PDB 3WVD (106.147, 23.095, 16.75, *xyz*) with 0.375 Å spacing and dimensions of 44 x 28 x 36 (*x* x *y* x *z*). Number of binding modes was set to 30 and mutation rate set to 0.05. All other parameters were left as default.

7.5.4 Synthesis of Amides for Cross-Coupling

7.5.4.1 2-(diethylamino)acetonitrile (198)[307]



198

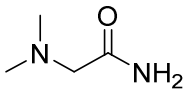
Diethylamine (4.0 mL, 38.7 mmol, 1.0 eq) was cooled on ice with stirring prior to slow addition of bromoacetonitrile (3.4 mL, 50.2 mmol, 1.3 eq) and stirring on ice overnight. The reaction was then poured into 2 M Na_2CO_3 (50 mL) and extracted into EtOAc (3 x 25 mL). The combined organics were then dried over $MgSO_4$ and concentrated *in vacuo*. After further drying *in vacuo*, the title compound was afforded as a yellow oil (4.17 g, 74 %); ν_{max}/cm^{-1} 3378 (br., N-H), 2230 (C≡N); ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.60 (2H, s, H1), 2.58 (H4, q, *J* = 7.2 Hz, H2), 1.09 (6H, t, *J* = 7.2 Hz, H3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 115.0 (C≡N) 48.1, 40.6, 12.8; *m/z* (ESI) 113.1 ($[M+H]^+$); HRMS *m/z* calcd for $C_6H_{13}N_2$ 113.1073 found: 113.1079.

7.5.4.2 General Method for the Hydrolysis of Nitriles

Crushed KOH (2.8 g, 50.0 mmol, 10 eq) was quickly transferred to an oven-dried flask and dissolved in anhydrous tertbutanol (20 mL) prior to addition of nitrile (5.0 mmol, 1.0 eq). After stirring at reflux for 1 hr, the reaction was cooled and diluted into brine (30 mL) prior to extraction

into CH_3Cl (3 x 30 mL). The combined organics were then dried over $MgSO_4$ and concentrated *in vacuo* to afford pure amide.

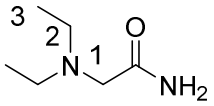
7.5.4.3 2-(dimethylamino)acetamide (207)[305]



207

Prepared from 2-(dimethylamino)acetonitrile (**197**) according to the general method above (7.5.4.2). Drying *in vacuo* afforded the title compound as an off-white solid (342 mg, 67 %); ν_{max}/cm^{-1} 3371 (br., N-H), 3195 (br., N-H), 2823 (w, C-H), 2777 (w, C-H), 1628 (s, C=O); ^{1}H NMR (400 MHz, DMSO) δ 7.15 (1H, br. s, N$\underline{H}_2$), 7.07 (1H, br. s, N$\underline{H}_2$), 2.78 (2H, s, C$\underline{H}_2$), 2.19 (6H, s, 2C$\underline{H}_3$); ^{13}C NMR (101 MHz, DMSO) δ 171.9 ($\underline{C}$=O), 62.8 ($\underline{C}H_2$), 45.5 ($\underline{C}H_3$); *m/z* (ESI) 103.5 ([M+H]$^+$).

7.5.4.4 2-(diethylamino)acetamide (208)[306]

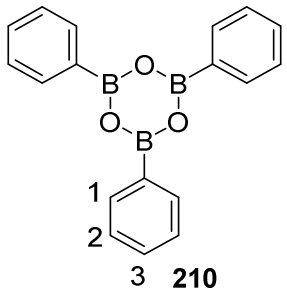


208

Prepared from 2-(diethylamino)acetonitrile (**198**) according to the general method above (7.5.4.2). Drying *in vacuo* afforded the title compound as an off-white solid (331 mg, 59 %); ν_{max}/cm^{-1} 3378 (br., N-H), 3174 (br., N-H), 2969 (m, C-H), 2931 (m, C-H), 1651 (s, C=O); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.26 (1H, br. s, N$\underline{H}_2$), 6.08 (1H, br. s, N$\underline{H}_2$), 2.99 (2H, s, C1), 2.54 (4H, q, *J* = 7.2 Hz, H2), 1.02 (6H, t, *J* = 7.2 Hz, H3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.7 ($\underline{C}$=O), 57.5 ($\underline{C}H_2$), 48.8 ($\underline{C}H_2$), 12.5 ($\underline{C}H_3$); *m/z* (ESI) 131.1 ([M+H]$^+$).

7.5.5 Copper-Catalysed C-N Bond Formation Methods

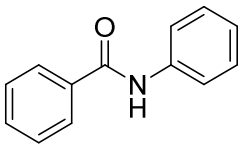
7.5.5.1 Triphenyl boroxine (210)[279]



210

Phenyl boronic acid (4.8 g, 39 mmol) was dissolved in toluene (100 mL) and the solution refluxed with azeotropic removal of water for 2 hrs. Upon cooling, the solution was concentrated *in vacuo* to ca 30 mL and the resultant precipitate collected by vacuum filtration. After washing with cold hexane (3 x 20 mL), the solid was dried *in vacuo* to afford the title compound as a white solid (3.36 g, 82 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (6H, app. d, H1), 7.61 (3H, app. t, H3), 7.53 (6H, app. t, H2); ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.8 (C1), 132.8 (C3), 128.1 (C2); *m/z* (APCI) 313.1 ([M+H]$^+$).

7.5.5.2 Synthesis of *N*-phenyl benzamide (209) standard using triphenyl boroxine (210)[279]



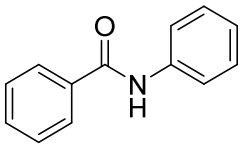
209

An oven-dried flask was charged with benzamide (**168**) (61 mg, 0.5 mmol, 1.0 eq) followed by $Cu(OTf)_2$ (18 mg, 50 µmol, 10 mol %), $Ph_3B_3O_3$ (**210**) (81.3 mg, 0.25 mmol, 0.5 eq) and anhydrous EtOH (5 mL). After stirring at 40 ^oC for 24 hrs, solvent was removed *in vacuo* and the residue dissolved in EtOAc (20 mL) prior to washing with saturated NH_4Cl (20 mL) and 4N NaOH (20 mL). The organic phase was then dried over $MgSO_4$ and concentrated *in vacuo* prior to purification by flash chromatography (SiO_2; 7:3, hexane:EtOAc) to afford the title compound as a white solid (35.5 mg, 36 %); ν_{max}/cm^{-1} 3360 (br., N-H) 2983 (w, C-H), 2967 (w, C-H), 1652 (s, C=O), 1623 (m, C=C); ^1H NMR (400 MHz, $CDCl_3$) δ 7.96 (1H, br. s, NH), 7.86 (2H, app. d, 2ArH), 7.64 (2H, app. d, 2ArH), 7.54 (1H, app t, ArH), 7.47 (2H, app. t, 2ArH), 7.36 (2H, app. t, 2ArH), 7.15 (1H, app. t, ArH); ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.9 (C=O), 138.1 (ArC), 135.1 (ArC), 132.0 (ArCH), 129.2 (ArCH), 128.9 (ArCH), 127.2 (ArCH), 124.7 (ArCH), 120.4 (ArCH); *m/z* (ESI) 198.0 ($[M+H]^+$, 100 %), 220.0 ($[M+Na]^+$, 90); HRMS *m/z* calcd for $C_{13}H_{10}NO$ 196.0768 found: 196.0758.

7.5.5.3 General Method for Preparation of CLE Product Standards

Benzamide (**168**) (61 mg, 0.5 mmol, 1.0 eq) was dissolved in CH_2Cl_2 (0.5 mL) and DMSO (0.5 mL) containing 4 Å molecular sieves (375 mg) and potassium aryltrifluoroborate salt (1.5 mmol, 1.5 eq). $Cu(OAc)_2$ (9 mg, 50 µmol, 10 mol %) was then added prior to heating at 40 ^oC with vigorous stirring overnight under a balloon of air. After cooling, the reaction was filtered through a celite plug with CH_2Cl_2 elution and the filtrate dried *in vacuo*. The residue was then dry-loaded onto SiO_2 and purified by flash chromatography.

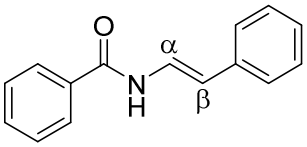
7.5.5.4 *N*-phenyl benzamide (209)[279]



209

Prepared according to the general method above (7.5.5.3) using potassium phenyltrifluoroborate. Purification by flash chromatography (SiO_2; 7:3, hexane:EtOAc) afforded the title compound as a white solid (77.8 mg, 79 %). Spectra were consistent with those reported herein.

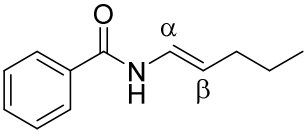
7.5.5.5 (E)-*N*-styryl benzamide (213)[308]



213

Prepared according to the general method above (7.5.5.3) using potassium (E)-styryltrifluoroborate. Purification by flash chromatography (SiO_2; 7:3 hexane:EtOAc) afforded the title compound as a yellow solid (74.7 mg, 70 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.50 (1H, br. d, *J* = 10.8, NH), 7.86 (2H, m, 2ArH), 7.72 (1H, dd, *J* = 14.6, 10.8 Hz, Hα), 7.50 (1H, app. t, ArH), 7.43 (2H, app. t, 2ArH), 7.32 (2H, m, 2ArH), 7.26 (2H, m, 2ArH), 7.26 (1H, app. t, ArH), 6.28 (1H, d, *J* = 14.6 Hz, Hβ); ^{13}C NMR ($CDCl_3$) δ, 164.7 (C=O), 136.2, 133.6, 132.1, 128.8, 127.3, 126.8, 125.7, 123.3, 113.7; *m/z* (ESI) 224.1 ($[M+H]^+$).

7.5.5.6 (E)-*N*-(pent-1-enyl) benzamide (214)[308]



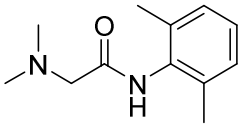
214

Prepared according to the general method above (7.5.4.3) using potassium (E)-pent-1-enyl trifluoroborate. Purification by flash chromatography (SiO_2; 7:3 hexane:EtOAc) afforded the title compound as a yellow solid (55.8 mg, 59 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (2H, app. d, 2ArH), 7.74 (1H, br. d, *J* = 10 Hz, NH), 7.50 (1H, m, ArH), 7.43 (2H, app. d, 2ArH), 6.95 (1H, m, Hα), 5.31 (1H, dt, *J* = 14.2, 7.2 Hz, Hβ), 2.05 (2H, qd, *J* = 7.2, 1.4 Hz, CH_2), 1.42 (2H, app. q, *J* = 7.4 Hz, CH_2), 0.92 (3H, t, *J* = 7.4 Hz, CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.4 (C=O), 134.0 (C), 131.9 (CH), 128.8 (CH), 127.1 (CH), 123.0 (CH), 114.3 (CH), 32.0 (CH_2), 23.2 (CH_2), 13.7 (CH_3); *m/z* (ESI) 190.1 ($[M+H]^+$, 60 %), 212.1 ($[M+Na]^+$, 100); HRMS *m/z* calcd for $C_{12}H_{15}NONa$ 212.1051 found: 212.1043.

7.5.5.7 General Method for the Preparation of *N*-2,6-dimethylphenyl Amides

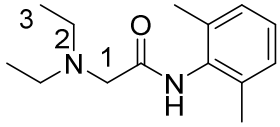
A flask was charged with amide (1.0 mmol, 1.0 eq), potassium 2,6-dimethylphenyltrifluoroborate (318 mg, 1.5 mmol, 1.5 eq), $Cu(OAc)_2$ (181 mg, 1.0 mmol, 1.0 eq) and CH_2Cl_2 (20 mL). The suspension was then stirred for 5 min prior to addition of Et_3N (700 µL, 5.0 mmol, 5.0 eq) and stirring at reflux overnight. Upon cooling, the reaction was dilued with CH_2Cl_2 (10 mL) and washed with saturated $NaHCO_3$ (20 mL). The aqueous phase was then extracted with further CH_2Cl_2 (2 x 20 mL) and the combined organics dried over $MgSO_4$ before concentration *in vacuo* and purification of the crude residue by flash chromatography.

7.5.5.8 2-(dimethylamino)-*N*-(2,6-dimethylphenyl) acetamide[309]



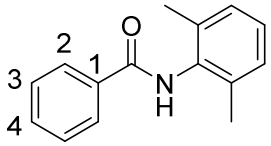
Prepared from 2-(dimethylamino) acetamide (**207**) according to the general method above (7.5.4.7). Purification by flash chromatography (SiO_2; 6:4 hexane:EtOAc + 1 % Et_3N) afforded the title compound as a yellow solid (14.9 mg, 11 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.75 (1H, br. s, NH), 7.15 – 7.02 (3H, m, 3ArH), 3.24 (2H, s, CH_2), 2.51 (6H, s, $2CH_3$), 2.25 (6H, s, $2CH_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.6 (C=O), 135.3 (ArC), 133.7 (ArC), 128.4 (ArCH), 127.3 (ArCH), 46.2 (CH_3), 29.9 (CH_2), 18.8 (CH_3); *m/z* (ESI) 207.1 ($[M+H]^+$); HRMS *m/z* calcd for $C_{12}H_{19}N_2O$ 207.1492 found: 207.1502.

7.5.5.9 2-(diethylamino)-*N*-(2,6-dimethylphenyl) acetamide[310]



Prepared from 2-(diethylamino) acetamide (**208**) according to the general method above (7.5.4.7). Purification by flash chromatography (SiO_2: 9:1 CH_2Cl_2:MeOH + 1 % Et_3N) afforded the title compound as a yellow solid (53.8 mg, 23 %); ν_{max}/cm^{-1} 3271 (br., N-H), 2971 (w, C-H), 2922 (w, C-H), 1567 (m, C=O), 1514 (m, C=C), 1473 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.97 (1H, br. s, NH), 7.11 – 7.06 (3H, m, 3 ArH), 3.23 (2H, s, H1), 2.70 (4H, q, *J* = 7.2 Hz, H2), 2.22 (6H, s, $2CH_3$), 1.21 (6H, s, $2CH_3$), 1.14 (6H, t, *J* = 7.2 Hz, CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.3 (C=O), 135.2 (ArC), 134.0 (ArC), 128.3 (ArCH), 127.2 (ArCH), 57.4 (C1), 49.0 (C2), 24.9 (CH_3), 18.7 (CH_3), 12.7 (C3); *m/z* (ESI) 235.2 ($[M+H]^+$); HRMS *m/z* calcd for $C_{14}H_{23}N_2O$ 235.1805 found: 235.1795

7.5.5.10 *N*-(2,6-dimethylphenyl) benzamide[311]



Prepared from benzamide (**168**) according to the general method above (7.5.4.7). Purification by flash chromatography (SiO_2: 7:3 hexane:EtOAc) afforded the title compound as a yellow solid (16.0 mg, 7.1 %); ν_{max}/cm^{-1} 3267 (br., N-H), 2954 (w, C-H), 2921 (w, C-H), 2853 (w, C-H), 1642 (m, C=O), 1596 (m, C=C), 1578 (m, C=C); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.93 (2H, app. d, H2), 7.57 (1H, app. t, H4), 7.51 (2H, app. t, H3), 7.37 (1H, br. s, NH), 7.17 – 7.11 (3H, m, 3ArH), 2.30 (6H, s, $2CH_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.7 (C=O), 135.7 (ArC), 134.7 (ArC), 133.9 (ArC), 132.0 (C4), 129.0 (C3), 128.5 (ArCH), 127.6 (ArCH), 127.3 (ArCH), 126.5 (ArCH), 18.7 ($2CH_3$); *m/z* (ESI) 226.1 ($[M+H]^+$, 100 %), 248.1 ($[M+Na]^+$, 70); HRMS *m/z* calcd for $C_{15}H_{16}NO$ 226.1226 found: 226.1238.

7.5.5.11 General Method for Screening of CLE Conditions

To oven-dried glass vials fitted with rubber septa was added base (50 – 500 mM) and aryl boron species (50 – 250 mM). A stock solution containing benzamide (50 mM) and copper source (0.5 – 50 mM) in the stated solvent was then added (1 mL) and the reactions heated at the specified temperature with stirring overnight. Upon cooling, reactions were washed with H_2O (1 mL) and organic phase concentrated *in vacuo*. The residue was then suspended in MeOH to a final substrate concentration of 5 mM prior to removal of particulates by centrifugation (13000 rpm, 10 min) and analysis of the supernatant by analytical HPLC method 2.

7.5.5.12 General Method A for Integrated NHase-CLE Reactions using PDMS Compartmentalisation

To a solution of benzamide (50 – 100 mM) in water (10 mL) was added a PDMS thimble containing phenyl boronic acid (2.0 – 5.0 eq), base (2.0 – 10 eq), copper source (0.1 – 1.0 eq), ligand (0.2 – 2.0 eq) and any additives suspended in the appropriate solvent (1 mL). After heating at the specified temperature overnight, the inner and outer compartments were combined and the thimble soaked in EtOAc (30 mL). The reaction was then extracted into further EtOAc (2 x 30 mL) and the combined organics dried over $MgSO_4$ prior to concentration *in vacuo*. The crude was analysed by ^{1}H and ^{13}C NMR and analytical HPLC method 2 prior to purification by flash chromatography. PDMS thimbles were prepared as described in section 7.3.10.

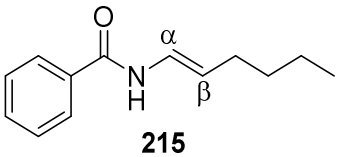
7.5.5.13 General Method B for Integrated NHase-CLE Reactions using PDMS Compartmentalisation

To a solution of benzamide (50 – 100 mM), aryl boron reagent (2.0 – 5.0 eq), base (2.0 – 5.0 eq), copper source (0.1 – 1.0 eq), ligand (0.2 – 2.0 eq) and any additives in the appropriate solvent (10 mL) was placed a PDMS thimble containing H_2O (1 mL). After heating at the specified temperature overnight, the inner and outer compartments were combined and the thimble soaked in EtOAc (30 mL). The reaction was then washed with saturated NH_4Cl (10 mL) and the aqueous phase extracted with further EtOAc (2 x 10 mL). The combined organics were then dried over $MgSO_4$ and concentrated *in vacuo*. The crude was analysed by ^{1}H and ^{13}C NMR and analytical HPLC method 2 prior to purification by flash chromatography. PDMS thimbles were prepared as described in section 7.3.10.

7.5.5.14 General Method for Goldberg Reactions

An oven-dried flask was charged with benzamide (121 mg, 1.0 mmol, 1.0 eq), base (1.0 – 5.0 eq) and aryl or vinyl bromide or iodide prior to backfill with N_2. Copper catalyst (10 – 20 mol %) was then added under a positive pressure of N_2 prior to addition of solvent (20 mL) and heating at the specified temperature overnight. Upon cooling, solvent was removed *in vacuo* and the residue dissolved in EtOAc (30 mL) prior to washing with saturated NH_4Cl (30 mL). The aqueous phase was then extracted into further EtOAc (2 x 30 mL) and the combined organics dried over $MgSO_4$ prior to concentration *in vacuo* and purification by flash chromatography.

7.5.5.15 (E)-*N*-(hex-1-enyl) benzamide (215)[277]

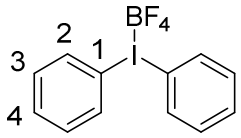


215

Prepared according to the general method above (7.5.5.14) using Cs_2CO_3 (1.0 eq), copper(I)-thiophene-2-caboxylate (**220**) (20 mol %), trans 1-iodohex-1-ene (1.0 eq) and N-methyl-2-pyrolidine with heating at 90 ^{O}C overnight. Purification by flash chromatography (SiO_2; 7:3 hexane:EtOAc) afforded the title compound as a yellow solid (30.5 mg, 15 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (2H, app. d, 2ArH), 7.73 (1H, br. d, *J* = 10.4 Hz, NH), 7.50 (1H, app. t, ArH), 7.43 (2H, app. d, 2ArH), 6.95 (1H, m, Hα), 5.31 (1H, dt, *J* = 14.2, 7.2 Hz, Hβ), 2.08 (2H, m, CH_2), 1.40 – 1.30 (4H, m, $2CH_2$), 0.90 (3H, t, *J* = 7.2 Hz, CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.4 (C=O), 134.0 (ArC), 131.9 (ArCH), 128.8 (ArCH), 127.1 (ArCH), 122.9 (Cα), 114.5 (Cβ), 32.2 (CH_2), 29.6 (CH_2), 22.3 (CH_2), 14.1 (CH_3); *m/z* (ESI) 204.1 ($[M+H]^+$).

7.5.6 Methods for the Coupling of Amides to Iodonium Salts

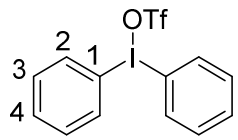
7.5.6.1 Diphenyliodonium tetrafluoroborate (222)[312]



222

To a solution of 3-chloroperbenzoic acid (640 mg, 3.0 mmol, 1.1 eq) in CH_2Cl_2 (10 mL) was added iodobenzene (310 µL, 2.7 mmol, 1.0 eq). BF_3OEt_2 (850 µL, 6.8 mmol, 2.5 eq) was then added dropwise and the resultant solution stirred for 1 hr. The reaction was then cooled on ice prior to addition of $PhB(OH)_2$ (370 mg, 3.0 mmol, 1.0 eq). After stirring at room temperature for 15 min, the reaction was applied to a silica plug and washed with CH_2Cl_2 (60 mL) prior to elution with 20:1 CH_2Cl_2:MeOH (120 mL). Solvent was then removed *in vacuo* and Et_2O (10 mL) added to the residue to induce precipitation. After removal of the Et_2O by decanting, the resultant white solid was stirred in further Et_2O (10 mL) for 10 min twice more with each Et_2O phase being discarded. After drying of the solid *in vacuo*, the title compound was afforded as a white solid (218 mg, 72 %); ^{1}H NMR (400 MHz, DMSO) δ 8.25 (2H, d, *J* = 8.0 Hz, H2), 7.67 (1H, t, *J* = 7.6, H4), 7.53 (2H, app. t, H3); ^{13}C NMR (101 MHz, DMSO) δ 135.2 (C2), 132.1 (ArCH), 131.8 (ArCH), 116.5 (C1); *m/z* (ESI) 280.9 ($[M-BF_4]^+$).

7.5.6.2 Diphenyliodonium triflate (223)[313]



223

To a solution of 3-chloroperbenzoic acid (640 mg, 3.0 mmol, 1.1 eq) in CH_2Cl_2 (10 mL) was added iodobenzene (310 µL, 2.7 mmol, 1.0 eq). BF_3OEt_2 (850 µL, 6.8 mmol, 2.5 eq) was then added dropwise and the resultant solution stirred for 1 hr. The reaction was then cooled on ice prior to

addition of $PhB(OH)_2$ (370 mg, 3.0 mmol, 1.0 eq). After stirring at room temperature for 15 min, TfOH (240 µL, 2.7 mmol, 1.1 eq) was added and stirring continued for 15 min. The reaction was then applied to a silica plug and washed with CH_2Cl_2 (60 mL) prior to elution with 20:1 CH_2Cl_2:MeOH (120 mL). Solvent was then removed *in vacuo* and Et_2O (10 mL) added to the residue to induce precipitation. After removal of the Et_2O by decanting, the resultant white solid was stirred in further Et_2O (10 mL) for 10 min twice more with each Et_2O phase being discarded. After drying of the solid *in vacuo*, the title compound was afforded as a white solid (840 mg, 72 %); ^{1}H NMR (400 MHz, DMSO) δ 8.25 (2H, app. d, H2), 7.66 (1H, app. t, H4), 7.53 (2H, app. t, H3); ^{13}C NMR (101 MHz, DMSO) δ 135.1 (C2), 132.1 (C4), 131.8 (C3), 122.7 (q, OTf), 116.5 (ArC); *m/z* (ESI) 280.9 ($[M\text{-}OTf]^+$).

7.5.6.3 General Method for Screening the Arylation of Amides using Iodonium Salts

Oven-dried glass vials fitted with rubber septa were charged with copper source (0.1 – 1.0 eq), diphenyliodonium tetrafluoroborate (**222**) or diphenyliodonium triflate (**223**) (1.1 – 5.0 eq) and base (2.0 – 5.0 eq) prior to backfill with N_2. A solution of benzamide (**168**) (50 mM) in the appropriate anhydrous solvent (1 mL) was then added to each vial and the reactions heated under N_2 with stirring at 40 ^oC overnight. Upon cooling, reactions were washed with H_2O (1 mL) and the organic phase concentrated *in vacuo*. The crude was then suspended in MeOH to a final substrate concentration of 5 mM and particulates removed by centrifugation (13000 rpm, 10 min) prior to analysis of the supernatant by analytical HPLC method 2.

7.5.6.4 General Method for Preparative Scale Arylation of Amides Using Iodonium Salts

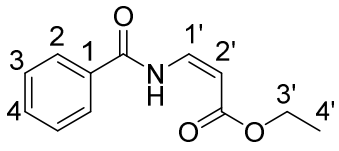
An oven-dried flask was charged with benzamide (**168**) (60 mg, 0.5 mmol, 1.0 eq), base (1.0 – 5.0 eq), copper source (0.1 to 1.0 eq) and either diphenyliodonium tetrafluoroborate (**222**) or diphenyliodonium triflate (**223**) (1.1 – 5.0 eq). The flask was then backfilled with N_2 prior to addition of solvent (10 mL) and stirring under N_2 at 40 ^oC overnight. Upon cooling, the reaction was washed with H_2O (10 mL) and the aqueous phase extracted with EtOAc (20 mL). The combined organics were then dried over $MgSO_4$ prior to concentration *in vacuo* and purification by flash chromatography.

7.5.6.5 General Method for Integrated NHase-Iodonium Reactions

Into a solution containing benzamide (60 mg, 0.5 mmol, 1.0 eq) in deoxygenated H_2O (10 mL) was palced a PDMS thimble containing copper source (0.1 – 1.0 eq), base (1.0 – 5.0 eq) and aryliodonium salt (1.1 – 5.0 eq) suspended in the appropriate solvent (1 mL). The reaction was then heated at the specified temperature overnight under N_2. After cooling, the inner and outer compartments were combined and the thimble soaked in EtOAc (10 mL). The aqueous phase was then extracted into further EtOAc (2 x 10 mL) and the combined organics dried over $MgSO_4$ prior to concentration *in vacuo* and purification by flash chromatography. PDMS thimbles were prepared as described in section 7.3.9.

7.5.7 Palladium-Catalysed Enamide Synthesis Methods

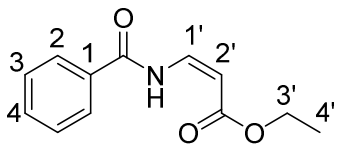
7.5.7.1 Ethyl (Z)-3-benzamidoacrylate (225) using ethyl acrylate (224)[289]



225

A flask was charged with 4 Å molecular sieves (100 mg), benzamide (**168**) (121 mg, 1.0 mmol, 1.0 eq), para-toluenesulfonic acid (95 mg, 0.5 mmol, 0.5 eq), $Pd(OAc)_2$ (22 mg, 100 µmol, 10 mol %) and 1,4-benzoquinone (118 mg, 1.1 mmol, 1.0 eq). Toluene (20 mL) was then added and the suspension stirred at room temperature for 15 min. Ethyl acrylate (160 µL, 1.5 mmol, 1.5 eq) was then added dropwise before stirring at room temperature overnight. The reaction was then diluted with CH_2Cl_2 (20 mL) and washed with H_2O (30 mL). The aqueous phase was then extracted into further CH_2Cl_2 (2 x 30 mL) and the combined organics dried over $MgSO_4$ before concentration *in vacuo*. Purification by flash chromatography (SiO_2; 8:2, hexane:EtOAc) afforded the title compound as an off-white solid (33 mg, 15 %). Spectra were consistent with those reported herein.

7.5.7.2 Ethyl (Z)-3-benzamidoacrylate (225) using ethyl propioloate (226)[289]



225

A flask was charged with benzamide (**168**) (121 mg, 1.0 mmol, 1.0 eq), NaOAc (164 mg, 2.0 mmol, 2.0 eq) and $Pd(OAc)_2$ (22 mg, 100 µmol, 10 mol %) prior to backfill with N_2. TFA (382 µL, 5.0 mmol, 5.0 eq) was then added under a positive pressure of N_2 prior to addition of deoxygenated toluene (10 mL) and the resultant mixture stirred for 15 min. Ethyl propiloate (152 µL, 1.5 mmol, 1.5 eq) was then added prior to heating at 70 ^{o}C with stirring overnight under N_2. Upon cooling, the reaction was diluted with CH_2Cl_2 (20 mL) and washed with saturated Na_2CO_3 (20 mL). The aqueous phase was then extracted with further CH_2Cl_2 and the combined organics dried over $MgSO_4$ prior to concentration *in vacuo*. Purification by flash chromatography (SiO_2; 8:2 hexane:EtOAc) afforded the title compound as an off-white solid (190 mg, 87 %); ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.51 (1H, br. d, *J* = 8.8 Hz, N<u>H</u>), 7.95 (2H, app. d, H2), 7.74 (1H, dd, *J* = 11.2, 8.8 Hz, H1'), 7.58 (1H, app. t, H4), 7.50 (2H, app. t, H3), 5.26 (1H, d, *J* = 8.8 Hz, H2'), 4.23 (2H, q, *J* = 7.2 Hz, H3'), 1.32 (3H, t, *J* = 7.1 Hz, H4'); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.8 (<u>C</u>=O), 164.64 (<u>C</u>=O), 138.9 (C1'), 133.0 (C4), 132.3 (C1), 129.0 (C3), 127.8 (C2), 97.3 (C2'), 60.4 (C3'), 14.4 (C4'); *m/z* (ESI) 220.0 ($[M+H]^+$).

7.5.7.3 General Method for the Integration of Pd-Catalysed Coupling of Alkynes to Amides with NHase

A PDMS thimble was charged with NaOAc (164 mg, 2.0 mmol, 2.0 eq), $Pd(OAc)_2$ (22 mg, 100 µmol, 10 mol %) and a stirrer bar before being placed inside a flask containing benzamide (**168**) (121 mg, 1.0 mmol, 1.0 eq). The flask was then backfilled with N_2 prior to addition of

deoxygenated water (20 mL) to the outside of the thimble and toluene (2 mL) to the inside of the thimble using long needles. TFA (382 µL, 5.0 mmol, 5.0 eq) was then added to the thimble under a positive pressure of N_2 and the contents of the thimble stirred for 15 min. Ethyl propiloate (152 µL, 1.5 mmol, 1.5 eq) was then added to the inside of the thimble using long needles and the reaction heated at 70 ^{o}C overnight under N_2. Upon cooling, the inner and outer compartments were combined and fragments of PDMS removed before being soaked in CH_2Cl_2. The aqueous portion was then extracted into further CH_2Cl_2 (2 x 30 mL) and the combined organics dried over $MgSO_4$ prior to concentration *in vacuo*. The crude resiude was then analysed by ^{1}H and ^{13}C NMR and analytical HPLC method 2 prior to purification by flash chromatography (SiO_2; 8:2 hexane:EtOAc) to afford the title compound as an off-white solid (57.2 mg, 26 %). Spectra were consistent with those reported herein. PDMS thimbles were prepared as described in section 7.3.9.

8. Appendix One

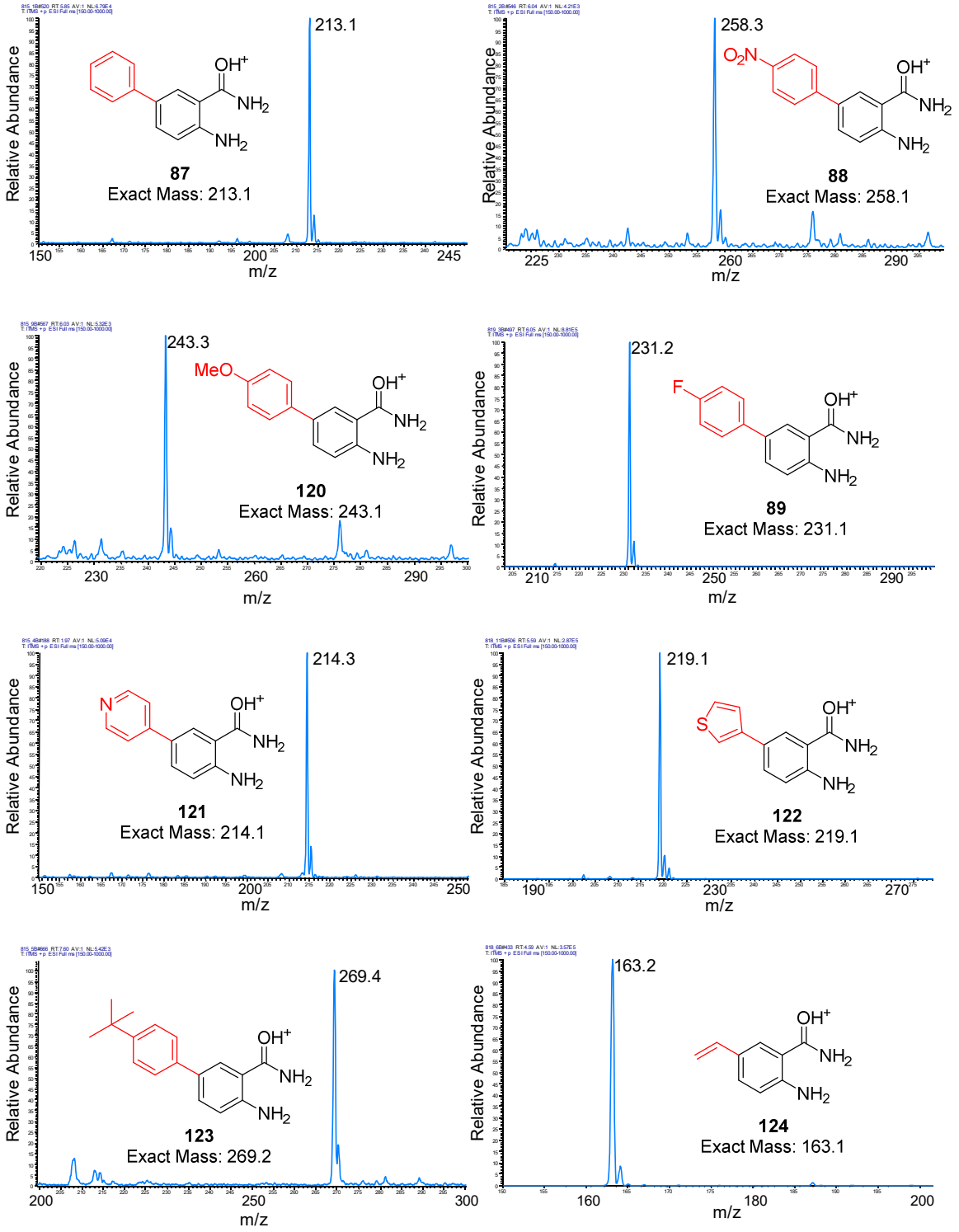


Figure S 3: Mass spectra of proposed arylated adducts from coupling partner screening of the integrated FIHal-SMC reaction using CLEA biotransformation supernatants (part 1).

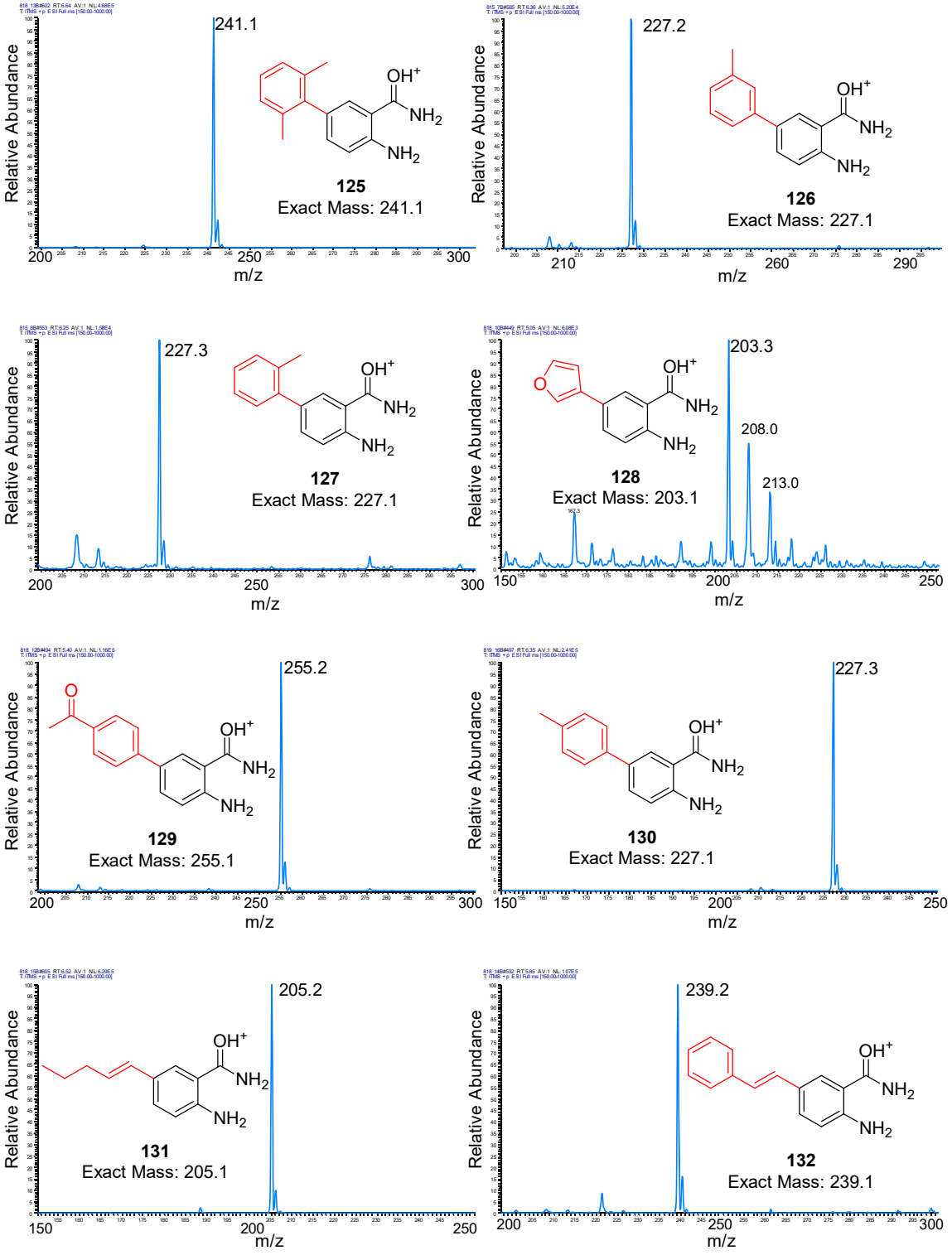


Figure S 4: Mass spectra of proposed arylated adducts from coupling partner screening of the integrated FIHal-SMC reaction using CLEA biotransformation supernatants (part 2).

9. Appendix Two

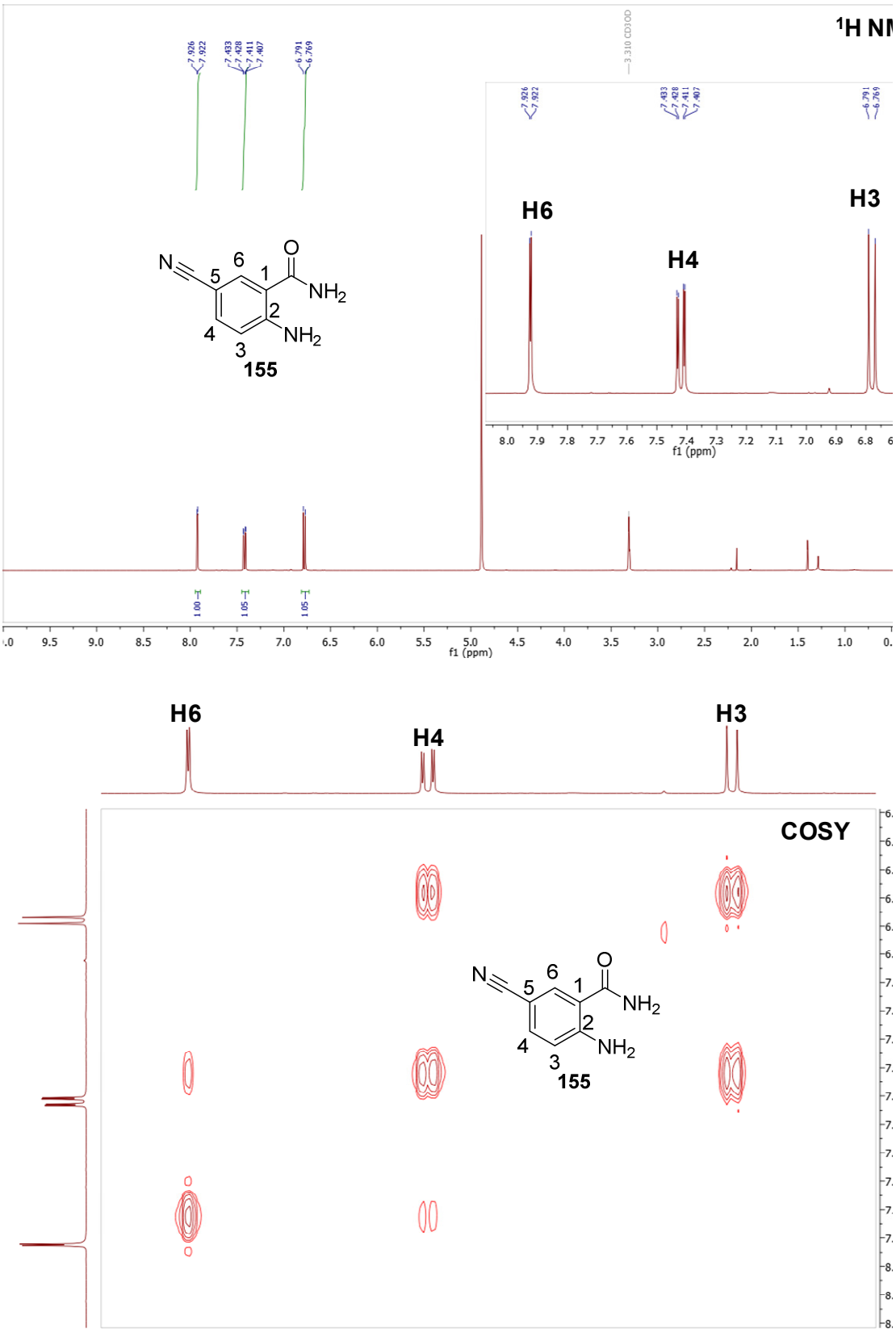


Figure S 5: ^{1}H and COSY NMR spectra (MeOD, 400 MHz) of 2-amino-5-cyano benzamide (**155**).

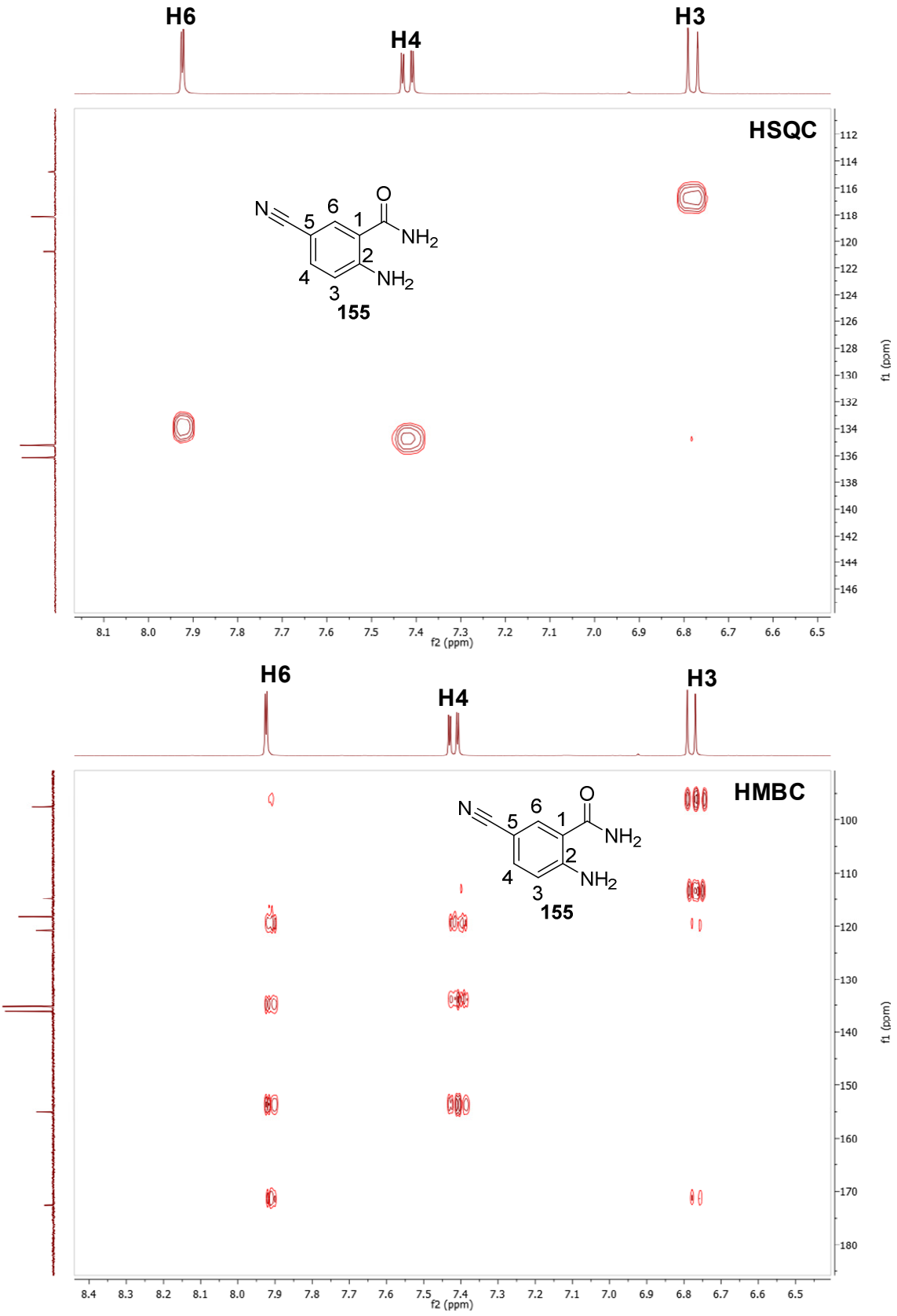


Figure S 6: HSQC and HMBC NMR spectra (MeOD, 101 MHz) of 2-amino-5-cyano benzamide (**155**).

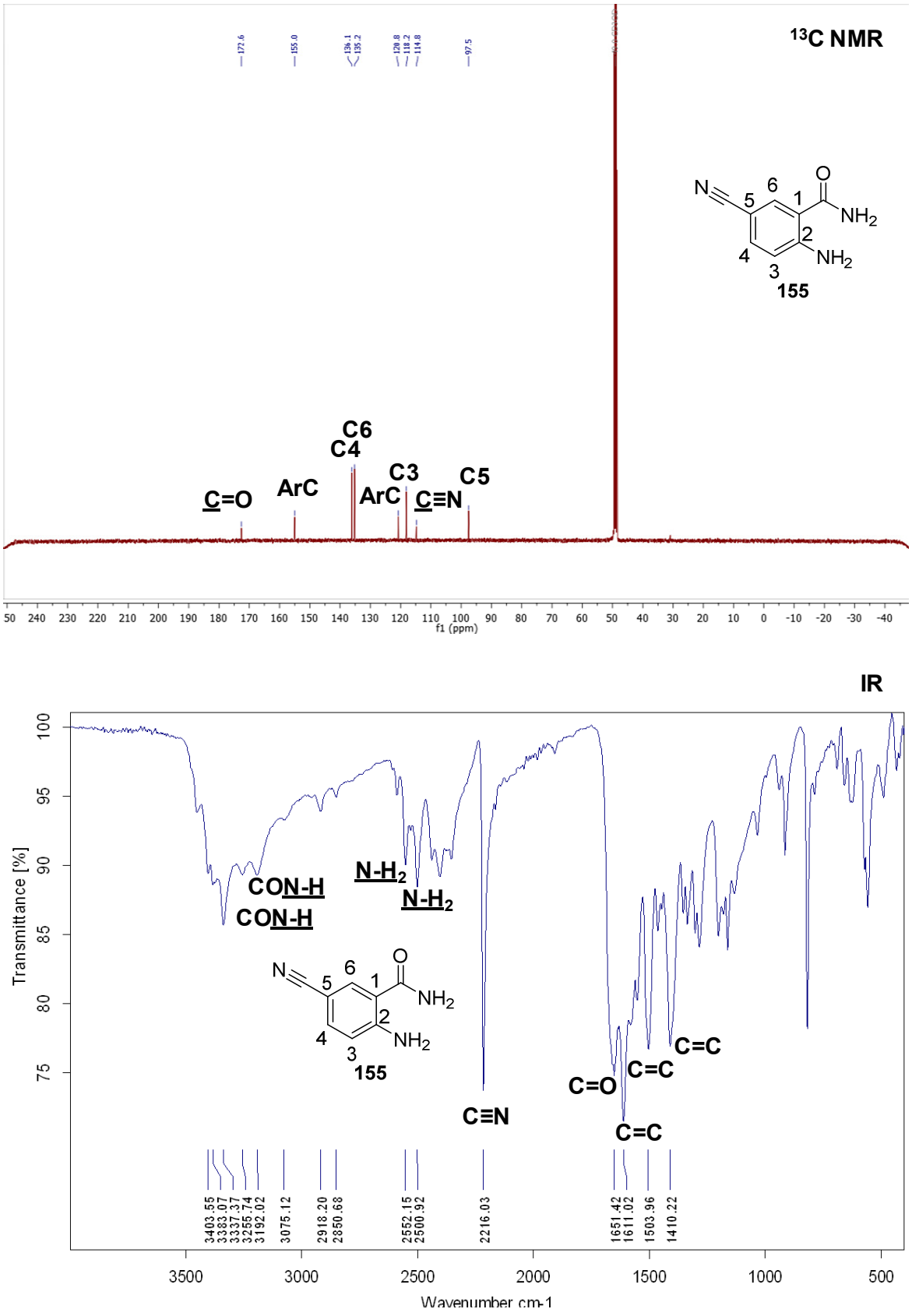


Figure S 7: ^{13}C NMR (MeOD, 101 MHz) and IR spectra of 2-amino-5-cyano benzamide (**155**).

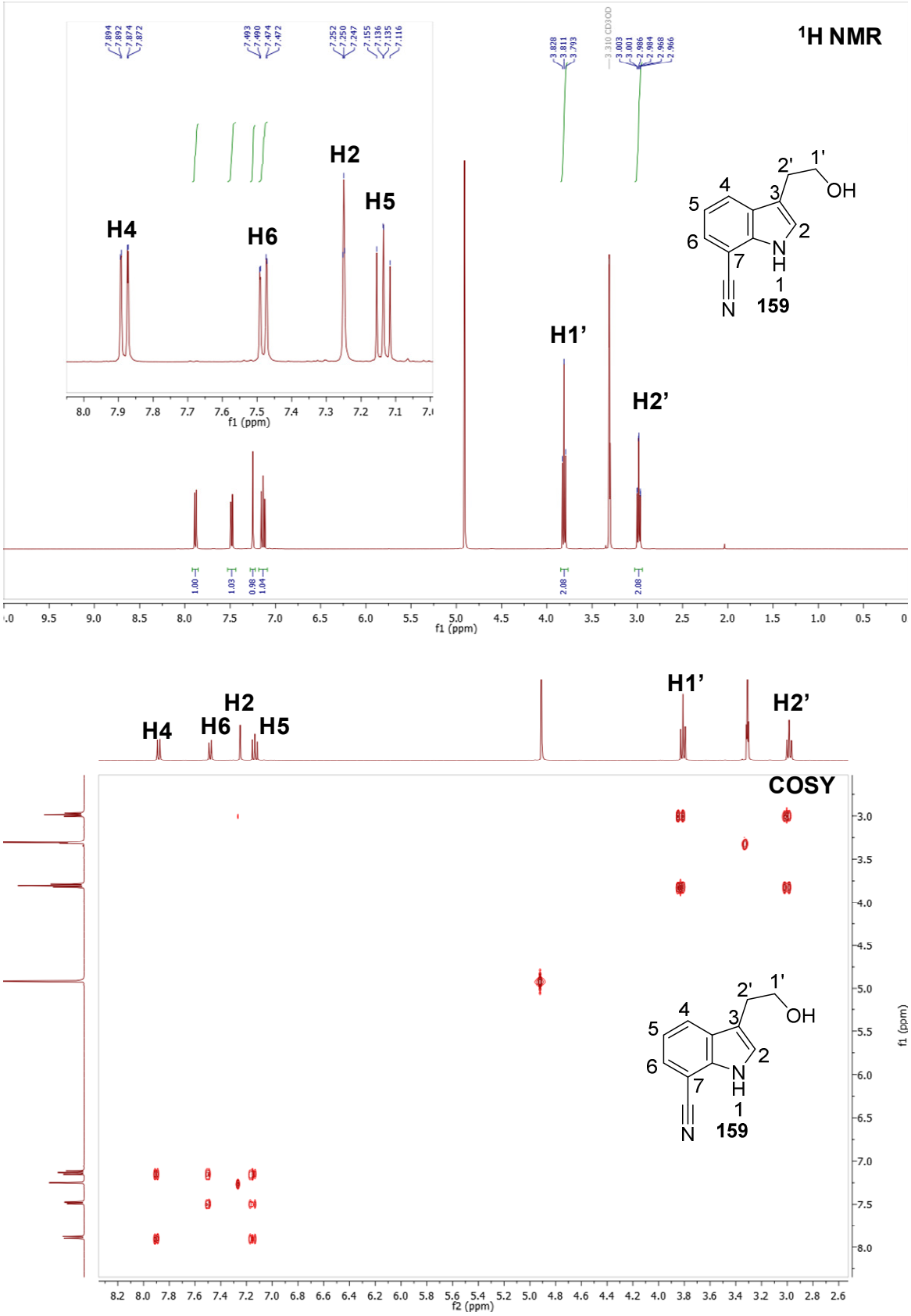


Figure S 8: [1]H NMR and COSY NMR spectra (MeOD, 400 MHz) of 7-cyano tryptophol (**159**).

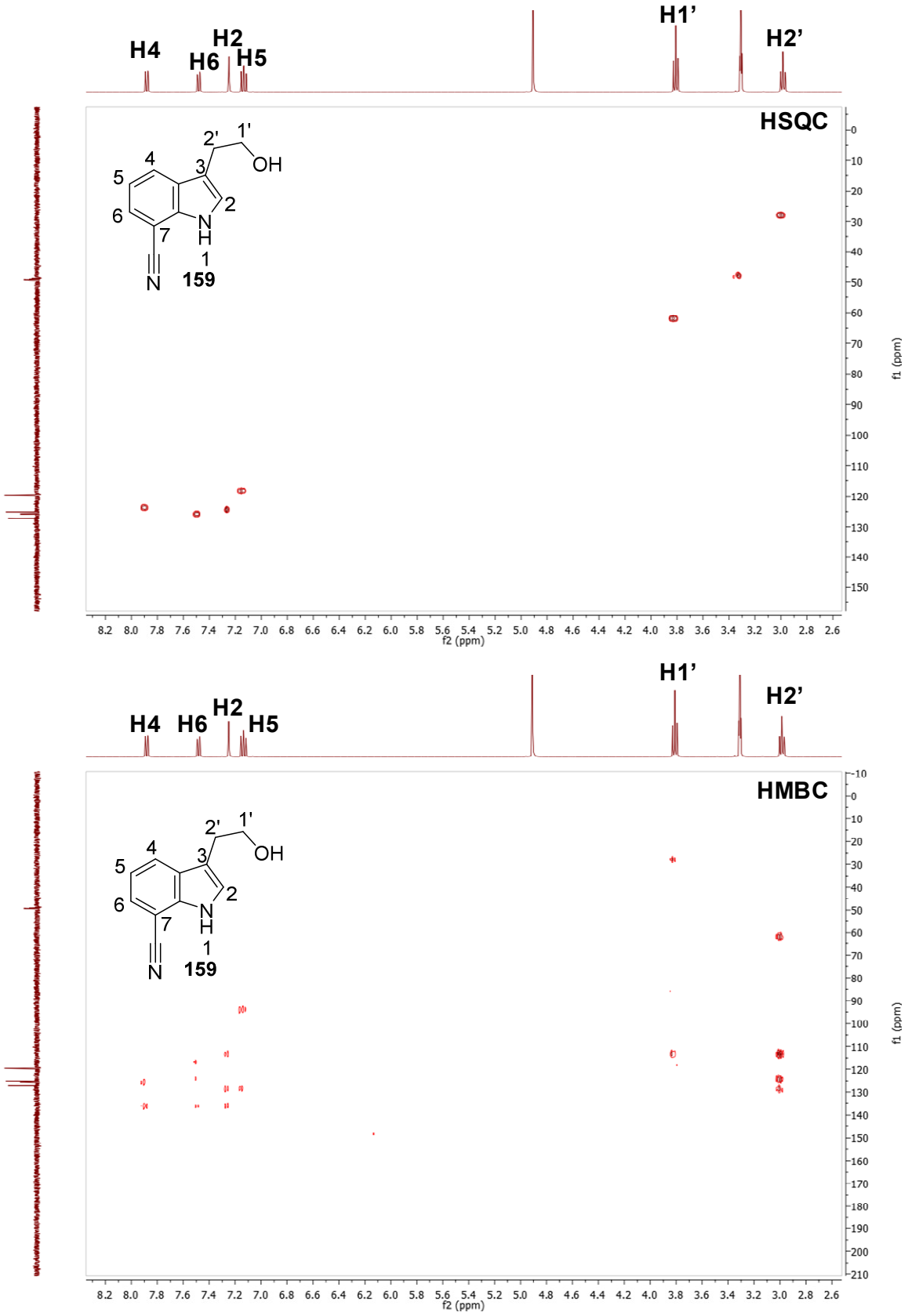


Figure S 9: HSQC and HMBC NMR spectra (MeOD, 101 MHz) of 7-cyano tryptophol (**159**).

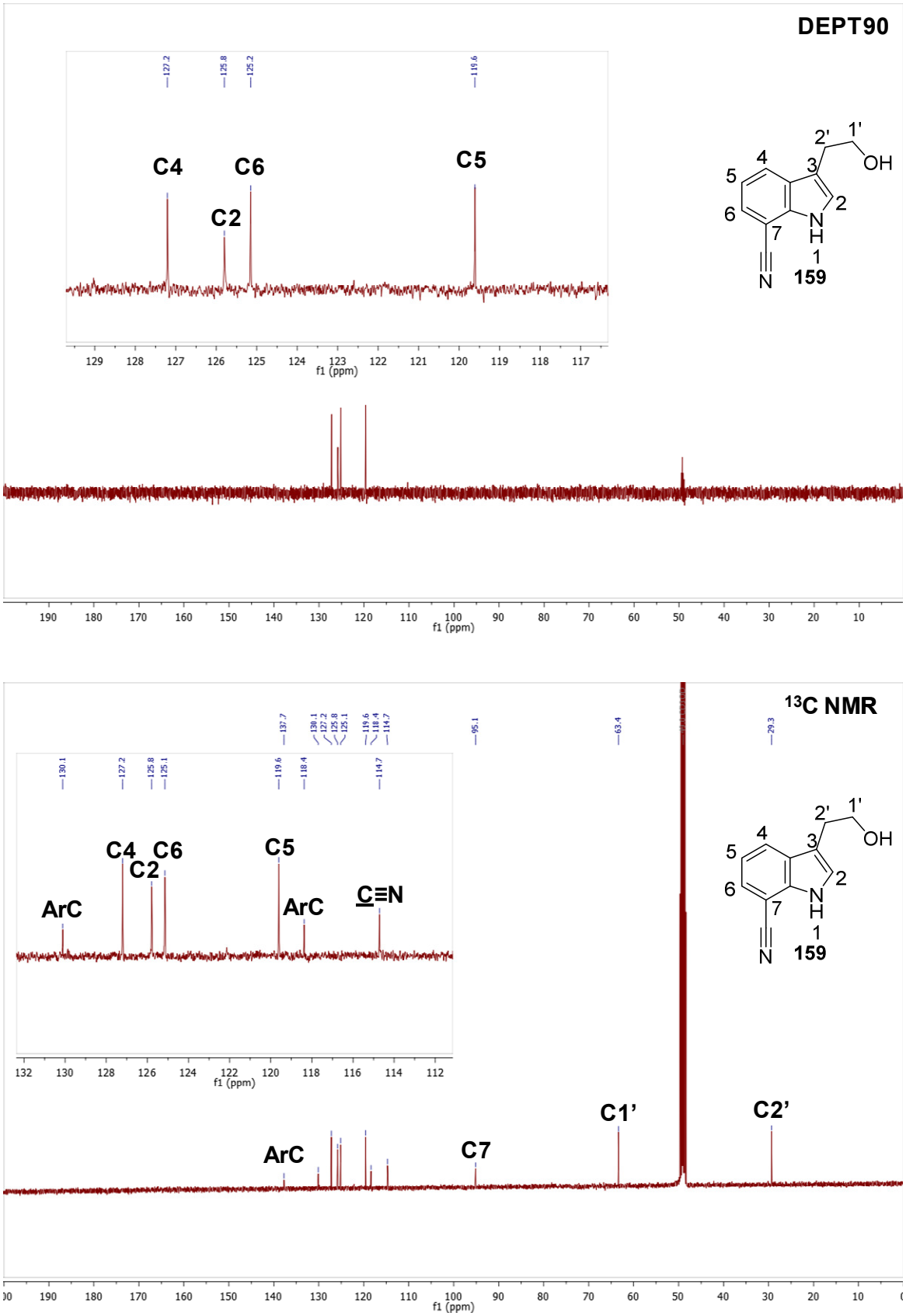


Figure S 10: DEPT90 and ^{13}C NMR spectra (MeOD, 101 MHz) of 7-cyano tryptophol (**159**).

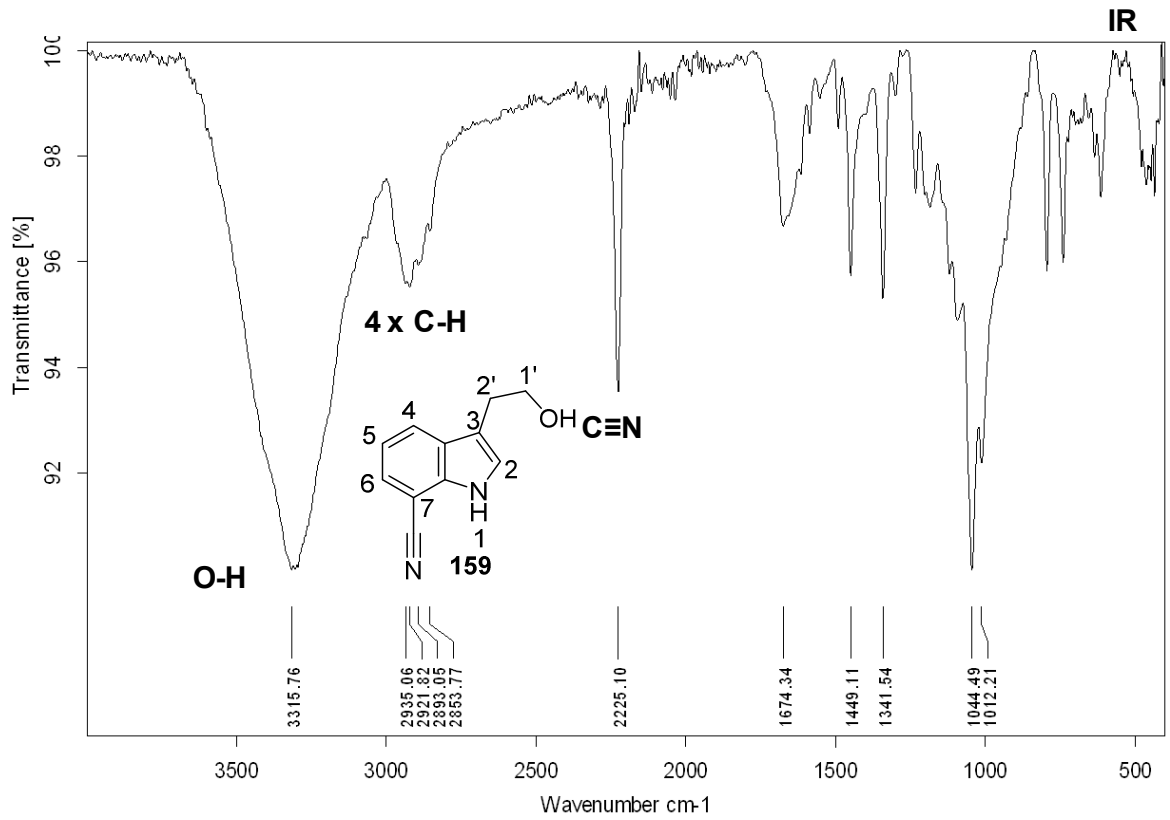


Figure S 11: IR spectrum of 7-cyano tryptophol (**159**).

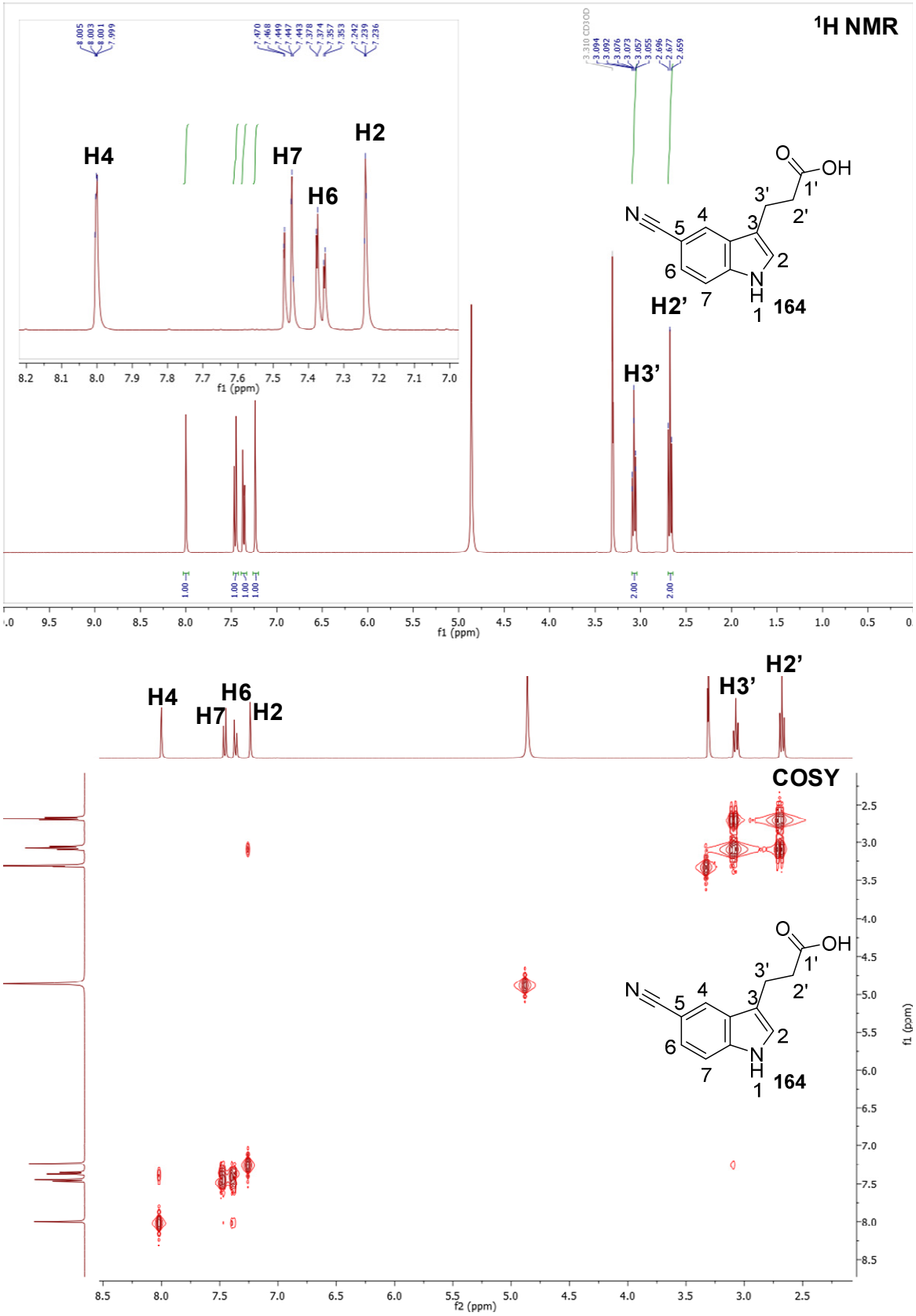


Figure S 12: [1]H and COSY NMR spectra (MeOD, 400 MHz) of 5-cyano indole-3-propionic acid (**164**).

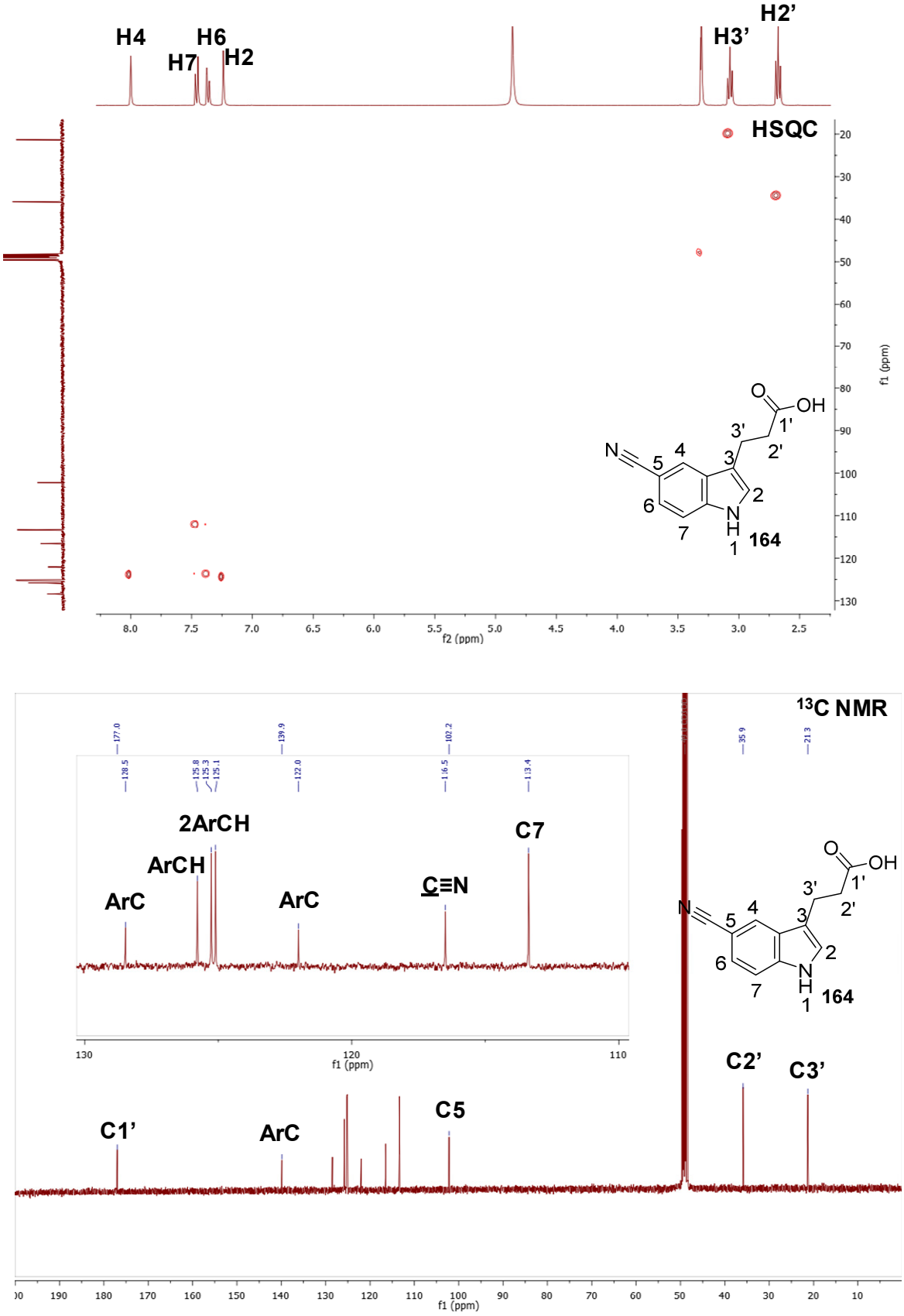


Figure S 13: HSQC and ^{13}C NMR spectra (MeOD, 101 MHz) of 5-cyano indole-3-propionic acid (**164**).

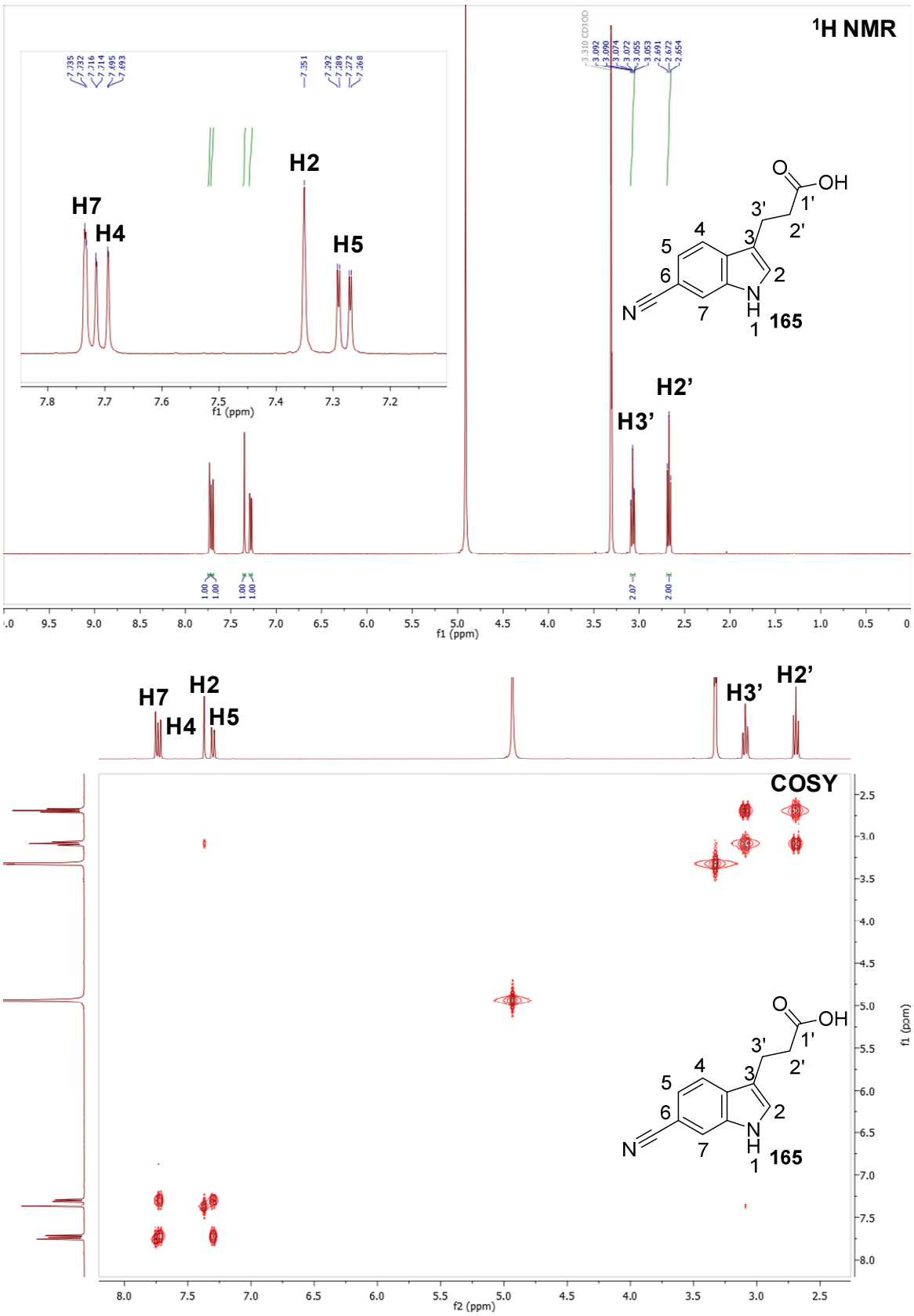


Figure S 14: ^{1}H and COSY NMR spectra (MeOD, 400 MHz) of 6-cyano indole-3-propionic acid (**165**).

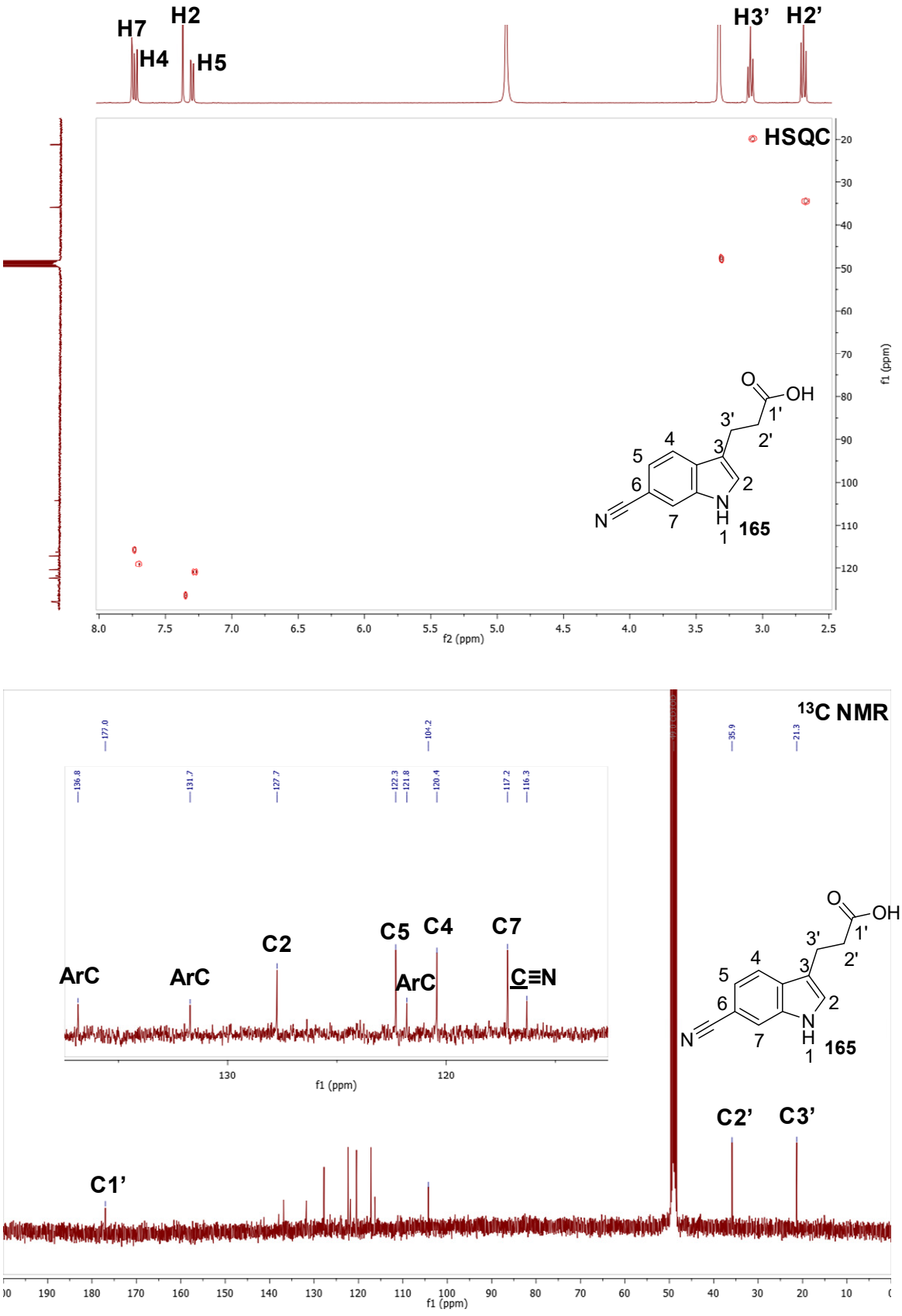


Figure S 15: HSQC and ^{13}C NMR spectra (MeOD, 101 MHz) of 6-cyano indole-3-propionic acid (**165**).

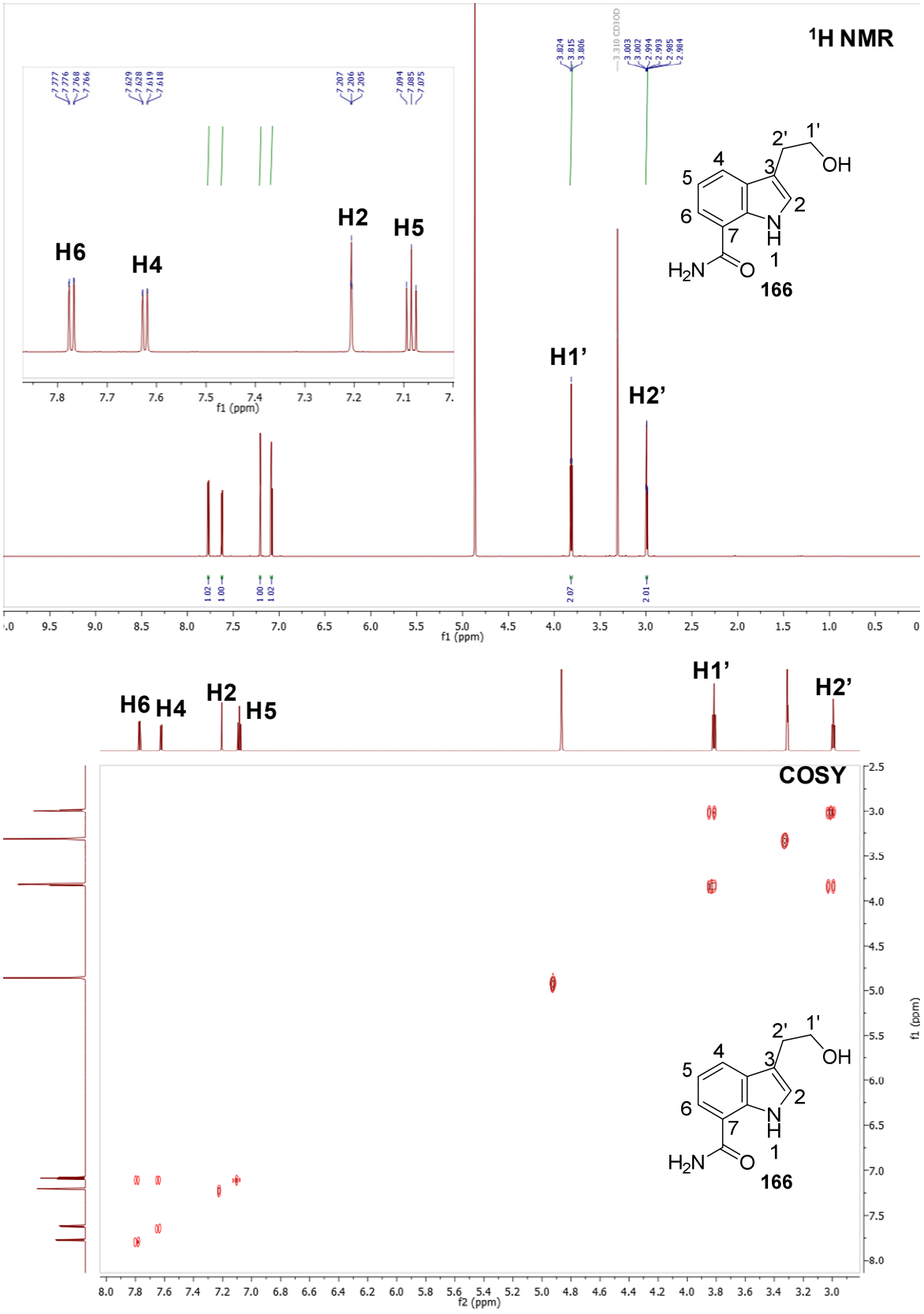


Figure S 16: [1]H NMR (MeOD, 800 MHz) and COSY (MeOD, 400 MHz) NMR spectra of 7-carboxamide tryptophol (**166**).

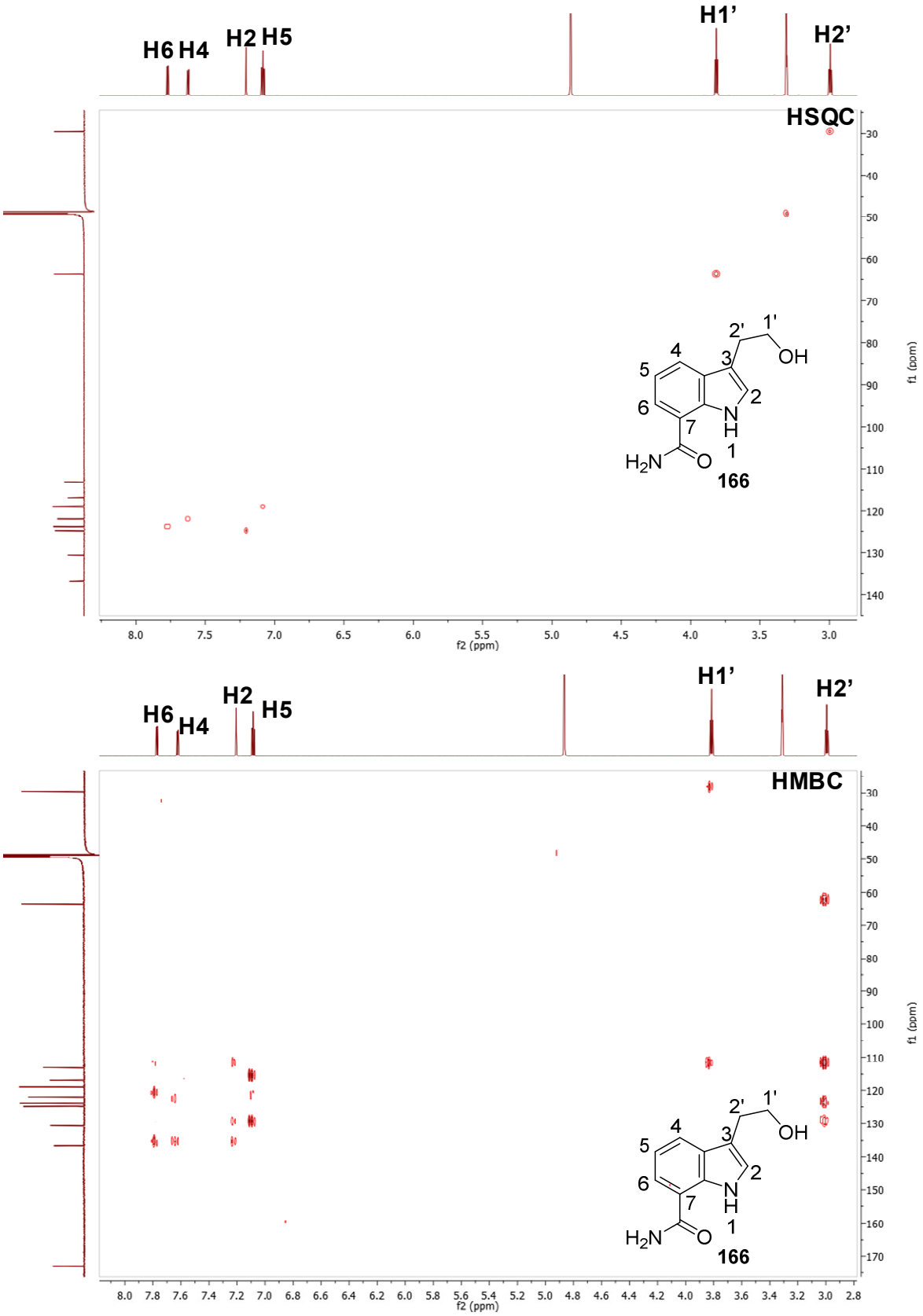


Figure S 17: HSQC and HMBC NMR spectra (MeOD, 101 MHz) of 7-carboxamide tryptophol (**166**).

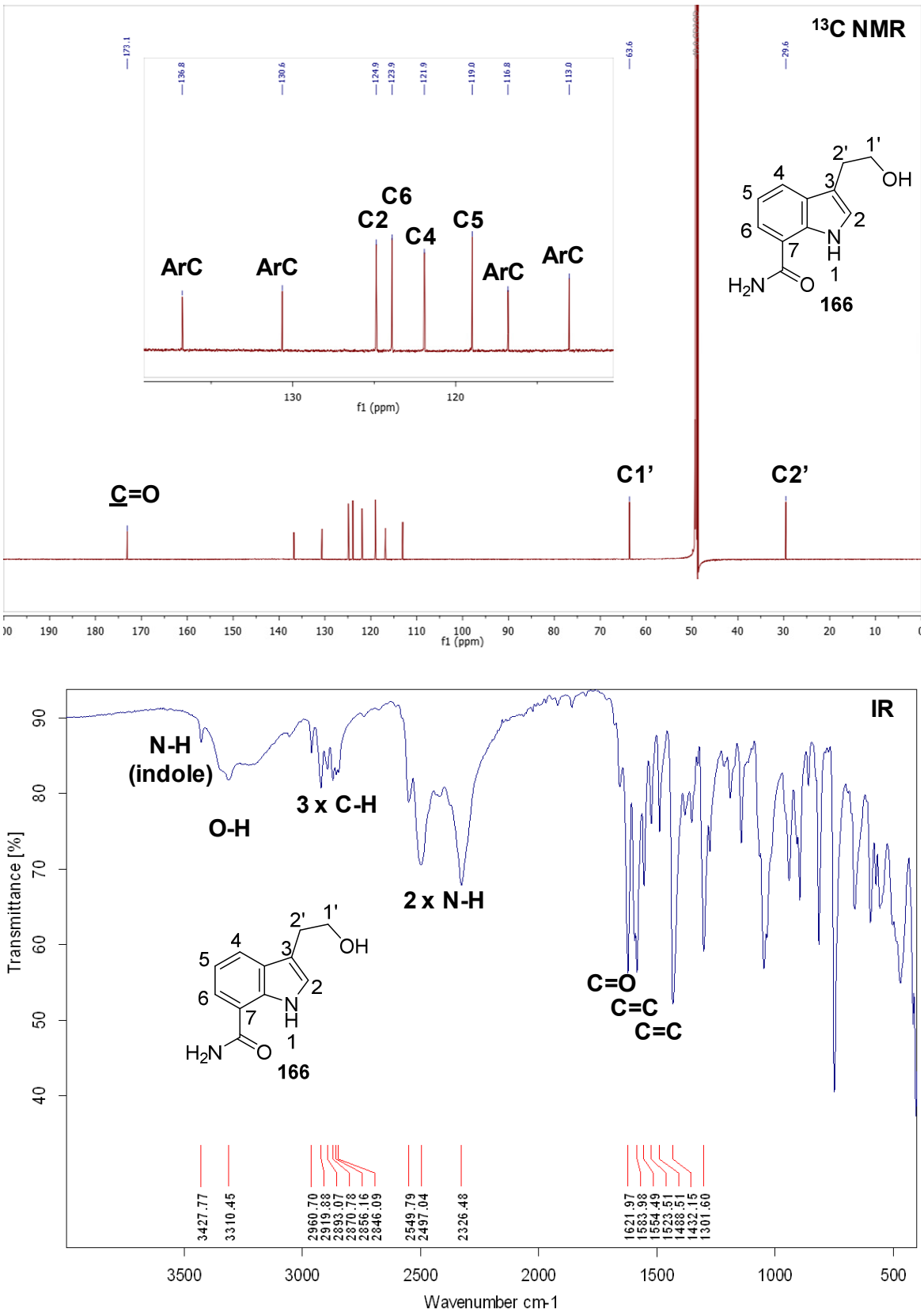


Figure S 18: ^{13}C NMR (MeOD, 202 MHz) and IR spectra of 7-carboxamide tryptophol (**166**).

10. Appendix Three

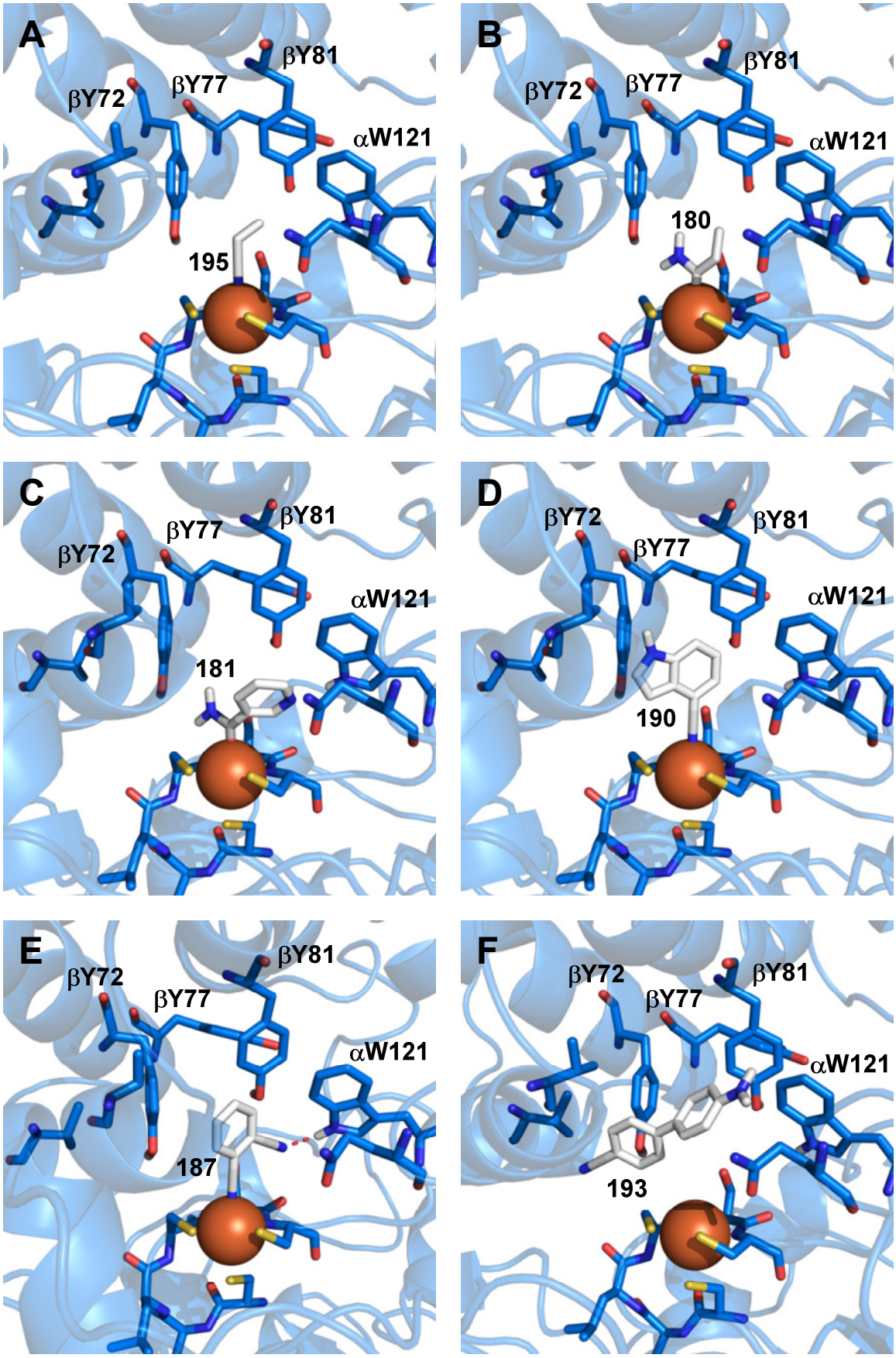


Figure S 19: Binding modes of additional NHase substrates and products generated using AutoDock 4.2 and the Equi_NHase homology model created herein.

11. References

(1) Lohr, T. L.; Marks, T. J. Orthogonal Tandem Catalysis. *Nature Chemistry* **2015,** *7* (6), 477-482.

(2) France, S. P.; Hepworth, L. J.; Turner, N. J.; Flitsch, S. L. Constructing Biocatalytic Cascades: In Vitro and in Vivo Approaches to De Novo Multi-Enzyme Pathways. *ACS Catalysis* **2017,** *7* (1), 710-724.

(3) Gröger, H.; Hummel, W. Combining the 'Two Worlds' of Chemocatalysis and Biocatalysis Towards Multi-Step One-Pot Processes in Aqueous Media. *Current Opinion in Chemical Biology* **2014,** *19* (0), 171-179.

(4) Wang, Y.; Zhou, H. Tandem Reactions Combining Biocatalysts and Chemical Catalysts for Asymmetric Synthesis. *Catalysts* **2016,** *6* (12), 194.

(5) Peterson, E. A.; Dillon, B.; Raheem, I.; Richardson, P.; Richter, D.; Schmidt, R.; Sneddon, H. F. Sustainable Chromatography (an Oxymoron?). *Green Chemistry* **2014,** *16* (9), 4060-4075.

(6) Sheldon, R. A. Fundamentals of Green Chemistry: Efficiency in Reaction Design. *Chem Soc Rev* **2012,** *41* (4), 1437-1451.

(7) Tietze, L. F. Domino Reactions in Organic Synthesis. *Chemical Reviews* **1996,** *96* (1), 115-136.

(8) Takahashi, T.; Katouda, W.; Sakamoto, Y.; Tomida, S.; Yamada, H. Stereochemical Prediction for Tandem Radical Cyclization Based on Mm2 Transition State Model. New Approach to Steroid Cd-Ring. *Tetrahedron Letters* **1995,** *36* (13), 2273-2276.

(9) Kim, K.; Vasu, D.; Im, H.; Hong, S. Palladium(II)-Catalyzed Tandem Synthesis of Acenes Using Carboxylic Acids as Traceless Directing Groups. *Angewandte Chemie International Edition* **2016,** *55* (30), 8652-8655.

(10) Johnston, A. J. S.; Ling, K. B.; Sale, D.; Lebrasseur, N.; Larrosa, I. Direct Ortho-Arylation of Pyridinecarboxylic Acids: Overcoming the Deactivating Effect of sp_2-Nitrogen. *Organic Letters* **2016,** *18* (23), 6094-6097.

(11) Teskey, C. J.; Lui, A. Y. W.; Greaney, M. F. Ruthenium-Catalyzed Meta-Selective C-H Bromination. *Angewandte Chemie International Edition* **2015,** *54* (40), 11677-11680.

(12) Huff, C. A.; Sanford, M. S. Cascade Catalysis for the Homogeneous Hydrogenation of CO_2 to Methanol. *Journal of the American Chemical Society* **2011,** *133* (45), 18122-18125.

(13) Helms, B.; Guillaudeu, S. J.; Xie, Y.; McMurdo, M.; Hawker, C. J.; Fréchet, J. M. J. One-Pot Reaction Cascades Using Star Polymers with Core-Confined Catalysts. *Angewandte Chemie International Edition* **2005,** *44* (39), 6384-6387.

(14) Lu, J.; Dimroth, J.; Weck, M. Compartmentalization of Incompatible Catalytic Transformations for Tandem Catalysis. *Journal of the American Chemical Society* **2015,** *137* (40), 12984-12989.

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

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

(17) France, S. P.; Hussain, S.; Hill, A. M.; Hepworth, L. J.; Howard, R. M.; Mulholland, K. R.; Flitsch, S. L.; Turner, N. J. One-Pot Cascade Synthesis of Mono- and Disubstituted Piperidines and Pyrrolidines Using Carboxylic Acid Reductase (Car), Ω-Transaminase (Ω-Ta), and Imine Reductase (IRed) Biocatalysts. *ACS Catalysis* **2016,** *6* (6), 3753-3759.

(18) Mutti, F. G.; Knaus, T.; Scrutton, N. S.; Breuer, M.; Turner, N. J. Conversion of Alcohols to Enantiopure Amines through Dual-Enzyme Hydrogen-Borrowing Cascades. *Science* **2015,** *349* (6255), 1525-1529.

(19) Both, P.; Busch, H.; Kelly, P. P.; Mutti, F. G.; Turner, N. J.; Flitsch, S. L. Whole-Cell Biocatalysts for Stereoselective C−H Amination Reactions. *Angewandte Chemie International Edition* **2016,** *55* (4), 1511-1513.

(20) Pàmies, O.; Bäckvall, J.-E. Combination of Enzymes and Metal Catalysts. A Powerful Approach in Asymmetric Catalysis. *Chemical Reviews* **2003,** *103* (8), 3247-3262.

(21) Långvik, O.; Saloranta, T.; Murzin, D. Y.; Leino, R. Heterogeneous Chemoenzymatic Catalyst Combinations for One-Pot Dynamic Kinetic Resolution Applications. *ChemCatChem* **2015,** *7* (24), 4004-4015.

(22) Martín-Matute, B.; Edin, M.; Bogár, K.; Bäckvall, J.-E. Highly Compatible Metal and Enzyme Catalysts for Efficient Dynamic Kinetic Resolution of Alcohols at Ambient Temperature. *Angewandte Chemie International Edition* **2004,** *43* (47), 6535-6539.

(23) Martín-Matute, B.; Edin, M.; Bogár, K.; Kaynak, F. B.; Bäckvall, J.-E. Combined Ruthenium(II) and Lipase Catalysis for Efficient Dynamic Kinetic Resolution of Secondary Alcohols. Insight into the Racemization Mechanism. *Journal of the American Chemical Society* **2005,** *127* (24), 8817-8825.

(24) Heidlindemann, M.; Rulli, G.; Berkessel, A.; Hummel, W.; Gröger, H. Combination of Asymmetric Organo- and Biocatalytic Reactions in Organic Media Using Immobilized Catalysts in Different Compartments. *ACS Catalysis* **2014,** *4* (4), 1099-1103.

(25) Rulli, G.; Heidlindemann, M.; Berkessel, A.; Hummel, W.; Gröger, H. Towards Catalyst Compartimentation in Combined Chemo- and Biocatalytic Processes: Immobilization of Alcohol Dehydrogenases for the Diastereoselective Reduction of a β-Hydroxy Ketone Obtained from an Organocatalytic Aldol Reaction. *Journal of Biotechnology* **2013,** *168* (3), 271-276.

(26) Burda, E.; Hummel, W.; Gröger, H. Modular Chemoenzymatic One-Pot Syntheses in Aqueous Media: Combination of a Palladium-Catalyzed Cross-Coupling with an Asymmetric Biotransformation. *Angewandte Chemie International Edition* **2008,** *47* (49), 9551-9554.

(27) Schnapperelle, I.; Hummel, W.; Gröger, H. Formal Asymmetric Hydration of Non-Activated Alkenes in Aqueous Medium through a “Chemoenzymatic Catalytic System”. *Chemistry – A European Journal* **2012,** *18* (4), 1073-1076.

(28) Boffi, A.; Cacchi, S.; Ceci, P.; Cirilli, R.; Fabrizi, G.; Prastaro, A.; Niembro, S.; Shafir, A.; Vallribera, A. The Heck Reaction of Allylic Alcohols Catalyzed by Palladium Nanoparticles in Water: Chemoenzymatic Synthesis of (R)-(−)-Rhododendrol. *ChemCatChem* **2011,** *3* (2), 347-353.

(29) Huang, H.; Denard, C. A.; Alamillo, R.; Crisci, A. J.; Miao, Y.; Dumesic, J. A.; Scott, S. L.; Zhao, H. Tandem Catalytic Conversion of Glucose to 5-Hydroxymethylfurfural with an Immobilized Enzyme and a Solid Acid. *ACS Catalysis* **2014,** *4* (7), 2165-2168.

(30) Huang, R.; Qi, W.; Su, R.; He, Z. Integrating Enzymatic and Acid Catalysis to Convert Glucose into 5-Hydroxymethylfurfural. *Chemical Communications* **2010,** *46* (7), 1115-1117.

(31) Simeonov, S. P.; Coelho, J. A. S.; Afonso, C. A. M. Integrated Chemo-Enzymatic Production of 5-Hydroxymethylfurfural from Glucose. *ChemSusChem* **2013,** *6* (6), 997-1000.

(32) Gómez Baraibar, Á.; Reichert, D.; Mügge, C.; Seger, S.; Gröger, H.; Kourist, R. A One-Pot Cascade Reaction Combining an Encapsulated Decarboxylase with a Metathesis Catalyst

for the Synthesis of Bio-Based Antioxidants. *Angewandte Chemie International Edition* **2016,** *55* (47), 14823-14827.

(33) Suljić, S.; Pietruszka, J. Synthesis of 3-Arylated 3,4-Dihydrocoumarins: Combining Continuous Flow Hydrogenation with Laccase-Catalysed Oxidation. *Advanced Synthesis & Catalysis* **2014,** *356* (5), 1007-1020.

(34) Fink, M. J.; Schön, M.; Rudroff, F.; Schnürch, M.; Mihovilovic, M. D. Single Operation Stereoselective Synthesis of Aerangis Lactones: Combining Continuous Flow Hydrogenation and Biocatalysts in a Chemoenzymatic Sequence. *ChemCatChem* **2013,** *5* (3), 724-727.

(35) Simon, R. C.; Busto, E.; Richter, N.; Resch, V.; Houk, K. N.; Kroutil, W. Biocatalytic Trifluoromethylation of Unprotected Phenols. *Nature Communications* **2016,** *7*, 13323.

(36) Long, T. R.; Bowden, N. B. Polydimethylsiloxane Thimbles. *Encyclopedia of Reagents for Organic Synthesis* **2001**.

(37) Millerli, A. L.; Bowden, N. B. Site-Isolation and Recycling of $PdCl_2$ Using Pdms Thimbles. *The Journal of Organic Chemistry* **2009,** *74* (13), 4834-4840.

(38) Runge, M. B.; Mwangi, M. T.; Miller, A. L.; Perring, M.; Bowden, N. B. Cascade Reactions Using $LiAlH_4$ and Grignard Reagents in the Presence of Water. *Angewandte Chemie International Edition* **2008,** *47* (5), 935-939.

(39) Sato, H.; Hummel, W.; Gröger, H. Cooperative Catalysis of Noncompatible Catalysts through Compartmentalization: Wacker Oxidation and Enzymatic Reduction in a One-Pot Process in Aqueous Media. *Angewandte Chemie International Edition* **2015,** *54* (15), 4488-4492.

(40) Denard, C. A.; Bartlett, M. J.; Wang, Y.; Lu, L.; Hartwig, J. F.; Zhao, H. Development of a One-Pot Tandem Reaction Combining Ruthenium-Catalyzed Alkene Metathesis and Enantioselective Enzymatic Oxidation to Produce Aryl Epoxides. *ACS Catalysis* **2015,** *5* (6), 3817-3822.

(41) Denard, C. A.; Huang, H.; Bartlett, M. J.; Lu, L.; Tan, Y.; Zhao, H.; Hartwig, J. F. Cooperative Tandem Catalysis by an Organometallic Complex and a Metalloenzyme. *Angewandte Chemie International Edition* **2014,** *53* (2), 465-469.

(42) Kohler, V.; Turner, N. J. Artificial Concurrent Catalytic Processes Involving Enzymes. *Chemical Communications* **2015,** *51* (3), 450-464.

(43) Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. Directed Evolution of Artificial Metalloenzymes for in Vivo Metathesis. *Nature* **2016,** *537* (7622), 661-665.

(44) Köhler V; Wilson, Y. M.; Dürrenberger M; Ghislieri D; Churakova E; Quinto T; Knörr L; Häussinger D; Hollmann F; Turner, N. J.et al. Synthetic Cascades Are Enabled by Combining Biocatalysts with Artificial Metalloenzymes. *Nat Chem* **2013,** *5* (2), 93-99.

(45) Yusop, R. M.; Unciti-Broceta, A.; Johansson, E. M. V.; Sánchez-Martín, R. M.; Bradley, M. Palladium-Mediated Intracellular Chemistry. *Nature Chemistry* **2011,** *3* (3), 239-243.

(46) Johansson Seechurn, C. C. C.; Kitching, M. O.; Colacot, T. J.; Snieckus, V. Palladium-Catalyzed Cross-Coupling: A Historical Contextual Perspective to the 2010 Nobel Prize. *Angewandte Chemie International Edition* **2012,** *51* (21), 5062-5085.

(47) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Palladium-Catalyzed Cross-Coupling Reactions in Total Synthesis. *Angewandte Chemie International Edition* **2005,** *44* (29), 4442-4489.

(48) Ford, M. C.; Ho, P. S. Computational Tools to Model Halogen Bonds in Medicinal Chemistry. *Journal of Medicinal Chemistry* **2016,** *59* (5), 1655-1670.

(49) Zhou, Y.; Wang, J.; Gu, Z.; Wang, S.; Zhu, W.; Aceña, J. L.; Soloshonok, V. A.; Izawa, K.; Liu, H. Next Generation of Fluorine-Containing Pharmaceuticals, Compounds Currently in Phase II–III Clinical Trials of Major Pharmaceutical Companies: New Structural Trends and Therapeutic Areas. *Chemical Reviews* **2016,** *116* (2), 422-518.

(50) Harris, C. M.; Kannan, R.; Kopecka, H.; Harris, T. M. The Role of the Chlorine Substituents in the Antibiotic Vancomycin - Preparation and Characterization of Monodechlorovancomycin and Didechlorovancomycin. *Journal of the American Chemical Society* **1985,** *107* (23), 6652-6658.

(51) Jeschke, P. The Unique Role of Halogen Substituents in the Design of Modern Agrochemicals. *Pest Management Science* **2010,** *66* (1), 10-27.

(52) Jeschke, P. The Unique Role of Fluorine in the Design of Active Ingredients for Modern Crop Protection. *ChemBioChem* **2004,** *5* (5), 570-589.

(53) Berger, G.; Soubhye, J.; Meyer, F. Halogen Bonding in Polymer Science: From Crystal Engineering to Functional Supramolecular Polymers and Materials. *Polymer Chemistry* **2015,** *6* (19), 3559-3580.

(54) Tang, M. L.; Bao, Z. Halogenated Materials as Organic Semiconductors. *Chemistry of Materials* **2011,** *23* (3), 446-455.

(55) Voskressensky, L. G.; Golantsov, N. E.; Maharramov, A. M. Recent Advances in Bromination of Aromatic and Heteroaromatic Compounds. *Synthesis* **2016,** *48* (05), 615-643.

(56) Withers, R. M. J.; Lees, F. P. The Assessment of Major Hazards: The Lethal Toxicity of Bromine. *Journal of Hazardous Materials* **1986,** *13* (3), 279-299.

(57) Kiyoshi, T.; Tsuneo, S.; Yoko, N.; Koko, S.; Takaaki, H. Halogenation of Aromatic Compounds by N-Chloro-, N-Bromo-, and N-Iodosuccinimide. *Chemistry Letters* **2003,** *32* (10), 932-933.

(58) Mo, F.; Yan, J. M.; Qiu, D.; Li, F.; Zhang, Y.; Wang, J. Gold-Catalyzed Halogenation of Aromatics by N-Halosuccinimides. *Angewandte Chemie International Edition* **2010,** *49* (11), 2028-2032.

(59) WHO *DDT and Its Derivatives: Environmental Aspects*; World Health Organisation: Geneva, 1989.

(60) Kürti, L.; Czakó, B. *Strategic Applications of Named Reactions in Organic Synthesis*; Elsevier: Oxford, 2005.

(61) Johnson, R. G.; Ingham, R. K. The Degradation of Carboxylic Acid Salts by Means of Halogen - the Hunsdiecker Reaction. *Chemical Reviews* **1956,** *56* (2), 219-269.

(62) Tsuji, J.; Ohno, K. Organic Syntheses by Means of Noble Metal Compounds Xxi. Decarbonylation of Aldehydes Using Rhodium Complex. *Tetrahedron Letters* **1965,** *6* (44), 3969-3971.

(63) Hodgson, H. H. The Sandmeyer Reaction. *Chemical Reviews* **1947,** *40* (2), 251-277.

(64) Laali, K. K.; Gettwert, V. J. Fluorodediazoniation in Ionic Liquid Solvents: New Life for the Balz–Schiemann Reaction. *Journal of Fluorine Chemistry* **2001,** *107* (1), 31-34.

(65) Macklin, T. K.; Snieckus, V. Directed Ortho Metalation Methodology. The N,N-Dialkyl Aryl O-Sulfamate as a New Directed Metalation Group and Cross-Coupling Partner for Grignard Reagents. *Organic Letters* **2005,** *7* (13), 2519-2522.

(66) Nguyen, T.-H.; Castanet, A.-S.; Mortier, J. Directed Ortho-Metalation of Unprotected Benzoic Acids. Methodology and Regioselective Synthesis of Useful Contiguously 3- and

6-Substituted 2-Methoxybenzoic Acid Building Blocks. *Organic Letters* **2006,** *8* (4), 765-768.

(67) Sandtorv, A. H. Transition Metal-Catalyzed C-H Activation of Indoles. *Advanced Synthesis & Catalysis* **2015,** *357* (11), 2403-2435.

(68) Chen, X.; Engle, K. M.; Wang, D. H.; Yu, J. Q. Palladium(II)-Catalyzed C-H Activation/C-C Cross-Coupling Reactions: Versatility and Practicality. *Angewandte Chemie International Edition* **2009,** *48* (28), 5094-5115.

(69) Kalyani, D.; Dick, A. R.; Anani, W. Q.; Sanford, M. S. A Simple Catalytic Method for the Regioselective Halogenation of Arenes. *Organic Letters* **2006,** *8* (12), 2523-2526.

(70) Du, B.; Jiang, X.; Sun, P. Palladium-Catalyzed Highly Selective Ortho-Halogenation (I, Br, Cl) of Arylnitriles Via sp2 C–H Bond Activation Using Cyano as Directing Group. *The Journal of Organic Chemistry* **2013,** *78* (6), 2786-2791.

(71) Sun, X.; Shan, G.; Sun, Y.; Rao, Y. Regio- and Chemoselective C-H Chlorination/Bromination of Electron-Deficient Arenes by Weak Coordination and Study of Relative Directing-Group Abilities. *Angewandte Chemie International Edition* **2013,** *52* (16), 4440-4444.

(72) Tian, Q.; Chen, X.; Liu, W.; Wang, Z.; Shi, S.; Kuang, C. Regioselective Halogenation of 2-Substituted-1,2,3-Triazoles Via sp2 C-H Activation. *Organic & Biomolecular Chemistry* **2013,** *11* (45), 7830-7833.

(73) Gribble, G. W. Natural Organohalogens: A New Frontier for Medicinal Agents? *Journal of Chemical Education* **2004,** *81* (10), 1441.

(74) Smith, D. R. M.; Grüschow, S.; Goss, R. J. M. Scope and Potential of Halogenases in Biosynthetic Applications. *Current Opinion in Chemical Biology* **2013,** *17* (2), 276-283.

(75) Brown, S.; O'Connor, S. E. Halogenase Engineering for the Generation of New Natural Product Analogues. *ChemBioChem* **2015,** *16* (15), 2129-2135.

(76) Weichold, V.; Milbredt, D.; van Pée, K.-H. Specific Enzymatic Halogenation—from the Discovery of Halogenated Enzymes to Their Applications in Vitro and in Vivo. *Angewandte Chemie International Edition* **2016,** *55* (22), 6374-6389.

(77) O'Hagan, D.; Deng, H. Enzymatic Fluorination and Biotechnological Developments of the Fluorinase. *Chemical Reviews* **2015,** *115* (2), 634-649.

(78) Latham, J.; Brandenburger, E.; Menon, B. R.; Shepherd, S. A.; Micklefield, J. Development of Halogenase Enzymes for Use in Synthesis. *Chemical Reviews* **2017,** *ASAP,* DOI: 10.1021/acs.chemrev.7b00032.

(79) Nuell, M. J.; Fang, G. H.; Axley, M. J.; Kenigsberg, P.; Hager, L. P. Isolation and Nucleotide Sequence of the Chloroperoxidase Gene from Caldariomyces Fumago. *Journal of Bacteriology* **1988,** *170* (2), 1007-1011.

(80) Morris, D. R.; Hager, L. P. Chloroperoxidase I: Isolation and Properties of Crystalline Glycoprotein. *Journal of Biological Chemistry* **1966,** *241* (8), 1763-&.

(81) Hager, L. P.; Morris, D. R.; Brown, F. S.; Eberwein, H. Chloroperoxidase Ii: Utilisation of Halogen Anions. *Journal of Biological Chemistry* **1966,** *241* (8), 1769-&.

(82) Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. Halogenation Strategies in Natural Product Biosynthesis. *Chemistry & Biology* **2008,** *15* (2), 99-109.

(83) Messerschmidt, A.; Prade, L.; Wever, R. Implications for the Catalytic Mechanism of the Vanadium-Containing Enzyme Chloroperoxidase from the Fungus Curvularia Inaequalis by

X-Ray Structures of the Native and Peroxide Form. *Biological Chemistry* **1997,** *378* (3-4), 309-315.

(84) Littlechild, J.; Garcia Rodriguez, E.; Isupov, M. Vanadium Containing Bromoperoxidase--Insights into the Enzymatic Mechanism Using X-Ray Crystallography. *J Inorg Biochem* **2009,** *103* (4), 617-621.

(85) Itoh, N.; Hasan, A. K. M. Q.; Izumi, Y.; Yamada, H. Substrate Specificity, Regiospecificity and Stereospecificity of Halogenation Reactions Catalyzed by Non-Heme-Type Bromoperoxidase of Corallina Pilulifera. *European Journal of Biochemistry* **1988,** *172* (2), 477-484.

(86) Wischang, D.; Brücher, O.; Hartung, J. Bromoperoxidases and Functional Enzyme Mimics as Catalysts for Oxidative Bromination—a Sustainable Synthetic Approach. *Coordination Chemistry Reviews* **2011,** *255* (19–20), 2204-2217.

(87) Bernhardt, P.; Okino, T.; Winter, J. M.; Miyanaga, A.; Moore, B. S. A Stereoselective Vanadium-Dependent Chloroperoxidase in Bacterial Antibiotic Biosynthesis. *Journal of the American Chemical Society* **2011,** *133* (12), 4268-4270.

(88) Carter-Franklin, J. N.; Butler, A. Vanadium Bromoperoxidase-Catalyzed Biosynthesis of Halogenated Marine Natural Products. *Journal of the American Chemical Society* **2004,** *126* (46), 15060-15066.

(89) Kaysser, L.; Bernhardt, P.; Nam, S. J.; Loesgen, S.; Ruby, J. G.; Skewes-Cox, P.; Jensen, P. R.; Fenical, W.; Moore, B. S. Merochlorins a-D, Cyclic Meroterpenoid Antibiotics Biosynthesized in Divergent Pathways with Vanadium-Dependent Chloroperoxidases. *Journal of the American Chemical Society* **2012,** *134* (29), 11988-11991.

(90) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. Syrb2 in Syringomycin E Biosynthesis Is a Nonheme Fe(II) α-Ketoglutarate- and O_2-Dependent Halogenase. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (29), 10111-10116.

(91) Galonić, D. P.; Vaillancourt, F. H.; Walsh, C. T. Halogenation of Unactivated Carbon Centers in Natural Product Biosynthesis: Trichlorination of Leucine During Barbamide Biosynthesis. *Journal of the American Chemical Society* **2006,** *128* (12), 3900-3901.

(92) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. Cryptic Chlorination by a Non-Haem Iron Enzyme During Cyclopropyl Amino Acid Biosynthesis. *Nature* **2005,** *436* (7054), 1191-1194.

(93) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. Crystal Structure of the Non-Haem Iron Halogenase SyrB2 in Syringomycin Biosynthesis. *Nature* **2006,** *440* (7082), 368-371.

(94) Hoffart, L. M.; Barr, E. W.; Guyer, R. B.; Bollinger, J. M.; Krebs, C. Direct Spectroscopic Detection of a C-H-Cleaving High-Spin Fe(IV) Complex in a Prolyl-4-Hydroxylase. *Proceedings of the National Academy of Sciences* **2006,** *103* (40), 14738-14743.

(95) Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C. The First Direct Characterization of a High-Valent Iron Intermediate in the Reaction of an α-Ketoglutarate-Dependent Dioxygenase: A High-Spin Fe(IV) Complex in Taurine/α-Ketoglutarate Dioxygenase (TauD) from Escherichia Coli. *Biochemistry* **2003,** *42* (24), 7497-7508.

(96) Puri, M.; Biswas, A. N.; Fan, R.; Guo, Y.; Que, L. Modeling Non-Heme Iron Halogenases: High-Spin Oxoiron(IV)–Halide Complexes That Halogenate C–H Bonds. *Journal of the American Chemical Society* **2016,** *138* (8), 2484-2487.

(97) Matthews, M. L.; Neumann, C. S.; Miles, L. A.; Grove, T. L.; Booker, S. J.; Krebs, C.; Walsh, C. T.; Bollinger, J. M. Substrate Positioning Controls the Partition between

Halogenation and Hydroxylation in the Aliphatic Halogenase, SyrB2. *Proceedings of the National Academy of Sciences* **2009,** *106* (42), 17723-17728.

(98) Hillwig, M. L.; Liu, X. A New Family of Iron-Dependent Halogenases Acts on Freestanding Substrates. *Nat Chem Biol* **2014,** *10* (11), 921-923.

(99) Hillwig, M. L.; Fuhrman, H. A.; Ittiamornkul, K.; Sevco, T. J.; Kwak, D. H.; Liu, X. Identification and Characterization of a Welwitindolinone Alkaloid Biosynthetic Gene Cluster in the Stigonematalean Cyanobacterium Hapalosiphon Welwitschii. *ChemBioChem* **2014,** *15* (5), 665-669.

(100) Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron- and 2-Oxo-Glutarate-Dependent Enzyme Welo5. *Nat Chem Biol* **2016,** *12* (8), 636-640.

(101) Mitchell, A. J.; Dunham, N. P.; Bergman, J. A.; Wang, B.; Zhu, Q.; Chang, W.-c.; Liu, X.; Boal, A. K. Structure-Guided Reprogramming of a Hydroxylase to Halogenate Its Small Molecule Substrate. *Biochemistry* **2017,** *56* (3), 441-444.

(102) Maguire, A. R.; Meng, W.-D.; Roberts, S. M.; Willetts, A. J. Synthetic Approaches Towards Nucleocidin and Selected Analogues; Anti-HIV Activity in 4'-Fluorinated Nucleoside Derivatives. *Journal of the Chemical Society, Perkin Transactions 1* **1993,** (15), 1795-1808.

(103) O'Hagan, D.; Schaffrath, C.; Cobb, S. L.; Hamilton, J. T. G.; Murphy, C. D. Biochemistry: Biosynthesis of an Organofluorine Molecule. *Nature* **2002,** *416* (6878), 279-279.

(104) Deng, H.; Ma, L.; Bandaranayaka, N.; Qin, Z.; Mann, G.; Kyeremeh, K.; Yu, Y.; Shepherd, T.; Naismith, J. H.; O'Hagan, D. Identification of Fluorinases from Streptomyces Sp MA37, Norcardia Brasiliensis, and Actinoplanes Sp N902-109 by Genome Mining. *ChemBioChem* **2014,** *15* (3), 364-368.

(105) Wang, Y.; Deng, Z.; Qu, X. Characterization of a Sam-Dependent Fluorinase from a Latent Biosynthetic Pathway for Fluoroacetate and 4-Fluorothreonine Formation in Nocardia Brasiliensis. *F1000Research* **2014,** *61*, 1.

(106) Huang, S.; Ma, L.; Tong, M. H.; Yu, Y.; O'Hagan, D.; Deng, H. Fluoroacetate Biosynthesis from the Marine-Derived Bacterium Streptomyces Xinghaiensis NRRL B-24674. *Organic & Biomolecular Chemistry* **2014,** *12* (27), 4828-4831.

(107) Dong, C.; Huang, F.; Deng, H.; Schaffrath, C.; Spencer, J. B.; O'Hagan, D.; Naismith, J. H. Crystal Structure and Mechanism of a Bacterial Fluorinating Enzyme. *Nature* **2004,** *427* (6974), 561-565.

(108) Cadicamo, C. D.; Courtieu, J.; Deng, H.; Meddour, A.; O'Hagan, D. Enzymatic Fluorination in Streptomyces Cattleya Takes Place with an Inversion of Configuration Consistent with an Sn2 Reaction Mechanism. *ChemBioChem* **2004,** *5* (5), 685-690.

(109) O'Hagan, D.; Goss, R. J. M.; Meddour, A.; Courtieu, J. Assay for the Enantiomeric Analysis of [^{2}H1]-Fluoroacetic Acid: Insight into the Stereochemical Course of Fluorination During Fluorometabolite Biosynthesis in Streptomyces Cattleya. *Journal of the American Chemical Society* **2003,** *125* (2), 379-387.

(110) Lohman, D. C.; Edwards, D. R.; Wolfenden, R. Catalysis by Desolvation: The Catalytic Prowess of SAM-Dependent Halide-Alkylating Enzymes. *Journal of the American Chemical Society* **2013,** *135* (39), 14473-14475.

(111) Eustaquio, A. S.; Pojer, F.; Noel, J. P.; Moore, B. S. Discovery and Characterization of a Marine Bacterial SAM-Dependent Chlorinase. *Nat Chem Biol* **2008,** *4* (1), 69-74.

(112) Deng, H.; Cobb, S. L.; McEwan, A. R.; McGlinchey, R. P.; Naismith, J. H.; O'Hagan, D.; Robinson, D. A.; Spencer, J. B. The Fluorinase from Streptomyces Cattleya Is Also a Chlorinase. *Angewandte Chemie International Edition* **2006,** *45* (5), 759-762.

(113) Deng, H.; McMahon, S. A.; Eustaquio, A. S.; Moore, B. S.; Naismith, J. H.; O'Hagan, D. Mechanistic Insights into Water Activation in Sam Hydroxide Adenosyltransferase (DUF-62). *ChemBioChem* **2009,** *10* (15), 2455-2459.

(114) Deng, H.; Cobb, S. L.; Gee, A. D.; Lockhart, A.; Martarello, L.; McGlinchey, R. P.; O'Hagan, D.; Onega, M. Fluorinase Mediated C-^{18}F Bond Formation, an Enzymatic Tool for PET Labelling. *Chemical Communications* **2006,** *6* (6), 652-654.

(115) Winkler, M.; Domarkas, J.; Schweiger, L. F.; O'Hagan, D. Fluorinase-Coupled Base Swaps: Synthesis of [^{18}F]-5′-Deoxy-5′-Fluorouridines. *Angewandte Chemie International Edition* **2008,** *47* (52), 10141-10143.

(116) Onega, M.; Domarkas, J.; Deng, H.; Schweiger, L. F.; Smith, T. A. D.; Welch, A. E.; Plisson, C.; Gee, A. D.; O'Hagan, D. An Enzymatic Route to 5-Deoxy-5-[^{18}F]Fluoro-D-Ribose, a [^{18}F]-Fluorinated Sugar for Pet Imaging. *Chemical Communications* **2010,** *46* (1), 139-141.

(117) Li, X.-G.; Domarkas, J.; O'Hagan, D. Fluorinase Mediated Chemoenzymatic Synthesis of [^{18}F]-Fluoroacetate. *Chemical Communications* **2010,** *46* (41), 7819-7821.

(118) Thompson, S.; Zhang, Q.; Onega, M.; McMahon, S.; Fleming, I.; Ashworth, S.; Naismith, J. H.; Passchier, J.; O'Hagan, D. A Localized Tolerance in the Substrate Specificity of the Fluorinase Enzyme Enables "Last-Step" ^{18}F Fluorination of a Rgd Peptide under Ambient Aqueous Conditions. *Angewandte Chemie International Edition* **2014,** *53* (34), 8913-8918.

(119) Eustáquio, A. S.; O'Hagan, D.; Moore, B. S. Engineering Fluorometabolite Production: Fluorinase Expression in Salinispora Tropica Yields Fluorosalinosporamide. *Journal of Natural Products* **2010,** *73* (3), 378-382.

(120) Holtmann, D.; Hollmann, F. The Oxygen Dilemma: A Severe Challenge for the Application of Monooxygenases? *ChemBioChem* **2016,** *17* (15), 1391-1398.

(121) Huijbers, M. M.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J. Flavin Dependent Monooxygenases. *Arch Biochem Biophys* **2014,** *544*, 2-17.

(122) Yeh, E.; Garneau, S.; Walsh, C. T. Robust in Vitro Activity of RebF and RebH, a Two-Component Reductase/Halogenase, Generating 7-Chlorotryptophan During Rebeccamycin Biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (11), 3960-3965.

(123) Keller, S.; Wage, T.; Hohaus, K.; Hölzer, M.; Eichhorn, E.; van Pée, K.-H. Purification and Partial Characterization of Tryptophan 7-Halogenase (PrnA) from Pseudomonas Fluorescens. *Angewandte Chemie International Edition* **2000,** *39* (13), 2300-2302.

(124) Zhu, X.; De Laurentis, W.; Leang, K.; Herrmann, J.; Ihlefeld, K.; van Pée, K.-H.; Naismith, J. H. Structural Insights into Regioselectivity in the Enzymatic Chlorination of Tryptophan. *Journal of Molecular Biology* **2009,** *391* (1), 74-85.

(125) Dong, C.; Kotzsch, A.; Dorward, M.; van Pee, K.-H.; Naismith, J. H. Crystallization and X-Ray Diffraction of a Halogenating Enzyme, Tryptophan 7-Halogenase, from Pseudomonas Fluorescens. *Acta Crystallographica Section D* **2004,** *60* (8), 1438-1440.

(126) Yeh, E.; Cole, L. J.; Barr, E. W.; Bollinger, J. M., Jr.; Ballou, D. P.; Walsh, C. T. Flavin Redox Chemistry Precedes Substrate Chlorination During the Reaction of the Flavin-Dependent Halogenase RebH. *Biochemistry* **2006,** *45* (25), 7904-7912.

(127) Yeh, E.; Blasiak, L. C.; Koglin, A.; Drennan, C. L.; Walsh, C. T. Chlorination by a Long-Lived Intermediate in the Mechanism of Flavin-Dependent Halogenases. *Biochemistry* **2007,** *46* (5), 1284-1292.

(128) Flecks, S.; Patallo, E. P.; Zhu, X.; Ernyei, A. J.; Seifert, G.; Schneider, A.; Dong, C.; Naismith, J. H.; van Pée, K.-H. New Insights into the Mechanism of Enzymatic Chlorination of Tryptophan. *Angewandte Chemie International Edition* **2008,** *47* (49), 9533-9536.

(129) Bitto, E.; Huang, Y.; Bingman, C. A.; Singh, S.; Thorson, J. S.; Phillips, G. N. The Structure of Flavin-Dependent Tryptophan 7-Halogenase RebH. *Proteins: Structure, Function, and Bioinformatics* **2008,** *70* (1), 289-293.

(130) Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K.-H.; Naismith, J. H. Tryptophan 7-Halogenase (PrnA) Structure Suggests a Mechanism for Regioselective Chlorination. *Science* **2005,** *309* (5744), 2216-2219.

(131) Kirner, S.; Hammer, P. E.; Hill, D. S.; Altmann, A.; Fischer, I.; Weislo, L. J.; Lanahan, M.; van Pee, K. H.; Ligon, J. M. Functions Encoded by Pyrrolnitrin Biosynthetic Genes from Pseudomonas Fluorescens. *J Bacteriol* **1998,** *180* (7), 1939-1943.

(132) Heemstra, J. R.; Walsh, C. T. Tandem Action of the O2- and Fadh2-Dependent Halogenases KtzQ and KtzR Produce 6,7-Dichlorotryptophan for Kutzneride Assembly. *Journal of the American Chemical Society* **2008,** *130* (43), 14024-14025.

(133) Milbredt, D.; Patallo, E. P.; van Pée, K.-H. A Tryptophan 6-Halogenase and an Amidotransferase Are Involved in Thienodolin Biosynthesis. *ChemBioChem* **2014,** *15* (7), 1011-1020.

(134) Zeng, J.; Zhan, J. Characterization of a Tryptophan 6-Halogenase from Streptomyces Toxytricini. *Biotechnol Lett* **2011,** *33* (8), 1607-1613.

(135) Xu, L.; Han, T.; Ge, M.; Zhu, L.; Qian, X. Discovery of the New Plant Growth-Regulating Compound Lyxlf2 Based on Manipulating the Halogenase in Amycolatopsis Orientalis. *Curr Microbiol* **2016,** *73* (3), 335-340.

(136) Smith, D. R. M.; Uria, A. R. R.; Helfrich, E. J. N.; Milbredt, D.; van-Pee, K.-H.; Piel, J.; Goss, R. J. M. An Unusual Flavin-Dependent Halogenase from the Metagenome of the Marine Sponge Theonella Swinhoei Wa. *ACS Chemical Biology* **2017**, DOI:10.1021/acschembio.6b01115.

(137) Zehner, S.; Kotzsch, A.; Bister, B.; Süssmuth, R. D.; Méndez, C.; Salas, J. A.; van Pée, K.-H. A Regioselective Tryptophan 5-Halogenase Is Involved in Pyrroindomycin Biosynthesis in Streptomyces Rugosporus LL-42D005. *Chemistry & Biology* **2005,** *12* (4), 445-452.

(138) Ryan, K. S. Biosynthetic Gene Cluster for the Cladoniamides, Bis-Indoles with a Rearranged Scaffold. *Plos One* **2011,** *6* (8).

(139) Ortega, M. A.; Cogan, D. P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G. N.; Maffioli, S. I.; Donadio, S.; Sosio, M.; Escano, J.et al. Two Flavoenzymes Catalyze the Post-Translational Generation of 5-Chlorotryptophan and 2-Aminovinyl-Cysteine During NAI-107 Biosynthesis. *ACS Chemical Biology* **2017,** *12* (2), 548-557.

(140) Wynands, I.; van Pee, K. H. A Novel Halogenase Gene from the Pentachloropseudilin Producer Actinoplanes Sp ATCC 33002 and Detection of in Vitro Halogenase Activity. *Fems Microbiology Letters* **2004,** *237* (2), 363-367.

(141) Mantovani, S. M.; Moore, B. S. Flavin-Linked Oxidase Catalyzes Pyrrolizine Formation of Dichloropyrrole-Containing Polyketide Extender Unit in Chlorizidine A. *Journal of the American Chemical Society* **2013,** *135* (48), 18032-18035.

(142) El Gamal, A.; Agarwal, V.; Diethelm, S.; Rahman, I.; Schorn, M. A.; Sneed, J. M.; Louie, G. V.; Whalen, K. E.; Mincer, T. J.; Noel, J. P.et al. Biosynthesis of Coral Settlement Cue Tetrabromopyrrole in Marine Bacteria by a Uniquely Adapted Brominase–Thioesterase Enzyme Pair. *Proceedings of the National Academy of Sciences* **2016,** *113* (14), 3797-3802.

(143) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. Biosynthesis of Polybrominated Aromatic Organic Compounds by Marine Bacteria. *Nature Chemical Biology* **2014,** *10* (8), 640-647.

(144) Zeng, J.; Zhan, J. A Novel Fungal Flavin-Dependent Halogenase for Natural Product Biosynthesis. *ChemBioChem* **2010,** *11* (15), 2119-2123.

(145) Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y. Insights into Radicicol Biosynthesis Via Heterologous Synthesis of Intermediates and Analogs. *Journal of Biological Chemistry* **2010,** *285* (53), 41412-41421.

(146) Zeng, J.; Lytle, A. K.; Gage, D.; Johnson, S. J.; Zhan, J. Specific Chlorination of Isoquinolines by a Fungal Flavin-Dependent Halogenase. *Bioorganic & Medicinal Chemistry Letters* **2013,** *23* (4), 1001-1003.

(147) Wang, S.; Xu, Y.; Maine, E. A.; Wijeratne, E. M.; Espinosa-Artiles, P.; Gunatilaka, A. A.; Molnar, I. Functional Characterization of the Biosynthesis of Radicicol, an Hsp90 Inhibitor Resorcylic Acid Lactone from Chaetomium Chiversii. *Chem Biol* **2008,** *15* (12), 1328-1338.

(148) Chooi, Y. H.; Cacho, R.; Tang, Y. Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from Penicillium Aethiopicum. *Chemistry & Biology* **2010,** *17* (5), 483-494.

(149) Chankhamjon, P.; Boettger-Schmidt, D.; Scherlach, K.; Urbansky, B.; Lackner, G.; Kalb, D.; Dahse, H. M.; Hoffmeister, D.; Hertweck, C. Biosynthesis of the Halogenated Mycotoxin Aspirochlorine in Koji Mold Involves a Cryptic Amino Acid Conversion. *Angew Chem Int Edit* **2014,** *53* (49), 13409-13413.

(150) Sato, M.; Winter, J. M.; Kishimoto, S.; Noguchi, H.; Tang, Y.; Watanabe, K. Combinatorial Generation of Chemical Diversity by Redox Enzymes in Chaetoviridin Biosynthesis. *Organic Letters* **2016,** *18* (6), 1446-1449.

(151) Ferrara, M.; Perrone, G.; Gambacorta, L.; Epifani, F.; Solfrizzo, M.; Gallo, A. Identification of a Halogenase Involved in the Biosynthesis of Ochratoxin a in Aspergillus Carbonarius. *Appl Environ Microbiol* **2016,** *82* (18), 5631-5641.

(152) Nielsen, M. T.; Nielsen, J. B.; Anyaogu, D. C.; Holm, D. K.; Nielsen, K. F.; Larsen, T. O.; Mortensen, U. H. Heterologous Reconstitution of the Intact Geodin Gene Cluster in Aspergillus Nidulans through a Simple and Versatile PCR Based Approach. *Plos One* **2013,** *8* (8).

(153) Podzelinska, K.; Latimer, R.; Bhattacharya, A.; Vining, L. C.; Zechel, D. L.; Jia, Z. Chloramphenicol Biosynthesis: The Structure of CmlS, a Flavin-Dependent Halogenase Showing a Covalent Flavin–Aspartate Bond. *Journal of Molecular Biology* **2010,** *397* (1), 316-331.

(154) Chankhamjon, P.; Tsunematsu, Y.; Ishida-Ito, M.; Sasa, Y.; Meyer, F.; Boettger-Schmidt, D.; Urbansky, B.; Menzel, K.-D.; Scherlach, K.; Watanabe, K.et al. Regioselective Dichlorination of a Non-Activated Aliphatic Carbon Atom and Phenolic Bismethylation by a Multifunctional Fungal Flavoenzyme. *Angewandte Chemie International Edition* **2016,** *55* (39), 11955-11959.

(155) Shepherd, S. A.; Karthikeyan, C.; Latham, J.; Struck, A.-W.; Thompson, M. L.; Menon, B. R. K.; Styles, M. Q.; Levy, C.; Leys, D.; Micklefield, J. Extending the Biocatalytic Scope of Regiocomplementary Flavin-Dependent Halogenase Enzymes. *Chemical Science* **2015,** *6* (6), 3454-3460.

(156) Shepherd, S. A.; Menon, B. R. K.; Fisk, H.; Struck, A.-W.; Levy, C.; Leys, D.; Micklefield, J. A Structure-Guided Switch in the Regioselectivity of a Tryptophan Halogenase. *ChemBioChem* **2016,** *17* (9), 821-824.

(157) Hölzer, M.; Burd, W.; Reißig, H.-U.; Pée, K.-H. v. Substrate Specificity and Regioselectivity of Tryptophan 7-Halogenase from Pseudomonas Fluorescens BL915. *Advanced Synthesis & Catalysis* **2001,** *343* (6-7), 591-595.

(158) Payne, J. T.; Poor, C. B.; Lewis, J. C. Directed Evolution of Rebh for Site-Selective Halogenation of Large Biologically Active Molecules. *Angewandte Chemie International Edition* **2015,** *54* (14), 4226-4230.

(159) Payne, J. T.; Andorfer, M. C.; Lewis, J. C. Regioselective Arene Halogenation Using the FAD-Dependent Halogenase RebH. *Angewandte Chemie International Edition* **2013,** *52* (20), 5271-5274.

(160) Frese, M.; Guzowska, P. H.; Voß, H.; Sewald, N. Regioselective Enzymatic Halogenation of Substituted Tryptophan Derivatives Using the Fad-Dependent Halogenase RebH. *ChemCatChem* **2014,** *6* (5), 1270-1276.

(161) Xu, F.; Merkley, A.; Yu, D.; Zhan, J. Selective Biochlorination of Hydroxyquinolines by a Flavin-Dependent Halogenase. *Tetrahedron Letters* **2016,** *57* (47), 5262-5265.

(162) Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C. Directed Evolution of RebH for Catalyst-Controlled Halogenation of Indole C-H Bonds. *Chemical Science* **2016,** *7* (6), 3720-3729.

(163) Payne, J. T.; Lewis, J. C. Upgrading Nature's Tools: Expression Enhancement and Preparative Utility of the Halogenase RebH. *Synlett* **2014,** *25*, A-E.

(164) Poor, C. B.; Andorfer, M. C.; Lewis, J. C. Improving the Stability and Catalyst Lifetime of the Halogenase RebH by Directed Evolution. *ChemBioChem* **2014,** *15* (9), 1286-1289.

(165) Schnepel, C.; Minges, H.; Frese, M.; Sewald, N. A High-Throughput Fluorescence Assay to Determine the Activity of Tryptophan Halogenases. *Angewandte Chemie International Edition* **2016,** *55* (45), 14159-14163.

(166) Hosford, J.; Shepherd, S. A.; Micklefield, J.; Wong, L. S. A High-Throughput Assay for Arylamine Halogenation Based on a Peroxidase-Mediated Quinone–Amine Coupling with Applications in the Screening of Enzymatic Halogenations. *Chemistry – A European Journal* **2014,** *20* (50), 16759-16763.

(167) Haki, G. D.; Rakshit, S. K. Developments in Industrially Important Thermostable Enzymes: A Review. *Bioresource Technology* **2003,** *89* (1), 17-34.

(168) Chen, X.; Zhang, B.; Zhang, W.; Wu, X.; Zhang, M.; Chen, T.; Liu, G.; Dyson, P. Genome Sequence of Streptomyces Violaceusniger Strain SPC6, a Halotolerant Streptomycete That Exhibits Rapid Growth and Development. *Genome Announcements* **2013,** *1* (4).

(169) Nagpure, A.; Gupta, R. K. Purification and Characterization of an Extracellular Chitinase from Antagonistic Streptomyces Violaceusniger. *Journal of Basic Microbiology* **2013,** *53* (5), 429-439.

(170) Takahashi, S.; Furuya, T.; Ishii, Y.; Kino, K.; Kirimura, K. Characterization of a Flavin Reductase from a Thermophilic Dibenzothiophene-Desulfurizing Bacterium, Bacillus Subtilis Wu-S2b. *Journal of Bioscience and Bioengineering* **2009,** *107* (1), 38-41.

(171) Menon, B. R. K.; Brandenburger, E.; Sharif, H. H.; Klemstein, U.; Greaney, M. F.; Micklefield, J. RadH: A Versitile Halogenase for Integration into Synthetic Pathways. *Submitted* **2016**.

(172) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry* **2009,** *30* (16), 2785-2791.

(173) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *Journal of the American Chemical Society* **2007,** *129* (33), 10110-10112.

(174) Belsare, K. D.; Andorfer, M. C.; Cardenas, F. S.; Chael, J. R.; Park, H. J.; Lewis, J. C. A Simple Combinatorial Codon Mutagenesis Method for Targeted Protein Engineering. *ACS Synthetic Biology* **2016**, DOI:10.1021/acssynbio.6b00297.

(175) Menon, B. R. K.; Latham, J.; Dunstan, M. S.; Brandenburger, E.; Klemstein, U.; Leys, D.; Karthikeyan, C.; Greaney, M. F.; Shepherd, S. A.; Micklefield, J. Structure and Biocatalytic Scope of Thermophilic Flavin-Dependent Halogenase and Flavin Reductase Enzymes. *Organic & Biomolecular Chemistry* **2016,** *14* (39), 9354-9361.

(176) Torborg, C.; Beller, M. Recent Applications of Palladium-Catalyzed Coupling Reactions in the Pharmaceutical, Agrochemical, and Fine Chemical Industries. *Advanced Synthesis & Catalysis* **2009,** *351* (18), 3027-3043.

(177) Ruiz-Castillo, P.; Buchwald, S. L. Applications of Palladium-Catalyzed C–N Cross-Coupling Reactions. *Chemical Reviews* **2016,** *116* (19), 12564-12649.

(178) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. A Convenient Catalyst for Aqueous and Protein Suzuki−Miyaura Cross-Coupling. *Journal of the American Chemical Society* **2009,** *131* (45), 16346-16347.

(179) Li, J.-H.; Zhang, X.-D.; Xie, Y.-X. Efficient and Copper-Free Sonogashira Cross-Coupling Reaction Catalyzed by $Pd(OAc)_2$/Pyrimidines Catalytic System. *European Journal of Organic Chemistry* **2005,** *2005* (20), 4256-4259.

(180) Spicer, C. D.; Triemer, T.; Davis, B. G. Palladium-Mediated Cell-Surface Labeling. *Journal of the American Chemical Society* **2011,** *134* (2), 800-803.

(181) Gao, Z.; Gouverneur, V.; Davis, B. G. Enhanced Aqueous Suzuki-Miyaura Coupling Allows Site-Specific Polypeptide ^{18}F-Labeling. *Journal of the American Chemical Society* **2013,** *135* (37), 13612-13615.

(182) Prastaro, A.; Ceci, P.; Chiancone, E.; Boffi, A.; Cirilli, R.; Colone, M.; Fabrizi, G.; Stringaro, A.; Cacchi, S. Suzuki-Miyaura Cross-Coupling Catalyzed by Protein-Stabilized Palladium Nanoparticles under Aerobic Conditions in Water: Application to a One-Pot Chemoenzymatic Enantioselective Synthesis of Chiral Biaryl Alcohols. *Green Chemistry* **2009,** *11* (12), 1929-1932.

(183) Roy, A. D.; Grüschow, S.; Cairns, N.; Goss, R. J. M. Gene Expression Enabling Synthetic Diversification of Natural Products: Chemogenetic Generation of Pacidamycin Analogs. *Journal of the American Chemical Society* **2010,** *132* (35), 12243-12245.

(184) Grüschow, S.; Rackham, E. J.; Elkins, B.; Newill, P. L. A.; Hill, L. M.; Goss, R. J. M. New Pacidamycin Antibiotics through Precursor-Directed Biosynthesis. *ChemBioChem* **2009,** *10* (2), 355-360.

(185) Sánchez, C.; Zhu, L.; Braña, A. F.; Salas, A. P.; Rohr, J.; Méndez, C.; Salas, J. A. Combinatorial Biosynthesis of Antitumor Indolocarbazole Compounds. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (2), 461-466.

(186) Glenn, W. S.; Nims, E.; O'Connor, S. E. Reengineering a Tryptophan Halogenase to Preferentially Chlorinate a Direct Alkaloid Precursor. *Journal of the American Chemical Society* **2011,** *133* (48), 19346-19349.

(187) Runguphan, W.; Qu, X.; O/'Connor, S. E. Integrating Carbon-Halogen Bond Formation into Medicinal Plant Metabolism. *Nature* **2010,** *468* (7322), 461-464.

(188) Corr, M. J.; Sharma, S. V.; Pubill-Ulldemolins, C.; Bown, R. T.; Poirot, P.; Smith, D. R. M.; Cartmell, C.; Abou Fayad, A.; Goss, R. J. M. Sonogashira Diversification of Unprotected Halotryptophans, Halotryptophan Containing Tripeptides; and Generation of a New to

Nature Bromo-Natural Product and Its Diversification in Water. *Chemical Science* **2017,** *8* (3), 2039-2046.

(189) Phipps, R. J.; Grimster, N. P.; Gaunt, M. J. Cu(Ii)-Catalyzed Direct and Site-Selective Arylation of Indoles under Mild Conditions. *Journal of the American Chemical Society* **2008,** *130* (26), 8172-8174.

(190) Willemse, T.; Van Imp, K.; Goss, R. J. M.; Van Vlijmen, H. W. T.; Schepens, W.; Maes, B. U. W.; Ballet, S. Suzuki-Miyaura Diversification of Amino Acids and Dipeptides in Aqueous Media. *ChemCatChem* **2015,** *7* (14), 2055-2070.

(191) Cheng, X.; Vellalath, S.; Goddard, R.; List, B. Direct Catalytic Asymmetric Synthesis of Cyclic Aminals from Aldehydes. *Journal of the American Chemical Society* **2008,** *130* (47), 15786-15787.

(192) Stokes, B. J.; Richert, K. J.; Driver, T. G. Examination of the Mechanism of Rh_2(II)-Catalyzed Carbazole Formation Using Intramolecular Competition Experiments. *The Journal of Organic Chemistry* **2009,** *74* (17), 6442-6451.

(193) Latham, J.; Henry, J.-M.; Sharif, H. H.; Menon, B. R. K.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J. Integrated Catalysis Opens New Arylation Pathways Via Regiodivergent Enzymatic C–H Activation. *Nature Communications* **2016,** *7*, 11873.

(194) Willwacher, J.; Raj, R.; Mohammed, S.; Davis, B. G. Selective Metal-Site-Guided Arylation of Proteins. *Journal of the American Chemical Society* **2016,** *138* (28), 8678-8681.

(195) Hall, D. G. In *Boronic Acids*; Wiley-VCH Verlag GmbH & Co. KGaA, 2006.

(196) James, T. D. In *Boronic Acids*; Wiley-VCH Verlag GmbH & Co. KGaA, 2006.

(197) Frese, M.; Sewald, N. Enzymatic Halogenation of Tryptophan on a Gram Scale. *Angewandte Chemie International Edition* **2015,** *54* (1), 298-301.

(198) Cao, L.; van Rantwijk, F.; Sheldon, R. A. Cross-Linked Enzyme Aggregates: A Simple and Effective Method for the Immobilization of Penicillin Acylase. *Organic Letters* **2000,** *2* (10), 1361-1364.

(199) Sheldon, R. A. Enzyme Immobilization: The Quest for Optimum Performance. *Advanced Synthesis & Catalysis* **2007,** *349* (8-9), 1289-1307.

(200) Lee, J. N.; Park, C.; Whitesides, G. M. Solvent Compatibility of Poly(Dimethylsiloxane)-Based Microfluidic Devices. *Analytical Chemistry* **2003,** *75* (23), 6544-6554.

(201) McDonald, J. C.; Whitesides, G. M. Poly(Dimethylsiloxane) as a Material for Fabricating Microfluidic Devices. *Accounts of Chemical Research* **2002,** *35* (7), 491-499.

(202) Bedford, R. B.; Hazelwood, S. L.; Albisson, D. A. Platinum Catalysts for Suzuki Biaryl Coupling Reactions. *Organometallics* **2002,** *21* (13), 2599-2600.

(203) Francis, D.; Winn, M.; Latham, J.; Greaney, M. F.; Micklefield, J. An Engineered Tryptophan Synthase Opens New Enzymatic Pathways to B-Methyltryptophan and Derivatives. *ChemBioChem* **2017,** *18* (4), 382-386.

(204) Dunn, M. F.; Niks, D.; Ngo, H.; Barends, T. R. M.; Schlichting, I. Tryptophan Synthase: The Workings of a Channeling Nanomachine. *Trends in Biochemical Sciences* **2008,** *33* (6), 254-264.

(205) Phillips, R. S. Synthetic Applications of Tryptophan Synthase. *Tetrahedron: Asymmetry* **2004,** *15* (18), 2787-2792.

(206) Goss, R. J. M.; Newill, P. L. A. A Convenient Enzymatic Synthesis of L-Halotryptophans. *Chemical Communications* **2006,** *47* (47), 4924-4925.

(207) Frese, M.; Schnepel, C.; Minges, H.; Voß, H.; Feiner, R.; Sewald, N. Modular Combination of Enzymatic Halogenation of Tryptophan with Suzuki–Miyaura Cross-Coupling Reactions. *ChemCatChem* **2016,** *8* (10), 1799-1803.

(208) Roff, G. J.; Lloyd, R. C.; Turner, N. J. A Versatile Chemo-Enzymatic Route to Enantiomerically Pure β-Branched α-Amino Acids. *Journal of the American Chemical Society* **2004,** *126* (13), 4098-4099.

(209) Parmeggiani, F.; Lovelock, S. L.; Weise, N. J.; Ahmed, S. T.; Turner, N. J. Synthesis of D- and L-Phenylalanine Derivatives by Phenylalanine Ammonia Lyases: A Multienzymatic Cascade Process. *Angewandte Chemie International Edition* **2015,** *54* (15), 4608-4611.

(210) Castanet, A.-S.; Colobert, F.; Schlama, T. Suzuki−Miyaura Coupling of Alkynylboronic Esters Generated in Situ from Acetylenic Derivatives. *Organic Letters* **2000,** *2* (23), 3559-3561.

(211) Durak, L. J.; Payne, J. T.; Lewis, J. C. Late-Stage Diversification of Biologically Active Molecules Via Chemoenzymatic C–H Functionalization. *ACS Catalysis* **2016,** *6* (3), 1451-1454.

(212) Yang, Y.; Li, R.; Zhao, Y.; Zhao, D.; Shi, Z. Cu-Catalyzed Direct C6-Arylation of Indoles. *Journal of the American Chemical Society* **2016,** *138* (28), 8734-8737.

(213) Yang, Y.; Qiu, X.; Zhao, Y.; Mu, Y.; Shi, Z. Palladium-Catalyzed C–H Arylation of Indoles at the C7 Position. *Journal of the American Chemical Society* **2016,** *138* (2), 495-498.

(214) Anbarasan, P.; Schareina, T.; Beller, M. Recent Developments and Perspectives in Palladium-Catalyzed Cyanation of Aryl Halides: Synthesis of Benzonitriles. *Chemical Society Reviews* **2011,** *40* (10), 5049-5067.

(215) Senecal, T. D.; Shu, W.; Buchwald, S. L. A General, Practical Palladium-Catalyzed Cyanation of (Hetero)Aryl Chlorides and Bromides. *Angewandte Chemie International Edition* **2013,** *52* (38), 10035-10039.

(216) Jones, L. H.; Summerhill, N. W.; Swain, N. A.; Mills, J. E. Aromatic Chloride to Nitrile Transformation: Medicinal and Synthetic Chemistry. *MedChemComm* **2010,** *1* (5), 309-318.

(217) Feng, J.; Zhang, Z.; Wallace, M. B.; Stafford, J. A.; Kaldor, S. W.; Kassel, D. B.; Navre, M.; Shi, L.; Skene, R. J.; Asakawa, T.et al. Discovery of Alogliptin: A Potent, Selective, Bioavailable, and Efficacious Inhibitor of Dipeptidyl Peptidase Iv. *Journal of Medicinal Chemistry* **2007,** *50* (10), 2297-2300.

(218) Okamoto, K.; Matsumoto, K.; Hille, R.; Eger, B. T.; Pai, E. F.; Nishino, T. The Crystal Structure of Xanthine Oxidoreductase During Catalysis: Implications for Reaction Mechanism and Enzyme Inhibition. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (21), 7931-7936.

(219) Zanon, J.; Klapars, A.; Buchwald, S. L. Copper-Catalyzed Domino Halide Exchange-Cyanation of Aryl Bromides. *Journal of the American Chemical Society* **2003,** *125* (10), 2890-2891.

(220) Monnier, F.; Taillefer, M. Catalytic C-C, C-N, and C-O Ullmann-Type Coupling Reactions. *Angewandte Chemie International Edition* **2009,** *48* (38), 6954-6971.

(221) Dobbs, K. D.; Marshall, W. J.; Grushin, V. V. Why Excess Cyanide Can Be Detrimental to Pd-Catalyzed Cyanation of Haloarenes. Facile Formation and Characterization of $[Pd(CN)_3(H)]^{2-}$ and $[Pd(CN)_3(Ph)]_2$. *Journal of the American Chemical Society* **2007,** *129* (1), 30-31.

(222) Erhardt, S.; Grushin, V. V.; Kilpatrick, A. H.; Macgregor, S. A.; Marshall, W. J.; Roe, D. C. Mechanisms of Catalyst Poisoning in Palladium-Catalyzed Cyanation of Haloarenes.

Remarkably Facile C−N Bond Activation in the [$(Ph_3P)_4Pd$]/[Bu_4n]$^+$ CN^- System. *Journal of the American Chemical Society* **2008,** *130* (14), 4828-4845.

(223) Kentaro, T.; Tadashi, O.; Yasumasa, S.; Shinzaburo, O. Palladium(II) Catalyzed Synthesis of Aryl Cyanaides from Aryl Halides. *Chemistry Letters* **1973,** *2* (5), 471-474.

(224) Sundermeier, M.; Zapf, A.; Mutyala, S.; Baumann, W.; Sans, J.; Weiss, S.; Beller, M. Progress in the Palladium-Catalyzed Cyanation of Aryl Chlorides. *Chemistry – A European Journal* **2003,** *9* (8), 1828-1836.

(225) Yang, C.; Williams, J. M. Palladium-Catalyzed Cyanation of Aryl Bromides Promoted by Low-Level Organotin Compounds. *Organic Letters* **2004,** *6* (17), 2837-2840.

(226) Tschaen, D. M.; Desmond, R.; King, A. O.; Fortin, M. C.; Pipik, B.; King, S.; Verhoeven, T. R. An Improved Procedure for Aromatic Cyanation. *Synthetic Communications* **1994,** *24* (6), 887-890.

(227) Zhang, D.; Sun, H.; Zhang, L.; Zhou, Y.; Li, C.; Jiang, H.; Chen, K.; Liu, H. An Expedient Pd/DBU Mediated Cyanation of Aryl/Heteroaryl Bromides with $K_4[Fe(CN)_6]$. *Chemical Communications* **2012,** *48* (23), 2909-2911.

(228) Cohen, D. T.; Buchwald, S. L. Mild Palladium-Catalyzed Cyanation of (Hetero)Aryl Halides and Triflates in Aqueous Media. *Org Lett* **2015,** *17* (2), 202-205.

(229) Jin, J.; Wen, Q.; Lu, P.; Wang, Y. Copper-Catalyzed Cyanation of Arenes Using Benzyl Nitrile as a Cyanide Anion Surrogate. *Chemical Communications* **2012,** *48* (79), 9933-9935.

(230) Zhang, W.; Wang, F.; McCann, S. D.; Wang, D.; Chen, P.; Stahl, S. S.; Liu, G. Enantioselective Cyanation of Benzylic C–H Bonds Via Copper-Catalyzed Radical Relay. *Science* **2016,** *353* (6303), 1014-1018.

(231) Yan, G.; Kuang, C.; Zhang, Y.; Wang, J. Palladium-Catalyzed Direct Cyanation of Indoles with $K_4[Fe(CN)_6]$. *Organic Letters* **2010,** *12* (5), 1052-1055.

(232) Subba Reddy, B. V.; Begum, Z.; Jayasudhan Reddy, Y.; Yadav, J. S. $Pd(OAc)_2$-Catalyzed C–H Activation of Indoles: A Facile Synthesis of 3-Cyanoindoles. *Tetrahedron Letters* **2010,** *51* (25), 3334-3336.

(233) Liu, W.; Richter, S. C.; Mei, R.; Feldt, M.; Ackermann, L. Synergistic Heterobimetallic Manifold for Expedient Manganese(I)-Catalyzed C−H Cyanation. *Chemistry - A European Journal* **2016,** *22* (50), 17958-17961.

(234) Peng, J.; Zhao, J.; Hu, Z.; Liang, D.; Huang, J.; Zhu, Q. Palladium-Catalyzed C(sp_2)–H Cyanation Using Tertiary Amine Derived Isocyanide as a Cyano Source. *Organic Letters* **2012,** *14* (18), 4966-4969.

(235) David, N.; Pasceri, R.; Kitson, R. R. A.; Pradal, A.; Moody, C. J. Formal Total Synthesis of Diazonamide a by Indole Oxidative Rearrangement. *Chemistry – A European Journal* **2016,** *22* (31), 10867-10876.

(236) Adachi, H.; Palaniappan, K. K.; Ivanov, A. A.; Bergman, N.; Gao, Z.-G.; Jacobson, K. A. Structure−Activity Relationships of 2,N6,5'-Substituted Adenosine Derivatives with Potent Activity at the A2b Adenosine Receptor. *Journal of Medicinal Chemistry* **2007,** *50* (8), 1810-1827.

(237) Pereira, R.; Benedetti, R.; Pérez-Rodríguez, S.; Nebbioso, A.; García-Rodríguez, J.; Carafa, V.; Stuhldreier, M.; Conte, M.; Rodríguez-Barrios, F.; Stunnenberg, H. G.et al. Indole-Derived Psammaplin a Analogues as Epigenetic Modulators with Multiple Inhibitory Activities. *Journal of Medicinal Chemistry* **2012,** *55* (22), 9467-9491.

(238) Kumar, V.; Kaur, S.; Kumar, S. Zrcl4 Catalyzed Highly Selective and Efficient Michael Addition of Heterocyclic Enamines with α,β-Unsaturated Olefins. *Tetrahedron Letters* **2006,** *47* (39), 7001-7005.

(239) Prasad, S.; Bhalla, T. C. Nitrile Hydratases (NHases): At the Interface of Academia and Industry. *Biotechnol Adv* **2010,** *28* (6), 725-741.

(240) Chen, J.; Zheng, R. C.; Zheng, Y. G.; Shen, Y. C. Microbial Transformation of Nitriles to High-Value Acids or Amides. *Adv Biochem Eng Biotechnol* **2009,** *113*, 33-77.

(241) Peng, J.; Liu, L.; Hu, Z.; Huang, J.; Zhu, Q. Direct Carboxamidation of Indoles by Palladium-Catalyzed C-H Activation and Isocyanide Insertion. *Chemical Communications* **2012,** *48* (31), 3772-3774.

(242) Xu, J.; Sharma, N.; Sharma, U. K.; Li, Z.; Song, G.; Van der Eycken, E. V. Cationic Rhodium(III)-Catalyzed Direct C-2 Carboxamidation of Indoles with Isocyanates Via C-H Bond Functionalization. *Advanced Synthesis & Catalysis* **2015,** *357* (12), 2615-2621.

(243) Qiu, G.; Ding, Q.; Wu, J. Recent Advances in Isocyanide Insertion Chemistry. *Chemical Society Reviews* **2013,** *42* (12), 5257-5269.

(244) Rzeznicka, K.; Schatzle, S.; Bottcher, D.; Klein, J.; Bornscheuer, U. T. Cloning and Functional Expression of a Nitrile Hydratase (NHase) from Rhodococcus Equi TG328-2 in Escherichia Coli, Its Purification and Biochemical Characterisation. *Appl Microbiol Biotechnol* **2010,** *85* (5), 1417-1425.

(245) Song, L.; Wang, M.; Yang, X.; Qian, S. Purification and Characterization of the Enantioselective Nitrile Hydratase from Rhodococcus Sp. AJ270. *Biotechnol J* **2007,** *2* (6), 717-724.

(246) Constable, D. J. C.; Dunn, P. J.; Hayler, J. D.; Humphrey, G. R.; Leazer, J. J. L.; Linderman, R. J.; Lorenz, K.; Manley, J.; Pearlman, B. A.; Wells, A.et al. Key Green Chemistry Research Areas-a Perspective from Pharmaceutical Manufacturers. *Green Chemistry* **2007,** *9* (5), 411-420.

(247) Bandichhor, R.; Bhattacharya, A.; Cosbie, A.; Diorazio, L.; Dunn, P.; Fraunhoffer, K.; Gallou, F.; Hayler, J.; Hickey, M.; Hinkley, B.et al. Green Chemistry Articles of Interest to the Pharmaceutical Industry. *Organic Process Research & Development* **2015,** *19* (12), 1924-1935.

(248) Pattabiraman, V. R.; Bode, J. W. Rethinking Amide Bond Synthesis. *Nature* **2011,** *480* (7378), 471-479.

(249) Dugger, R. W.; Ragan, J. A.; Ripin, D. H. B. Survey of Gmp Bulk Reactions Run in a Research Facility between 1985 and 2002. *Organic Process Research & Development* **2005,** *9* (3), 253-258.

(250) Carey, J. S.; Laffan, D.; Thomson, C.; Williams, M. T. Analysis of the Reactions Used for the Preparation of Drug Candidate Molecules. *Organic & Biomolecular Chemistry* **2006,** *4* (12), 2337-2347.

(251) Dunn, P. J. Synthesis of Commercial Phosphodiesterase(V) Inhibitors. *Organic Process Research & Development* **2005,** *9* (1), 88-97.

(252) Sindhuja, E.; Ramesh, R.; Balaji, S.; Liu, Y. Direct Synthesis of Amides from Coupling of Alcohols and Amines Catalyzed by Ruthenium(II) Thiocarboxamide Complexes under Aerobic Conditions. *Organometallics* **2014,** *33* (16), 4269-4278.

(253) Al-Zoubi, R. M.; Marion, O.; Hall, D. G. Direct and Waste-Free Amidations and Cycloadditions by Organocatalytic Activation of Carboxylic Acids at Room Temperature. *Angewandte Chemie International Edition* **2008,** *47* (15), 2876-2879.

(254) Goswami, A.; Van Lanen, S. G. Enzymatic Strategies and Biocatalysts for Amide Bond Formation: Tricks of the Trade Outside of the Ribosome. *Mol Biosyst* **2015,** *11* (2), 338-353.

(255) Reetz, M. T.; Dreisbach, C. Highly Efficient Lipase-Catalyzed Kinetic Resolution of Chiral Amines. *CHIMIA International Journal for Chemistry* **1994,** *48* (12), 570-570.

(256) Palo-Nieto, C.; Afewerki, S.; Anderson, M.; Tai, C.-W.; Berglund, P.; Córdova, A. Integrated Heterogeneous Metal/Enzymatic Multiple Relay Catalysis for Eco-Friendly and Asymmetric Synthesis. *ACS Catalysis* **2016,** *6* (6), 3932-3940.

(257) Nagasawa, T.; Takeuchi, K.; Yamada, H. Characterization of a New Cobalt-Containing Nitrile Hydratase Purified from Urea-Induced Cells of Rhodococcus Rhodochrous J1. *European Journal of Biochemistry* **1991,** *196* (3), 581-589.

(258) Nagasawa, T.; Takeuchi, K.; Nardi-Dei, V.; Mihara, Y.; Yamada, H. Optimum Culture Conditions for the Production of Cobalt-Containing Nitrile Hydratase by Rhodococcus Rhodochrous J1. *Appl Microbiol Biotechnol* **1991,** *34* (6), 783-788.

(259) Komeda, H.; Kobayashi, M.; Shimizu, S. Characterization of the Gene Cluster of High-Molecular-Mass Nitrile Hydratase (H-NHase) Induced by Its Reaction Product in Rhodococcus Rhodochrous J1. *Proceedings of the National Academy of Sciences* **1996,** *93* (9), 4267-4272.

(260) Nagashima, S.; Nakasako, M.; Dohmae, N.; Tsujimura, M.; Takio, K.; Odaka, M.; Yohda, M.; Kamiya, N.; Endo, I. Novel Non-Heme Iron Center of Nitrile Hydratase with a Claw Setting of Oxygen Atoms. *Nature structural biology* **1998,** *5* (5), 347-351.

(261) Miyanaga, A.; Fushinobu, S.; Ito, K.; Wakagi, T. Crystal Structure of Cobalt-Containing Nitrile Hydratase. *Biochem Biophys Res Commun* **2001,** *288* (5), 1169-1174.

(262) Payne, M. S.; Wu, S.; Fallon, R. D.; Tudor, G.; Stieglitz, B.; Turner, I. M.; Nelson, M. J. A Stereoselective Cobalt-Containing Nitrile Hydratase. *Biochemistry* **1997,** *36* (18), 5447-5454.

(263) Martinez, S.; Wu, R.; Sanishvili, R.; Liu, D.; Holz, R. The Active Site Sulfenic Acid Ligand in Nitrile Hydratases Can Function as a Nucleophile. *J Am Chem Soc* **2014,** *136* (4), 1186-1189.

(264) Yamanaka, Y.; Kato, Y.; Hashimoto, K.; Iida, K.; Nagasawa, K.; Nakayama, H.; Dohmae, N.; Noguchi, K.; Noguchi, T.; Yohda, M.et al. Time-Resolved Crystallography of the Reaction Intermediate of Nitrile Hydratase: Revealing a Role for the Cysteinesulfenic Acid Ligand as a Catalytic Nucleophile. *Angewandte Chemie International Edition* **2015,** *54* (37), 10763-10767.

(265) van Pelt, S.; Zhang, M.; Otten, L. G.; Holt, J.; Sorokin, D. Y.; van Rantwijk, F.; Black, G. W.; Perry, J. J.; Sheldon, R. A. Probing the Enantioselectivity of a Diverse Group of Purified Cobalt-Centred Nitrile Hydratases. *Org Biomol Chem* **2011,** *9* (8), 3011-3019.

(266) Yasukawa, K.; Hasemi, R.; Asano, Y. Dynamic Kinetic Resolution of A-Aminonitriles to Form Chiral A-Amino Acids. *Advanced Synthesis & Catalysis* **2011,** *353* (13), 2328-2332.

(267) Wang, M. X. Enantioselective Biotransformations of Nitriles in Organic Synthesis. *Acc Chem Res* **2015,** *48* (3), 602-611.

(268) Wang, M.-X.; Lu, G.; Ji, G.-J.; Huang, Z.-T.; Meth-Cohn, O.; Colby, J. Enantioselective Biotransformations of Racemic A-Substituted Phenylacetonitriles and Phenylacetamides Using Rhodococcus Sp. AJ270. *Tetrahedron: Asymmetry* **2000,** *11* (5), 1123-1135.

(269) Wang, M.-X.; Deng, G.; Wang, D.-X.; Zheng, Q.-Y. Nitrile Biotransformations for Highly Enantioselective Synthesis of Oxiranecarboxamides with Tertiary and Quaternary

Stereocenters; Efficient Chemoenzymatic Approaches to Enantiopure A-Methylated Serine and Isoserine Derivatives. *The Journal of Organic Chemistry* **2005,** *70* (7), 2439-2444.

(270) Wang, M.-X.; Feng, G.-Q. Enantioselective Synthesis of Chiral Cyclopropane Compounds through Microbial Transformations of Trans-2-Arylcyclopropanecarbonitriles. *Tetrahedron Letters* **2000,** *41* (33), 6501-6505.

(271) Kuranaga, T.; Sesoko, Y.; Inoue, M. Cu-Mediated Enamide Formation in the Total Synthesis of Complex Peptide Natural Products. *Natural Product Reports* **2014,** *31* (4), 514-532.

(272) Strieter, E. R.; Blackmond, D. G.; Buchwald, S. L. The Role of Chelating Diamine Ligands in the Goldberg Reaction: A Kinetic Study on the Copper-Catalyzed Amidation of Aryl Iodides. *Journal of the American Chemical Society* **2005,** *127* (12), 4120-4121.

(273) Chan, D. M. T.; Monaco, K. L.; Wang, R.-P.; Winters, M. P. New N- and O-Arylations with Phenylboronic Acids and Cupric Acetate. *Tetrahedron Letters* **1998,** *39* (19), 2933-2936.

(274) Lam, P. Y. S.; Clark, C. G.; Saubern, S.; Adams, J.; Winters, M. P.; Chan, D. M. T.; Combs, A. New Aryl/Heteroaryl C≡N Bond Cross-Coupling Reactions Via Arylboronic Acid/Cupric Acetate Arylation. *Tetrahedron Letters* **1998,** *39* (19), 2941-2944.

(275) Evans, D. A.; Katz, J. L.; West, T. R. Synthesis of Diaryl Ethers through the Copper-Promoted Arylation of Phenols with Arylboronic Acids. An Expedient Synthesis of Thyroxine. *Tetrahedron Letters* **1998,** *39* (19), 2937-2940.

(276) Qiao, J.; Lam, P. Copper-Promoted Carbon-Heteroatom Bond Cross-Coupling with Boronic Acids and Derivatives. *Synthesis* **2010,** *2011* (06), 829-856.

(277) Bolshan, Y.; Batey, R. A. Copper-Catalyzed Cross-Coupling of Amides and Potassium Alkenyltrifluoroborate Salts: A General Approach to the Synthesis of Enamides. *Tetrahedron* **2010,** *66* (27–28), 5283-5294.

(278) Bolshan, Y.; Batey, R. A. Enamide Synthesis by Copper-Catalyzed Cross-Coupling of Amides and Potassium Alkenyltrifluoroborate Salts. *Angew Chem Int Ed Engl* **2008,** *47* (11), 2109-2112.

(279) Zheng, Z. G.; Wen, J.; Wang, N.; Wu, B.; Yu, X. Q. N-Arylation of Amines, Amides, Imides and Sulfonamides with Arylboroxines Catalyzed by Simple Copper Salt/EtOH System. *Beilstein J Org Chem* **2008,** *4*, 40.

(280) Caine, D. Potassium Tert-Butoxide–18-Crown-6. *Encyclopedia of Reagents for Organic Synthesis* **2001**.

(281) Caine, D. Potassium t-Butoxide–Dimethyl Sulfoxide. *Encyclopedia of Reagents for Organic Synthesis* **2001**.

(282) Caine, D. Potassium Tert-Butoxide. *Encyclopedia of Reagents for Organic Synthesis* **2001**.

(283) Deng, W.; Wang, Y.-F.; Zou, Y.; Liu, L.; Guo, Q.-X. Amino Acid-Mediated Goldberg Reactions between Amides and Aryl Iodides. *Tetrahedron Letters* **2004,** *45* (11), 2311-2315.

(284) Quan, Z.-J.; Xia, H.-D.; Zhang, Z.; Da, Y.-X.; Wang, X.-C. Ligand-Free Cutc-Catalyzedn-Arylation of Amides, Anilines and 4-Aminoantipyrine: Synthesis Ofn-Arylacrylamides, 4-Amido-N-Phenylbenzamides and 4-Amino(N-Phenyl)Antipyrenes. *Applied Organometallic Chemistry* **2014,** *28* (2), 81-85.

(285) Shen, R.; Lin, C. T.; Bowman, E. J.; Bowman, B. J.; Porco, J. A. Lobatamide C: Total Synthesis, Stereochemical Assignment, Preparation of Simplified Analogues, and V-ATPase Inhibition Studies. *Journal of the American Chemical Society* **2003,** *125* (26), 7889-7901.

(286) Liu, X.; Mao, D.; Wu, S.; Yu, J.; Hong, G.; Zhao, Q.; Wang, L. Copper-Catalyzed Direct Oxidation and N-Arylation of Benzylamines with Diaryliodonium Salts. *Science China Chemistry* **2014,** *57* (8), 1132-1136.

(287) Cahard, E.; Bremeyer, N.; Gaunt, M. J. Copper-Catalyzed Intramolecular Electrophilic Carbofunctionalization of Allylic Amides. *Angewandte Chemie International Edition* **2013,** *52* (35), 9284-9288.

(288) Suarez, L. L.; Greaney, M. F. Tandem Indole C-H Alkenylation/Arylation for Tetra-Substituted Alkene Synthesis. *Chemical Communications* **2011,** *47* (28), 7992-7994.

(289) Panda, N.; Jena, A. K.; Raghavender, M. Stereoselective Synthesis of Enamides by Palladium Catalyzed Coupling of Amides with Electron Deficient Olefins. *ACS Catalysis* **2012,** *2* (4), 539-543.

(290) Panda, N.; Mothkuri, R. Stereoselective Synthesis of Enamides by Pd-Catalyzed Hydroamidation of Electron Deficient Terminal Alkynes. *The Journal of Organic Chemistry* **2012,** *77* (20), 9407-9412.

(291) Qiao, J. X.; Lam, P. Y. S. In *Boronic Acids*; Wiley-VCH Verlag GmbH & Co. KGaA, 2011.

(292) Liu, R.; Zhang, P.; Gan, T.; Cook, J. M. Regiospecific Bromination of 3-Methylindoles with Nbs and Its Application to the Concise Synthesis of Optically Active Unusual Tryptophans Present in Marine Cyclic Peptides1. *The Journal of Organic Chemistry* **1997,** *62* (21), 7447-7456.

(293) Vaughan, J.; Christopher, J. O-Hydroxy and O-Amino Benzamide Derivatives as IKK2 Inhibitors. **2007**, WO2007025575 A2007025571.

(294) Beugelmans, R.; Roussi, G.; González Zamora, E.; Carbonnelle, A.-C. Synthetic Studies Towards Western and Eastern Macropolypeptide Subunits of Kistamycin. *Tetrahedron* **1999,** *55* (16), 5089-5112.

(295) Plettenburg, O.; Hofmeister, A.; Kadereit, D.; Peukert, S.; Ruf, S.; Ritter, K.; Löhn, M.; Ivashchenko Monecke, P.; Dreyer, M.; Kannt, A. (Wo2007000240) Isoquinoline Derrivatives Ad Inhibitors of Rho-Kinase. **2007**.

(296) Xu, B.; Jiang, Q.; Zhao, A.; Jia, J.; Liu, Q.; Luo, W.; Guo, C. Copper-Catalyzed Aerobic Conversion of the C=O Bond of Ketones to a C≡N Bond Using Ammonium Salts as the Nitrogen Source. *Chemical Communications* **2015,** *51* (56), 11264-11267.

(297) Fu, R.; Yang, Y.; Feng, W.; Ge, Q.; Feng, Y.; Zeng, X.; Chai, W.; Yi, J.; Yuan, R. An Efficient, Eco-Friendly and Sustainable Tandem Oxidative Amidation of Alcohols with Amines Catalyzed by Heteropolyanion-Based Ionic Liquids Via a Bifunctional Catalysis Process. *Tetrahedron* **2016,** *72* (50), 8319-8326.

(298) Song, Q.; Feng, Q.; Yang, K. Synthesis of Primary Amides Via Copper-Catalyzed Aerobic Decarboxylative Ammoxidation of Phenylacetic Acids and α-Hydroxyphenylacetic Acids with Ammonia in Water. *Organic Letters* **2014,** *16* (2), 624-627.

(299) Kumar, S.; Sharma, S.; Das, P. Supported Gold Nanoparticles-Catalyzed Microwave-Assisted Hydration of Nitriles to Amides under Base-Free Conditions. *Advanced Synthesis & Catalysis* **2016,** *358* (18), 2889-2894.

(300) Xin, X.; Xiang, D.; Yang, J.; Zhang, Q.; Zhou, F.; Dong, D. Homogeneous and Stereoselective Copper(II)-Catalyzed Monohydration of Methylenemalononitriles to 2-Cyanoacrylamides. *The Journal of Organic Chemistry* **2013,** *78* (23), 11956-11961.

(301) Shimokawa, S.; Kawagoe, Y.; Moriyama, K.; Togo, H. Direct Transformation of Ethylarenes into Primary Aromatic Amides with N-Bromosuccinimide and I_2–Aqueous NH_3. *Organic Letters* **2016,** *18* (4), 784-787.

(302) Marce, P.; Lynch, J.; Blacker, A. J.; Williams, J. M. A Mild Hydration of Nitriles Catalysed by Copper(II) Acetate. *Chemical Communications* **2016,** *52* (7), 1436-1438.

(303) Ambreen, N.; Wirth, T. High-Temperature Synthesis of Amides from Alcohols or Aldehydes by Using Flow Chemistry. *European Journal of Organic Chemistry* **2014,** *2014* (34), 7590-7593.

(304) Lee, W.-C.; Frost, B. J. Aqueous and Biphasic Nitrile Hydration Catalyzed by a Recyclable Ru(Ii) Complex under Atmospheric Conditions. *Green Chem.* **2012,** *14* (1), 62-66.

(305) Salonen, L. M.; Holland, M. C.; Kaib, P. S. J.; Haap, W.; Benz, J.; Mary, J.-L.; Kuster, O.; Schweizer, W. B.; Banner, D. W.; Diederich, F. Molecular Recognition at the Active Site of Factor Xa: Cation–Π Interactions, Stacking on Planar Peptide Surfaces, and Replacement of Structural Water. *Chemistry – A European Journal* **2012,** *18* (1), 213-222.

(306) Kracke, G. R.; VanGordon, M. R.; Sevryugina, Y. V.; Kueffer, P. J.; Kabytaev, K.; Jalisatgi, S. S.; Hawthorne, M. F. Carborane-Derived Local Anesthetics Are Isomer Dependent. *ChemMedChem* **2015,** *10* (1), 62-67.

(307) Li, Y.-T.; Liao, B.-S.; Chen, H.-P.; Liu, S.-T. Ligand-Free Nickel-Catalyzed Conversion of Aldoximes into Nitriles. *Synthesis* **2011,** *2011* (16), 2639-2643.

(308) Feng, C.; Loh, T.-P. Rhodium(III)-Catalyzed Cross-Coupling of Alkenylboronic Acids and N-Pivaloyloxylamides. *Organic Letters* **2014,** *16* (13), 3444-3447.

(309) Baker, W. R.; Stasiak, M.; Macleod, D.; WO/2005/025498: US, 2005.

(310) Srinivas, M.; Hudwekar, A. D.; Venkateswarlu, V.; Reddy, G. L.; Kumar, K. A. A.; Vishwakarma, R. A.; Sawant, S. D. A Metal-Free Approach for Transamidation of Amides with Amines in Aqueous Media. *Tetrahedron Letters* **2015,** *56* (33), 4775-4779.

(311) Lim, D. S. W.; Lew, T. T. S.; Zhang, Y. Direct Amidation of N-Boc- and N-Cbz-Protected Amines Via Rhodium-Catalyzed Coupling of Arylboroxines and Carbamates. *Organic Letters* **2015,** *17* (24), 6054-6057.

(312) Bielawski, M.; Aili, D.; Olofsson, B. Regiospecific One-Pot Synthesis of Diaryliodonium Tetrafluoroborates from Arylboronic Acids and Aryl Iodides. *The Journal of Organic Chemistry* **2008,** *73* (12), 4602-4607.

(313) Xu, Y.; Young, M. C.; Wang, C.; Magness, D. M.; Dong, G. Catalytic C(sp_3)−H Arylation of Free Primary Amines with an Exo Directing Group Generated in Situ. *Angewandte Chemie International Edition* **2016,** *55* (31), 9084-9087.

nature COMMUNICATIONS

ARTICLE

Received 11 Jan 2016 | Accepted 9 May 2016 | Published 10 Jun 2016

DOI: 10.1038/ncomms11873 OPEN

Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation

Jonathan Latham[1,2], Jean-Marc Henry[1,2], Humera H. Sharif[1,2], Binuraj R.K. Menon[1,2], Sarah A. Shepherd[1,2], Michael F. Greaney[1] & Jason Micklefield[1,2]

Despite major recent advances in C–H activation, discrimination between two similar, unactivated C–H positions is beyond the scope of current chemocatalytic methods. Here we demonstrate that integration of regioselective halogenase enzymes with Pd-catalysed cross-coupling chemistry, in one-pot reactions, successfully addresses this problem for the indole heterocycle. The resultant 'chemobio-transformation' delivers a range of functionally diverse arylated products that are impossible to access using separate enzymatic or chemocatalytic C–H activation, under mild, aqueous conditions. This use of different biocatalysts to select different C–H positions contrasts with the prevailing substrate-control approach to the area, and presents opportunities for new pathways in C–H activation chemistry. The issues of enzyme and transition metal compatibility are overcome through membrane compartmentalization, with the optimized process requiring no intermediate work-up or purification steps.

[1] School of Chemistry, The University of Manchester, Oxford Road, Manchester M13 9PL, UK. [2] Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK. Correspondence and requests for materials should be addressed to M.F.G. (email: michael.greaney@manchester.ac.uk) or to J.M. (email: jason.micklefield@manchester.ac.uk).

Selective activation of C–H positions has transformed synthesis in recent years. Using transition metal catalysis, a plethora of direct, fundamental bond formations are now achievable under mild, catalytic conditions, changing the way chemists make molecules[1]. The field is currently defined by substrate control, with a relatively small number of C–H motifs undergoing predictable and reliable metalation by a transition metal catalyst; for example, electron-rich heteroarenes, electron-poor fluoroarenes, activation ortho (and more recently, meta), to a directing group via cyclometalation all being dominant examples (Fig. 1a)[2–9]. Effective catalyst control, where choice of catalyst can discriminate between neighbouring C–H bonds in the absence of strong substrate-directing effects, is very rare and remains an outstanding challenge in the field[10–13]. Progress in this area will profoundly enhance the scope of C–H activation, as the majority of C–H bonds do not fit the current substrate-control criteria for chemocatalytic activation chemistry.

We were interested in addressing this problem using a combination of bio- and chemo-catalysis. The complex coordination sphere of an enzyme active site enables exquisite selectivities that cannot be achieved by simple metal coordination complexes. Correspondingly, the manifold power of noble metal catalysis in C–C bond formation presents synthesis pathways that are not possible through biocatalysis. Their integration offers a complementarity that could create new C–H activation strategies, whereby C–H positions beyond the scope of chemocatalysis alone can be targeted for C–C, C–N, C–Hal and other fundamental bond-forming steps. In addition, the merging of bio- and chemo-catalysis into single process, one-pot operations creates a step-change in efficiency through enhanced space-time yields, better energy consumption and elimination of auxiliary chemicals[14,15].

Despite the potential advantages of combining biocatalysts with chemocatalysts, their successful integration has been largely restricted to organic solvents[15–17], which are incompatible with many enzymes and cofactors—therefore limiting the number of biocatalysts that can be utilized. In contrast, integration of chemocatalysts with enzymes in their preferred aqueous media is relatively rare[14]. Indeed, the inherent differences in operating conditions of enzymes and chemocatalysts present serious challenges, often requiring compartmentalization or immobilization of at least one of the catalysts[18–23].

Our plans for a catalyst-controlled, integrated C–H activation system involve the merging of flavin-dependent halogenase (Fl-Hal) enzymes with palladium-catalysed cross-coupling. The pivotal role played by aryl halide functionality in metal-catalysed transformations suggested to us that a regioselective, biohalogenation would have great versatility, with the incipient Ar-X group being adaptable to a multitude of Pd-catalysed chemistries. We chose to integrate Fl-Hals and Suzuki–Miyaura coupling (SMC) with boronic acids in the first instance, given the critical role of SMC in producing biaryl products for pharmaceutical, agrochemical and materials science applications.

The most widely studied Fl-Hals are tryptophan halogenases, which are responsible for chlorinating different positions of the indole ring of tryptophan *en route* to various bacterial natural

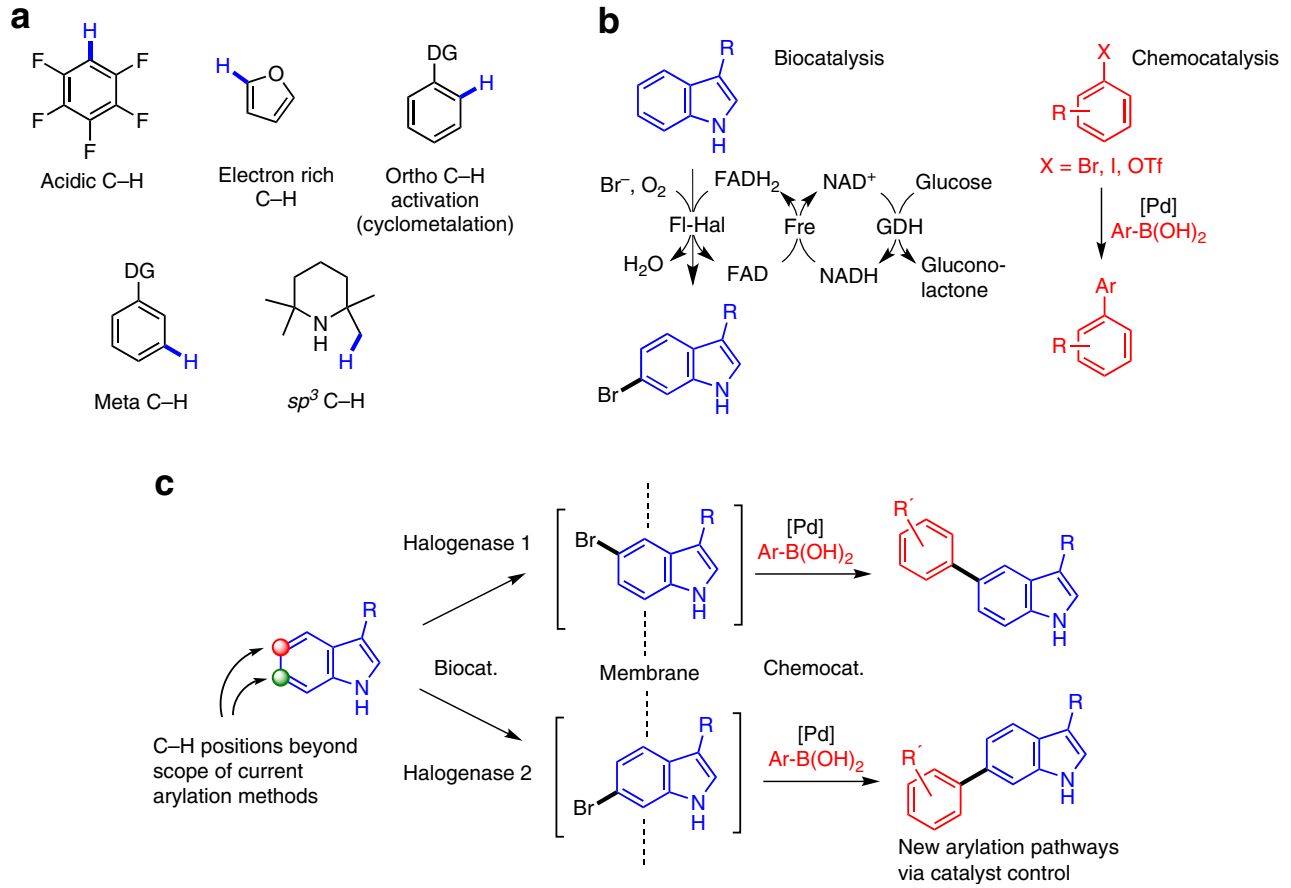


Figure 1 | Previous work on C–H activation and proposed scheme of catalyst control using biohalogenation and Pd-catalysed cross-couplings. (**a**) Previous work on C–H activation utilizing substrate control of regioselectivity. (**b**) Catalytic cycle of halogenase biocatalysis using a Fl-Hal, along with flavin reductase (Fre) and glucose dehydrogenase (GDH) enzymes for cofactor recycling. (**c**) Proposed scheme of regioselective C–H activation using catalyst control by integration of a Fl-Hal and palladium-catalysed SMC. Biocat., biocatalysis; chemocat., chemocatalysis.

products[24–29]. Several fungal Fl-Hals have also been reported that chlorinate phenolic natural products[30–32]. Together this suite of enzymes have the potential to regioselectively halogenate a range of aromatic scaffolds. With appropriate cofactor-recycling systems (Fig. 1b) Fl-Hals can afford halogenation reactions using benign inorganic halide (typically $MgCl_2$ or NaBr), oxygen (from air), and glucose or isopropanol as the only stoichiometric reagents[33,34]. Studies on broadening the substrate scope of the tryptophan halogenases RebH, PrnA and PyrH[33–36], as well as improving their productivity and scaleability[37], have meant that Fl-Hals are becoming increasingly viable as biocatalysts for aromatic halogenation. In addition, Fl-Hals have been used to halogenate natural products and synthetic substrates, which have been isolated or extracted and then derivatized in separate cross-coupling reactions under standard synthetic operating procedures[38–40]. Biohalogenation is also attractive, given that traditional aromatic halogenation chemistry utilizes deleterious reagents, lacks regiocontrol and can afford mixtures, including halogenated by-products that can be toxic and persistent in the environment.

We anticipated that the unique regiocontrol inherent to Fl-Hal enzymatic halogenation could set-up an overall catalyst-controlled C–H arylation system, if a working SMC could be successfully integrated under aqueous conditions. The indole heterocycle has been central to the development of contemporary C–H activation chemistry, driving the discovery of new chemocatalytic methods for functionalizing the C2 and C3 positions (the 'innate' positions for functionalization with an electrophile)[41–46]. Direct C–C bond formation at the arene C_6–H and C_7–H positions, however, is not currently possible using chemocatalysis—something we hoped to achieve with the integrated approach.

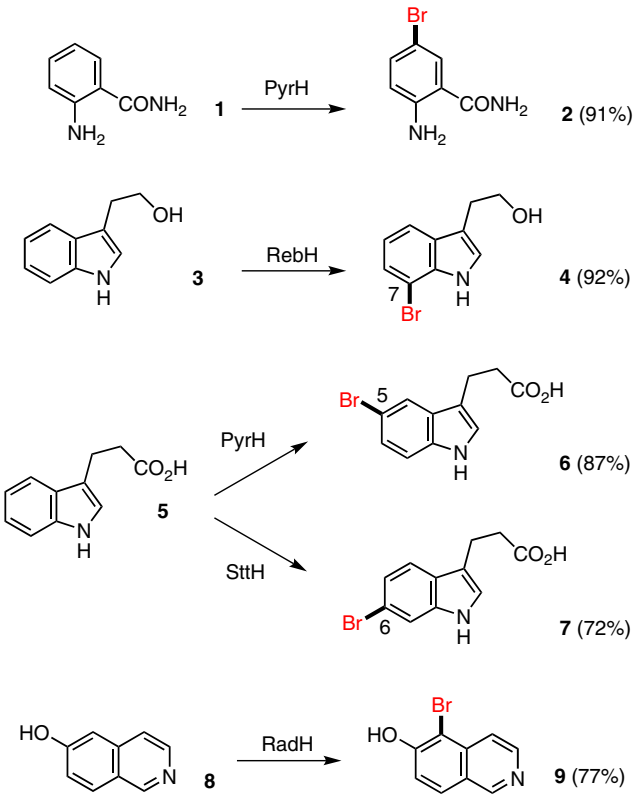


Figure 2 | The regioselective and regio-divergent bromination of a range of aromatic scaffolds. Regioselective halogenation of anthranilamide (**1**), tryptophol (**3**), 3-indole propionate (**5**) and 6-hydroxy isoquinoline (**8**) using a range of Fl-Hals to access different regiochemistries.

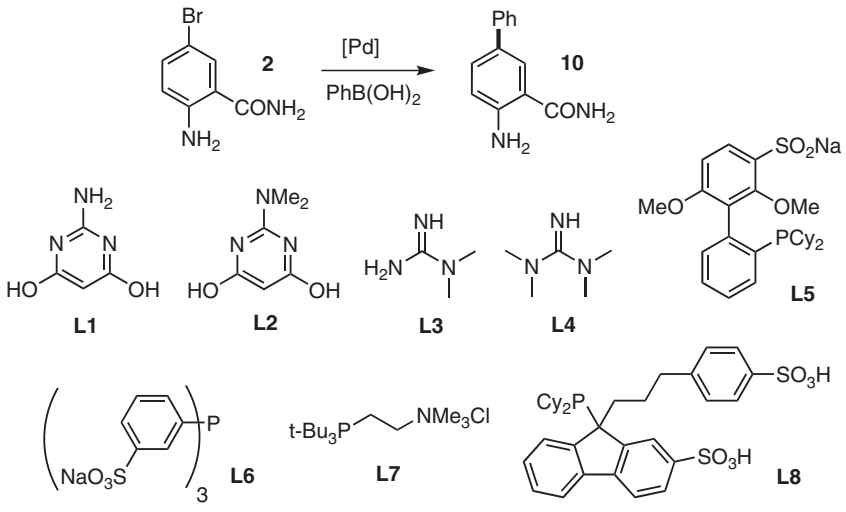
Results

Biocatalytic halogenation. To effect one-pot Fl-Hal-SMC reactions (Fig. 1c), we required enzymes and substrates that afford high yields of structurally diverse brominated products. First, we exploited our previous findings[33] that the tryptophan 5-halogenase PyrH can accept anthranilamide (**1**) to develop conditions for effective enzymatic synthesis of 5-bromoanthranilamide (**2**) (Fig. 2). Given the exquisite natural regio-complementarity of the flavin-dependent tryptophan halogenases we then set out to identify potential indole substrates for biocatalyst-controlled bromination and subsequent arylation. We found that tryptophol (**3**) could be efficiently brominated by RebH (a tryptophan 7-halogenase) to afford 7-bromotryptophol (**4**), after optimizing the conditions (Supplementary Fig. 1) that had been previously reported for chlorination of this substrate[34]. Initial attempts to brominate the 5-position of tryptophol (**3**) using PyrH in place of RebH gave very low conversion. The larger substrate 3-indolepropionic acid (**5**) however, was transformed by PyrH to give 5-bromo-3-indolepropionate (**6**) as a single regioisomer in good yield. In addition, the 6-bromo-3-indolepropionate (**7**) could be obtained in good regioisomeric excess and yield by using the tryptophan 6-halogenase SttH[27]. Collectively, these three halogenase enzymes established a clean, regiodivergent synthesis of C_5, C_6 and C_7 bromoindoles.

Finally, to extend the substrate diversity further we over-produced the radicicol halogenase RadH enzyme from the fungus *Chaetomium chiversii*[30]. RadH is similar in sequence to Rdc2—a related fungal halogenase that had been shown previously to chlorinate 6-hydroxy-isoquinoline (**8**)[31]. Accordingly conditions were developed whereby RadH could be used to prepare 5-bromo-6-hydroxyisoquinoline (**9**) in good yields.

Cross-coupling optimization. We next sought conditions whereby the SMC could be integrated with the biocatalysts, using the coupling of 5-bromoanthranilamide (**2**) with phenyl boronic acid as a test reaction (Fig. 3). On the basis of previous studies of SMCs in aqueous media with an open atmosphere[47,48], a number of pyrimidine and guanidine ligands were selected for screening (**L1**–**L4**). Of these, **L2** afforded highest conversion to 5-phenylanthranilamide (**10**) at 2.0 mM substrate (**2**) concentration, 2.5 mol% Pd loading, in potassium phosphate buffer in the absence of any cofactors or enzymes. In addition, a number of water-soluble phosphine ligands were also screened (**L5**–**L8**). Of these, tppts (**L6**) and sulfonated SPhos (**L5**) gave full conversion to the desired product (**10**). These two ligands were therefore carried forward in a base and palladium source screen (Supplementary Fig. 2). The reaction was found to be tolerant of many bases, including DIPEA, whilst the lack of any base afforded no product. In addition, conducting the reaction using **L5** and **L6** under air as opposed to inert atmosphere afforded no conversion, whilst the same reaction using **L2** under open air afforded good conversion.

The tolerance of the selected Pd catalysts towards the biocatalytic halogenation conditions was then probed. As NAD^+/NADH and FAD/$FADH_2$ are rich in heteroatoms and redox active they could possibly interfere with Pd-catalysis either through redox chemistry or coordination to metal species. In addition, it has been previously reported that soft donor atoms (for example, sulfur) on the surface of proteins can coordinate to transition metals potentially inhibiting their catalytic activity. Pd^{II} is also known to react with glucose under certain conditions (utilized to recycle NADH). The test reaction (Fig. 3) was therefore run in the presence of each of these components to



Ligand	Temperature (°C)	Conversion
L1	80	32
L2	80	86[a]
L3	80	56
L4	80	32
L5	50	100
L6	50	100[b]
L7	50	n.c.
L8	50	n.c.

Figure 3 | Ligand screening for the SMC of 2 with phenyl boronic acid. $Pd(OAc)_2$ was used as the palladium source in each case in a 2:1 ligand:Pd ratio. Conversions determined by analytical high-performance liquid chromatography (HPLC) or liquid chromatography mass spectrometry (LC-MS). n.c., no conversion. [a]81% isolated yield; [b]80% isolated yield. Further optimization results are available in the Supplementary Material (Supplementary Fig. 2).

determine their effect on SMC conversion. This revealed a somewhat deleterious effect from the presence of the cofactors and a major reduction in cross-coupling yields in the presence of proteins (Supplementary Fig. 3). Increasing the loading of palladium catalyst was generally advantageous in terms of conversion in the presence of NADH, FAD and glucose, but did not enhance cross-coupling when either the reductase or dehydrogenase proteins were present (Supplementary Fig. 4).

Halogenase SMC integration. Given that the presence of proteins caused the greatest reduction in cross-coupling yields, we reasoned that the enzymes would need to be compartmentalized or removed before cross-coupling, at least in the first instance. Pleasingly, we found that passing the PyrH biotransformation of anthranilamide (**1**) through a 10-kDa molecular-weight cut-off (MWCO) cellulose membrane to remove protein before cross-coupling was effective in this regard, affording very good conversion to arylated product (**10**). Following further optimization around this process (Supplementary Fig. 5), we could achieve an overall isolated yield of 79% of arylated product (**10**) using tppts ligand **L6**, $Pd(OAc)_2$ (50 mol%), and an excess of phenyl boronic acid (Fig. 4). Moreover, we showed that the enzymes that had been removed using the membrane filter could be recycled affording intermediate bromide (**2**) in 62% and 42% in the second and third cycles, respectively. Subsequent arylation of the combined biotransformation filtrates afforded arylated product (**10**) in 54% overall yield; a ca. twofold increase in the amount of material that would be obtained from the same amount of enzyme in a single arylation reaction. After determining the scope of this method with respect to boronic acids (Supplementary Fig. 6) and finding the SMC compatible with the conditions required for RadH-catalysed bromination of **8** (Supplementary Fig. 7), we then extended this membrane filtration method (method A) to the selective 5-arylation of 6-hydroxyisoquinoline (**8**), using RadH to produce **15**–**17** in good yields.

With a workable integrated arylation process in hand, we looked at ways to improve our method, principally with a view to increasing scale and reducing the high Pd content in the SMC reaction. The requirement for high catalyst loadings appeared to arise from a combination of dilute reaction volumes and inhibitory effect of cofactors present in the reaction medium. Indeed, other attempts to conduct palladium-catalysed chemistry in biological media and in the presence of proteins have required very high loading or suffered from poor yields[47,48]. Enhancing enzyme efficiency would mitigate these two issues, by affording higher concentrations of aryl bromide for cross-coupling. The group of Sewald have recently reported improvements in RebH productivity through the use of crosslinked enzyme aggregates (CLEAs)[37]. The individual subunits of RebH are crosslinked with each other, as well as the partner reductase and dehydrogenase proteins, via surface lysines using glutaraldehyde forming a RebH-reductase-dehydrogenase CLEA that is more efficient and tractable to handle[37]. Accordingly, we successfully prepared these

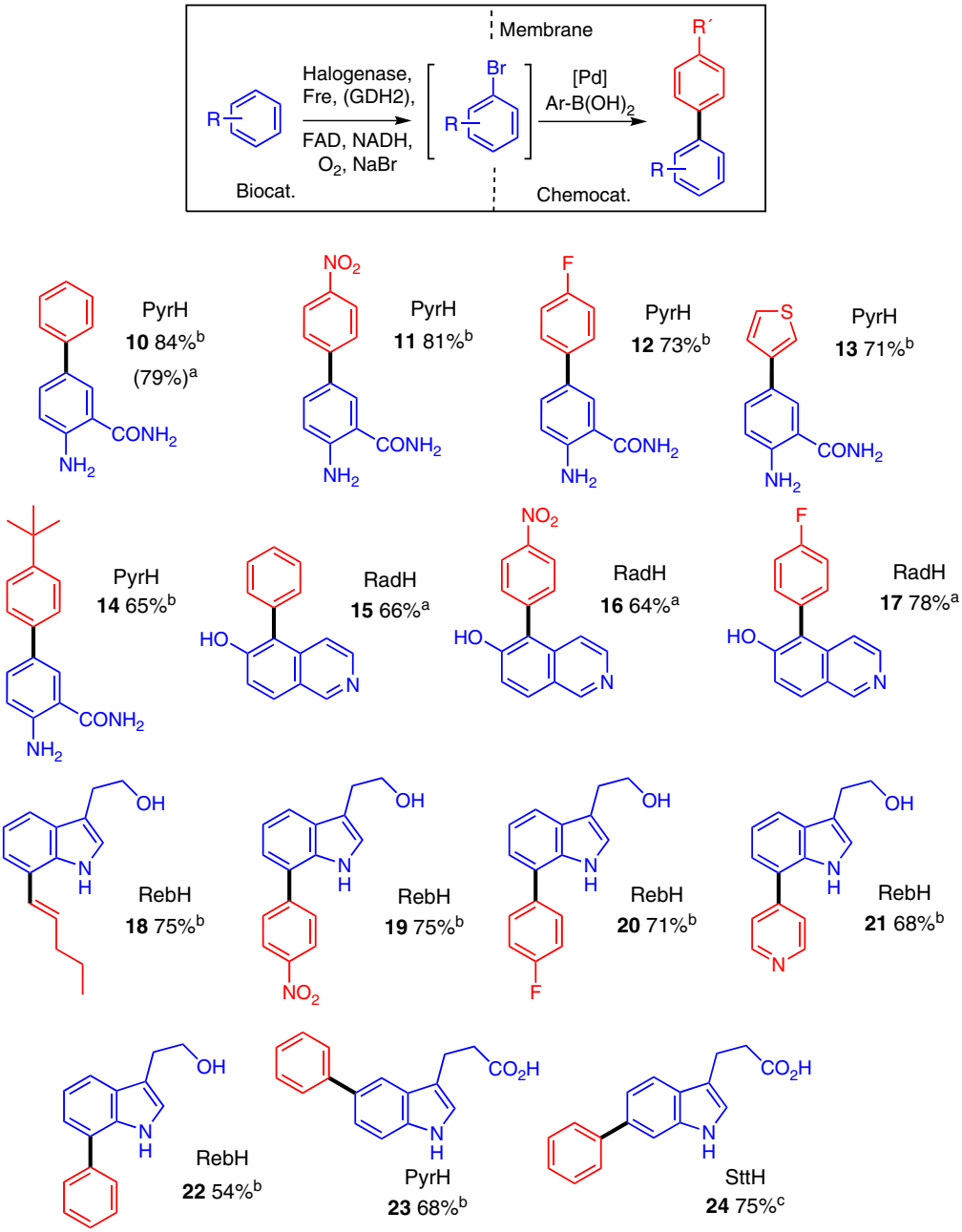


Figure 4 | Isolated yield of the regioselective and regiodivergent arylation. [a]Obtained using purified enzymes and a MWCO membrane (method A). [b]Obtained using a CLEA of the appropriate Fl-Hal and 10 mol% Pd-catalyst loading (method B). [c]Obtained from three successive cycles of halogenation using the same Fl-Hal. Biocat., biocatalysis; chemocat., chemocatalysis.

materials as active heterogeneous biocatalysts containing either PyrH, SttH or RebH. Although active CLEAs of the phenolic halogenase RadH were also obtained, they did not prove to be significantly more efficient than purified enzyme.

Initial screening of the PyrH-CLEA with anthranilamide (**1**) demonstrated improved efficiency, affording up to 3.0 mM aryl bromide. This enabled one-pot halogenase-SMC reactions to proceed to reasonable conversion (ca. 50%) without the need to remove or compartmentalize the CLEA from the Pd catalyst and other reagents, presumably as many heteroatoms of the protein had been effectively protected during the crosslinking process (Supplementary Fig. 8). We also found that the arylation of tryptophol (**3**) could be carried out in the same one-pot manner, with the RebH-CLEA and the Pd catalyst both present throughout (Supplementary Fig. 8). However, highest conversions were obtained when the CLEA was removed, by centrifugation or filtration, before cross-coupling of the supernatant from biocatalytic halogenation (Supplementary Figs 8 and 9). The scope of this method (method B) was found to be broad with respect to boronic acids (Supplementary Fig. 10), allowing the efficient arylation of anthranilamide (**1**) with both electron-rich, electon-poor and heterocyclic boronic acids using PyrH-CLEA with reduced palladium loading (10 mol%), affording the 5-arylated compounds in very good yield (**10**–**14**, Fig. 4). Tryptophol (**3**) was also efficiently coupled using electron-rich, electron-poor, heterocyclic and alkenyl boronic acids to afford 7-substituted products **18** to **22** in good yield with RebH-CLEA at lower Pd loading. 3-Indole propionate (**5**) was likewise effectively arylated to biaryl products **23** and **24**, with PyrH- and SttH-CLEAs, respectively. Along with the tryptophol examples

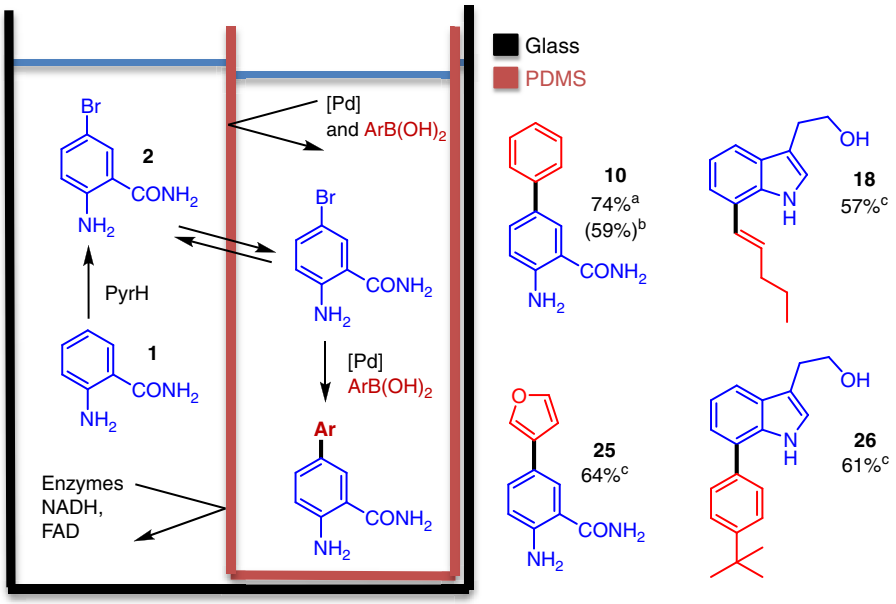


Figure 5 | PDMS compartmentalization of the regioselective halogenation cross-coupling cascade. [a]Obtained using purified enzyme and 10 mol% of **L2**.Pd(OAc)$_2$. [b]Obtained using halogenase CLEA and 2 mol% of **L2**.Pd(OAc)$_2$. [c]Obtained using halogenase CLEA and 10 mol% of **L2**.Pd(OAc)$_2$.

(**18**–**22**), this demonstrates the efficient and highly selective manipulation of each of the 5-, 6- and 7- C–H positions of the indole nucleus, under catalyst control, which is unattainable using existing synthetic methodology. In addition, the heterogeneous nature of the CLEA biocatalyst allowed it to be recycled over repeated cycles of bromination, with little loss in efficiency (Supplementary Fig. 11). In the case of the 6-arylation of 3-indolepropionic acid (**5**), this allowed an overall yield of 75% of **24** over three cycles with the SttH-CLEA. The increased stability and ease of preparation of the heterogeneous biocatalyst allowed this methodology to be scaled up to a 1.0-mmol scale to afford reasonable yields of **10** (>100 mg, 52%).

Polydimethylsiloxane compartmentalization. To further enhance the scope and efficiency of the halogenase-SMC process, we looked for alternative ways of separating the proteins from the palladium catalyst. While membrane filtration worked well, it did introduce an additional operation in the middle of the process and necessitated a deoxygenation step before SMC. To better streamline the process and develop a more efficient one-pot procedure, we examined polydimethylsiloxane (PDMS) thimbles as a compartmentalization strategy. This approach had been reported for a number of reactions that require the separation of incompatible reagents[19,49], including the combination of a Pd-catalyzed oxidation with a biocatalytic reduction[19].

Typically, one of the reactions is assembled inside a small thimble made of PDMS, which is then placed inside the larger bulk reaction vessel. As PDMS is a hydrophobic polymer, non-polar compounds (that is, substrates and intermediates in the cascade) should dissolve sufficiently well to pass through the PDMS walls, whilst charged reagents (Pd catalysts, enzymes and cofactors) would be unable to diffuse through PDMS and might therefore be compartmentalized. Initial tests showed that 5-bromoanthranilamide (**2**) would flux through PDMS well while glucose, NADH, FAD and the **L2**:Pd(OAc)$_2$ complex were all unable to effectively penetrate PDMS (Supplementary Fig. 12). SMC conditions using **L2**:Pd(OAc)$_2$ were therefore chosen to test the PDMS compartmentalization approach, as this catalyst functions well under open air and because alternative phosphine ligands have been shown to promote flux through PDMS. Accordingly conditions were optimized, which allowed the PyrH-halogenation of anthranilamide (**2**) using pure protein to be carried out with the cross-coupling reagents within the PDMS thimble (Fig. 5), resulting in a 'one-pot' reaction to give 5-phenylanthranilamide (**10**) in 74% yield. The immediate advantages of this approach (method C) are that protein removal and deoxygenation of the reaction buffer are not required. Moreover a lower catalyst loading (10 mol%) is achievable under this regime, compared with the removal of pure protein using filtration (50 mol%), due to the compartmentalization of cofactors and enzymes from the Pd catalyst. By using a PyrH-CLEA to afford higher concentration of aryl bromide in conjunction with PDMS compartmentalization (method D), the Pd loading was further reduced to 2 mol% to afford **10** in good yield (59%). This combination also allowed the coupling of deactivated, electron-rich, boronic acids such as 3-furyl boronic acid to give **25** in good yield (64%), a reaction that was unsuccessful using a cellulose membrane (MWCO filter). Finally, this method was also extended to include coupling of the 7-position of tryptophol (**3**), using RebH-CLEA, to give 7-pent-1-enyl tryptophol (**18**) and 4′-tertbutyl-7-phenyl tryptophol (**26**) in good yields.

Discussion

In summary, we have combined four regioselective flavin-dependent halogenase enzymes with a palladium-catalysed Suzuki–Miyaura cross-coupling to afford the regio-controlled arylation of a number of aromatic scaffolds (benzamides, isoquinolines and indoles). In each case, the regioselective C–H arylation transformations were previously inaccessible using stand-alone chemocatalysis or biocatalysis methods. Specifically, we report the selective C–H arylation of the 5-, 6- and 7- positions of the indole nucleus under catalyst control.

To overcome issues of catalyst incompatibility, we initially developed a filtration method using a 10-kDa MWCO filter (method A). This allowed efficient biocatalyst recycling for a number of cycles, affording approximately double the overall productivity that the same batch would in a single cycle. Biocatalyst efficiency was further improved by using CLEAs, which afforded higher aryl bromide concentration—enabling lower palladium loading to be employed (method B). We also

developed a PDMS thimble system to compartmentalize the chemo- and bio-catalysts. This method afforded similar yields of regioselectively arylated product as the MWCO filter system, but without the need for protein removal or solution deoxygenation before the cross-coupling chemistry (method C). Finally, the implementation of CLEAs with PDMS compartmentalization led to a further reduction in the required Pd loading down to 2 mol% (method D). Combined with directed evolution and other research aimed at improving the productivity and scalability of halogenase enzymes[37,50], the integrated method described here sets out a pathway to new C–H activation transformations that cannot be achieved through stand-alone biocatalysis or chemocatalysis. Research in this area is ongoing in our laboratories.

Methods

Method A: regioselective arylation 6-hydroxyisoquinoline. 6-hydroxyisoquinoline (**8**) (0.5 mM), NaBr (10 mM), FAD (1 μM), Fre (4 μM), RadH (25 μM) and NADH (2.5 mM) were added to 10 mM potassium phosphate buffer (pH 7.2) containing 1% v/v EtOH (40 ml) and incubated with shaking overnight (30 °C) before filtration through a 10-kDa MWCO filter (Vivaspin 20). CsF (4.8 eq) and boronic acid (30 eq) were then added to the filtrate. Following freeze-thaw degassing and backfill with N_2, tppts (0.5 mM) and Na_2PdCl_4 (0.25 mM) were added as deoxygenated stock solutions in water. The resultant solution was heated to 80 °C with stirring for 24 h. After cooling, pH was adjusted to 6 using HCl, and the reaction partitioned into CH_2Cl_2 before the removal of solvent *in vacuo* and purification.

Method B: regioselective arylation using CLEAs. Substrate (3.0 mM), FAD (10 μM), NaBr (30 mM), NADH (100 μM) were dissolved in 15 mM sodium phosphate buffer with 5% v/v isopropanol (30 ml, pH 7.4). Halogenase CLEA from 1.5 l culture (prepared as described on page S46 of the Supplementary Methods) was then resuspended into the reaction buffer and the resultant suspension incubated at room temperature with orbital shaking overnight. After removal of CLEA by centrifugation (10,000 r.p.m., 4 °C, 30 min), K_3PO_4/K_2CO_3 (10 eq) and boronic acid (5 eq) were added to the supernatant before degassing of the solution with a stream of nitrogen. Tppts (20 mol%) and Na_2PdCl_4 (10 mol%) were then added and the solution was stirred at 80 °C overnight. On cooling, the reaction was partitioned into EtOAc/CH_2Cl_2 and concentrated before purification.

Method C: regioselective arylation using PDMS thimbles. To a solution containing anthranilamide (2.0 mM), NaBr (100 mM), FAD (1 μM), glucose (20 mM), PyrH (20 μM), Fre (2 μM) and GDH2 (12 μM) in 10 mM potassium phosphate buffer (30 ml, pH 7.2) was added NADH (100 μM). A PDMS thimble containing **L2.**$Pd(OAc)_2$ (10 mol%), CsF (10 eq) and $PhB(OH)_2$ (5 eq) in water was then placed inside the Erlenmeyer flask containing the biotransformation. After incubation at room temperature overnight, the reaction was heated to 80 °C for 8 h. After cooling, the inner and outer chambers were combined and basified with 4 N NaOH. The thimble was then soaked in EtOAc to flux out any residual product, and the reaction mixture extracted into EtOAc before concentration and purification.

Method D: regioselective arylation using CLEAs and PDMS. Halogenase CLEA from 3 l of culture (prepared as described on page S46 of the Supplementary Methods) was suspended into 30 ml of reaction buffer containing anthranilamide (**1**) or tryptophol (**3**) (5.0 mM), NADH (100 μM), FAD (10 μM) and NaBr (30 mM) in 15 mM sodium phosphate buffer with 5% v/v isopropanol (30 ml, pH 7.4). To this suspension was added a PDMS thimble containing **L2.**$Pd(OAc)_2$ (10 mol%), CsF (10 eq) and $ArB(OH)_2$ (5 eq). After incubation with shaking at room temperature overnight, the reaction was heated to 80 °C for a further 24 h. Reactions were then worked up as previously described. To facilitate the recycling of the biocatalysts, the CLEA can be removed using centrifugation (10,000 r.p.m., 30 min., 4 °C) after overnight incubation at room temperature before addition of the PDMS thimble containing the SMC components.

Data availability. The authors declare that the data supporting the findings of this study are available within the article and its Supplementary Information Files.

References

1. Yu, J.-Q. & Shi, Z. *C-H Activation* Vol. 292 (Springer, 2010).
2. Clot, E. *et al.* C–F and C–H bond activation of fluorobenzenes and fluoropyridines at transition metal centers: how fluorine tips the scales. *Acc. Chem. Res.* **44,** 333–348 (2011).
3. Rossi, R., Bellina, F., Lessi, M. & Manzini, C. Cross-coupling of heteroarenes by C–H Functionalization: recent progress towards direct arylation and heteroarylation reactions involving heteroarenes containing one heteroatom. *Adv. Synth. Catal.* **356,** 17–117 (2014).
4. Engle, K. M., Mei, T. S., Wasa, M. & Yu, J. Q. Weak coordination as a powerful means for developing broadly useful C–H functionalization reactions. *Acc. Chem. Res.* **45,** 788–802 (2012).
5. Kuhl, N., Hopkinson, M. N., Wencel-Delord, J. & Glorius, F. Beyond directing groups: transition-metal-catalyzed C–H activation of simple arenes. *Angew. Chem. Int. Ed.* **51,** 10236–10254 (2012).
6. Leow, D., Li, G., Mei, T. S. & Yu, J. Q. Activation of remote meta-C–H bonds assisted by an end-on template. *Nature* **486,** 518–522 (2012).
7. Wang, X. C. *et al.* Ligand-enabled meta-C–H activation using a transient mediator. *Nature* **519,** 334–338 (2015).
8. Saidi, O. *et al.* Ruthenium-catalyzed meta sulfonation of 2-phenylpyridines. *J. Am. Chem. Soc.* **133,** 19298–19301 (2011).
9. McNally, A., Haffemayer, B., Collins, B. S. L. & Gaunt, M. J. Palladium-catalysed C–H activation of aliphatic amines to give strained nitrogen heterocycles. *Nature* **512,** 338–338 (2014).
10. Ueda, K., Yanagisawa, S., Yamaguchi, J. & Itami, K. A general catalyst for the beta-selective C–H Bond arylation of thiophenes with iodoarenes. *Angew. Chem. Int. Ed.* **49,** 8946–8949 (2010).
11. Hickman, A. J. & Sanford, M. S. Catalyst control of site selectivity in the Pd-II/IV-catalyzed direct arylation of naphthalene. *ACS Catal.* **1,** 170–174 (2011).
12. Neufeldt, S. R. & Sanford, M. S. Controlling site selectivity in palladium-catalyzed C–H bond functionalization. *Acc. Chem. Res.* **45,** 936–946 (2012).
13. Gormisky, P. E. & White, M. C. Catalyst-controlled aliphatic C–H oxidations with a predictive model for site-selectivity. *J. Am. Chem. Soc.* **135,** 14052–14055 (2013).
14. Gröger, H. & Hummel, W. Combining the 'two worlds' of chemocatalysis and biocatalysis towards multi-step one-pot processes in aqueous media. *Curr. Opin. Chem. Bio.* **19,** 171–179 (2014).
15. Lohr, T. L. & Marks, T. J. Orthogonal tandem catalysis. *Nat. Chem.* **7,** 477–482 (2015).
16. Allen, J. V. & Williams, J. M. J. Dynamic kinetic resolution with enzyme and palladium combinations. *Tetrahedron Lett.* **37,** 1859–1862 (1996).
17. Larsson, A. L. E., Persson, B. A. & Bäckvall, J.-E. Enzymatic resolution of alcohols coupled with ruthenium-catalyzed racemization of the substrate alcohol. *Angew. Chem. Int. Ed.* **36,** 1211–1212 (1997).
18. Denard, C. A. *et al.* Cooperative tandem catalysis by an organometallic complex and a metalloenzyme. *Angew. Chem. Int. Ed.* **53,** 465–469 (2014).
19. Sato, H., Hummel, W. & Gröger, H. Cooperative catalysis of noncompatible catalysts through compartmentalization: Wacker oxidation and enzymatic reduction in a one-pot process in aqueous media. *Angew. Chem. Int. Ed.* **54,** 4488–4492 (2015).
20. Makkee, M., Kieboom, A. P. G., Van Bekkum, H. & Roels, J. A. Combined action of enzyme and metal catalyst, applied to the preparation of D-mannitol. *J. Chem. Soc. Chem. Commun.* **19,** 930–931 (1980).
21. Köhler, V. *et al.* Synthetic cascades are enabled by combining biocatalysts with artificial metalloenzymes. *Nat. Chem.* **5,** 93–99 (2013).
22. Wang, Z. J., Clary, K. N., Bergman, R. G., Raymond, K. N. & Toste, F. D. A supramolecular approach to combining enzymatic and transition metal catalysis. *Nat. Chem.* **5,** 100–103 (2013).
23. Tenbrink, K., Seßler, M., Schatz, J. & Gröger, H. Combination of olefin metathesis and enzymatic ester hydrolysis in aqueous media in a one-pot synthesis. *Adv. Synth. Catal.* **353,** 2363–2367 (2011).
24. Keller, S. *et al.* Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens*. *Angew. Chem. Int. Ed.* **39,** 2300–2302 (2000).
25. Heemstra, J. R. & Walsh, C. T. Tandem action of the O_2- and FADH2-dependent halogenases KtzQ and KtzR produce 6,7-dichlorotryptophan for kutzneride assembly. *J. Am. Chem. Soc.* **130,** 14024–14025 (2008).
26. Yeh, E., Garneau, S. & Walsh, C. T. Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc. Natl Acad. Sci. USA* **102,** 3960–3965 (2005).
27. Zeng, J. & Zhan, J. Characterization of a tryptophan 6-halogenase from *Streptomyces toxytricini*. *Biotechnol. Lett.* **33,** 1607–1613 (2011).
28. Zehner, S. *et al.* A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005. *Chem. Biol.* **12,** 445–452 (2005).
29. Wagner, C., El Omari, M. & König, G. M. Biohalogenation: nature's way to synthesize halogenated metabolites. *J. Nat. Prod.* **72,** 540–553 (2009).
30. Wang, S. *et al.* Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. *Chem. Biol.* **15,** 1328–1338 (2008).
31. Zeng, J., Lytle, A. K., Gage, D., Johnson, S. J. & Zhan, J. Specific chlorination of isoquinolines by a fungal flavin-dependent halogenase. *Bioorg. Med. Chem. Lett.* **23,** 1001–1003 (2013).

32. Zhou, H., Qiao, K., Gao, Z., Vederas, J. C. & Tang, Y. Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. *J. Biol. Chem.* **285,** 41412–41421 (2010).
33. Shepherd, S. A. *et al.* Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. *Chem. Sci.* **6,** 3454–3460 (2015).
34. Payne, J. T., Andorfer, M. C. & Lewis, J. C. Regioselective arene halogenation using the FAD-dependent halogenase RebH. *Angew. Chem. Int. Ed.* **52,** 5271–5274 (2013).
35. Hölzer, M., Burd, W., Reißig, H.-U. & Pée, K.-H. V. Substrate specificity and regioselectivity of tryptophan 7-halogenase from *Pseudomonas fluorescens* BL915. *Adv. Synth. Catal.* **343,** 591–595 (2001).
36. Frese, M., Guzowska, P. H., Voß, H. & Sewald, N. Regioselective enzymatic halogenation of substituted tryptophan derivatives using the FAD-dependent halogenase RebH. *ChemCatChem.* **6,** 1270–1276 (2014).
37. Frese, M. & Sewald, N. Enzymatic halogenation of tryptophan on a gram scale. *Angew. Chem. Int. Ed.* **54,** 298–301 (2015).
38. Glenn, W. S., Nims, E. & O'Connor, S. E. Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **133,** 19346–19349 (2011).
39. Roy, A. D., Grüschow, S., Cairns, N. & Goss, R. J. M. Gene expression enabling synthetic diversification of natural products: chemogenetic generation of pacidamycin analogs. *J. Am. Chem. Soc.* **132,** 12243–12245 (2010).
40. Durak, L. J., Payne, J. T. & Lewis, J. C. Late-stage diversification of biologically active molecules via chemoenzymatic C–H activation. *ACS Catal.* **6,** 1451–1454 (2016).
41. Sandtorv, A. H. Transition metal-catalyzed C–H activation of indoles. *Adv. Synth. Catal.* **357,** 2403–2435 (2015).
42. Itahara, T., Ikeda, M. & Sakakibara, T. Alkenylation of 1-acylindoles with olefins bearing electron-withdrawing substituents and palladium acetate. *J. Chem. Soc. Perkin. Trans.* **1,** 1361–1363 (1983).
43. Grimster, N. P., Gauntlett, C., Godfrey, C. R. A. & Gaunt, M. J. Palladium-catalyzed intermolecular alkenylation of indoles by solvent-controlled regioselective C–H functionalization. *Angew. Chem. Int. Ed.* **44,** 3125–3129 (2005).
44. Stuart, D. R. & Fagnou, K. The catalytic cross-coupling of unactivated arenes. *Science* **316,** 1172–1175 (2007).
45. Kalyani, D., Dick, A. R., Anani, W. Q. & Sanford, M. S. A simple catalytic method for the regioselective halogenation of arenes. *Org. Lett.* **8,** 2523–2526 (2006).
46. Lebrasseur, N. & Larrosa, I. Room temperature and phosphine free palladium catalyzed direct C-2 arylation of indoles. *J. Am. Chem. Soc.* **130,** 2926–2927 (2008).
47. Chalker, J. M., Wood, C. S. C. & Davis, B. G. A convenient catalyst for aqueous and protein Suzuki-Miyaura cross-coupling. *J. Am. Chem. Soc.* **131,** 16346–16347 (2009).
48. Gao, Z., Gouverneur, V. & Davis, B. G. Enhanced aqueous Suzuki-Miyaura coupling allows site-specific polypeptide 18F-labeling. *J. Am. Chem. Soc.* **135,** 13612–13615 (2013).
49. Runge, M. B., Mwangi, M. T., Miller, A. L., Perring, M. & Bowden, N. B. Cascade reactions using $LiAlH_4$ and grignard reagents in the presence of water. *Angew. Chem. Int. Ed.* **47,** 935–939 (2008).
50. Poor, C. B., Andorfer, M. C. & Lewis, J. C. Improving the stability and catalyst lifetime of the halogenase RebH by directed evolution. *ChemBioChem.* **15,** 1286–1289 (2014).

Acknowledgements

We acknowledge BBSRC (grant BB/K00199X/1), EPSRC, CoEBio3 and GlaxoSmithKline for support. We also thank Dr Anna-Winona Struck for helpful conversations.

Author contributions

J.M. and M.F.G. conceived and supervised the project; J.L., J.-M.H. and H.H.S. performed integrated arylation reactions; J.L. optimized the integrated process whilst J.L. and H.H.S. screened conditions for the SMC; B.R.K.M. and S.A.S. developed expression and initial reaction conditions for RadH and SttH, respectively; J.L. and S.A.S. prepared functional CLEAS; J.L. prepared and carried out integrated reactions with PDMS thimbles; J.L., M.F.G. and J.M. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Latham, J. *et al.* Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation. *Nat. Commun.* 7:11873 doi: 10.1038/ncomms11873 (2016).

Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2016, **14**, 9354

Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes†

Binuraj R. K. Menon, Jonathan Latham, Mark S. Dunstan, Eileen Brandenburger, Ulrike Klemstein, David Leys, Chinnan Karthikeyan, Michael F. Greaney, Sarah A. Shepherd and Jason Micklefield*

Received 24th August 2016,
Accepted 2nd September 2016

DOI: 10.1039/c6ob01861k

www.rsc.org/obc

Flavin-dependent halogenase (Fl-Hal) enzymes have been shown to halogenate a range of synthetic as well as natural aromatic compounds. The exquisite regioselectively of Fl-Hal enzymes can provide halogenated building blocks which are inaccessible using standard halogenation chemistries. Consequently, Fl-Hal are potentially useful biocatalysts for the chemoenzymatic synthesis of pharmaceuticals and other valuable products, which are derived from haloaromatic precursors. However, the application of Fl-Hal enzymes, *in vitro*, has been hampered by their poor catalytic activity and lack of stability. To overcome these issues, we identified a thermophilic tryptophan halogenase (Th-Hal), which has significantly improved catalytic activity and stability, compared with other Fl-Hal characterised to date. When used in combination with a thermostable flavin reductase, Th-Hal can efficiently halogenate a number of aromatic substrates. X-ray crystal structures of Th-Hal, and the reductase partner (Th-Fre), provide insights into the factors that contribute to enzyme stability, which could guide the discovery and engineering of more robust and productive halogenase biocatalysts.

Introduction

Halogen substituents are present in a high proportion of the leading pharmaceuticals, as well as many agrochemicals, polymers and other useful synthetic materials.[1–3] Moreover, organohalogens are frequently used as intermediates in synthesis, with haloaromatic compounds serving as the key components for a wide range of transition metal catalysed cross-coupling reactions.[4–8] Traditional halogenation chemistries often utilise toxic reagents, employing reaction conditions that are becoming increasingly unsustainable, and often lack regioselectivity with aromatic substrates.[9–13] Consequently, there has been an interest in developing enzymes that can utilise benign inorganic halides in water to deliver haloaromatic compounds.[4–16]

Nature has evolved halogenase enzymes for the biosynthesis of a diverse array of halogenated natural products including many antibiotics, antitumor agents and other bioactive compounds, with the halogen substituents often contributing to the biological activity of these secondary metabolites.[17–20] The Flavin-dependent halogenase (Fl-Hal) enzymes, which typically halogenate aromatic precursors, are particularly common in secondary metabolism. Fl-Hals require an NADPH-dependent flavin reductase enzyme that can reduce flavin-adenine dinucleotide (FAD) to $FADH_2$, which is utilised to oxidise inorganic chloride or bromide to hypohalous acid.[21] X-ray crystal structures of the flavin-dependent tryptophan halogenases,[22,23] suggest that hypohalous acid is then channelled through an intra-enzymatic tunnel to the substrate binding site, where an active site lysine positions this electrophile in a spatially defined orientation to the aromatic substrate, either through formation of an intermediate chloramine or hydrogen bonding, effectively controlling the regioselectivity.[22,24] Fl-Hal have emerged as promising enzymes for use as biocatalysts due to their potential to regioselectively chlorinate or brominate synthetic, as well as natural, aromatic compounds that can also be further derivatised using cross-coupling chemistry.[4–8,25–27]

The potential of Fl-Hals is best exemplified by the tryptophan 5-, 6- and 7-halogenases that can catalyse the regiodivergent halogenation of the C5, C6 or C7 positions of the substrate indole moiety, which is beyond the scope of current non-enzymatic halogenation methods. The tryptophan halogenases have been shown to exhibit some substrate promiscuity and accept a number of non-native electron rich aromatic substrates. Targeted and random mutagenesis strategies have

School of Chemistry and Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester, M1 7DN, UK.
E-mail: jason.micklefield@manchester.ac.uk

†Electronic supplementary information (ESI) available: For experimental methods, supplementary figures, tables and NMR spectra. See DOI: 10.1039/c6ob01861k

also been employed to further broaden the substrate scope and switch the regioselectivities of the tryptophan halogenases.[28–31] In addition, we recently showed how tryptophan halogenases and other Fl-Hal enzymes can be integrated with palladium-catalyzed cross-coupling chemistry, in one-pot reactions, to affect the direct regioselective arylation or alkenylation of C–H positions in various aromatic scaffolds;[32] such transformations are inaccessible using stand-alone chemocatalysis or biocatalysis and further demonstrate the potential of Fl-Hal in synthetic applications (Fig. 1).

Despite their promise, the synthetic utility of Fl-Hal enzymes is limited by their relatively low activity and poor stability.[30,33] In common with other enzymes from secondary metabolism, Fl-Hal exhibit lower catalytic efficiencies than many of the enzymes from primary metabolism, which have presumably been under greater evolutionary pressure.[34] Thus, while Fl-Hal have been used *in vivo* to produce large quantities of halogenated natural products including antibiotics, such as vancomycin and griseovulfin,[35–38] their low activity and instability currently precludes their deployment *in vitro* for larger scale synthetic applications. To address the low productivity of Fl-Hal, cross-linked enzyme aggregates (CLEAS) containing Fl-Hal have been developed,[39] which have enhanced stability and can be recycled to increase the scale of biohalogenation reactions.[31,32,39] It has also been demonstrated that increasing the thermal stability of Fl-Hal by directed evolution could be a way to improve the limited product yield and catalytic lifetime of these enzymes.[33] However, this approach is usually time consuming and requires multiple rounds of mutagenesis to achieve the desired levels of improvement. In this paper, we take the alternative strategy of exploring natural biosynthetic pathways in thermophilic bacteria, to identify more robust and productive variants of Fl-Hal. Thermostable enzyme variants have been shown to possess many advantages, including prolonged catalyst life, as well as increased tolerance to proteolysis and organic solvents;[33,40–43] all of which would make these enzymes better candidates for biocatalysis and further engineering. In addition, a thermostable Fl-Hal would be more compatible with chemocatalysts that often require elevated temperatures, which could enable development of more integrated one-pot chemobio-transformations.[32]

Results and discussion

Substrate scope and regioselectivity of a thermophilic halogenase

BLAST analysis of published genome sequences revealed the presence of putative Fl-Hal encoding gene from a thermophilic halotolerant *Streptomyces violaceusniger* strain SPC6 (Table S1†).[44,45] The putative halogenase (Th-Hal) shows similarity to tryptophan halogenases (Fig. S1†), and is present within an uncharacterised biosynthetic gene cluster that includes gene(s) encoding a hybrid nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) assembly line. To explore the properties of Th-Hal, the protein was overproduced in *E. coli*. In addition, a thermophilic flavin reductase (Th-Fre) enzyme from *Bacillus subtilis* WU-S2B, which had been identified previously,[46,47] was also overproduced in *E. coli* to provide a potentially thermostable halogenase system.

The substrate specificity of Th-Hal was tested with tryptophan (**1**) and a series of related compounds including 1-methyltryptophan (**2**), 5-hydroxytryptophan (**3**), kynurenine (**4**), anthranilic acid (**5**) and anthranilamide (**6**) (Fig. 2). With tryptophan, halogenation occurs solely at the C6 position indicating that Th-Hal is a tryptophan 6-halogenase enzyme,

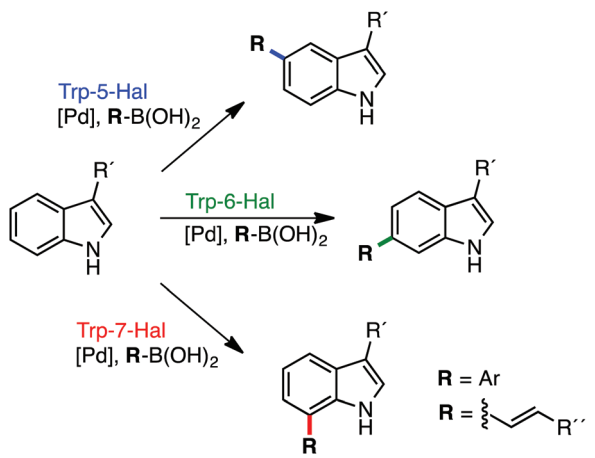


Fig. 1 Regiodivergent functionalisation of aromatic scaffolds using Fl-Hal C–H activation. Tryptophan 5-, 6- & 7-halogenases and other Fl-Hal enzymes have been integrated with Pd-catalysis to deliver functionally diverse arylated and alkenylated products, in one-pot reactions, under mild, aqueous conditions, *via* aryl bromide intermediate formed *in situ*.[32]

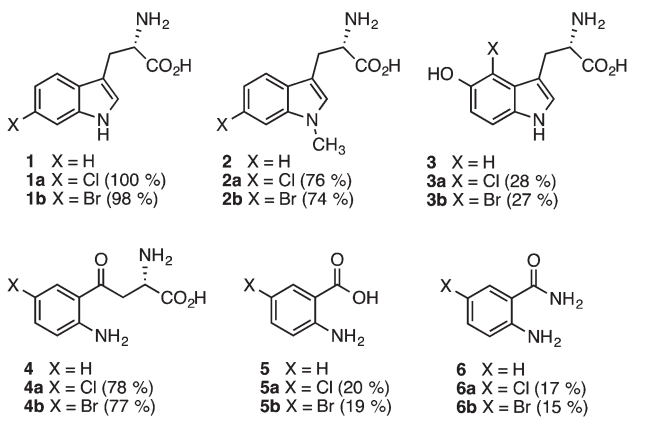


Fig. 2 Substrates accepted by Th-Hal and products obtained. HPLC conversations are shown in brackets. Reaction conditions: Th-Hal (2.5 μM), *E. coli* flavin reductase (Fre) (1 μM), FAD (1 μM), NADH (2.5 mM), $MgCl_2$ (10 mM) and substrate (0.5 mM) in 10 mM potassium phosphate buffer, pH 7.4, Reactions (**2–6**) were measured in HPLC after 3 hours incubation at 30 °C and 200 rpm. Reactions with tryptophan (**1**) were complete after incubation for 1 hour.

similar to SttH from *Streptomyces toxytricini* and KtzR from Kutzneride biosynthetic pathways.[48,49] Halogenation of 1-methyltryptophan (**2**) also occurred at the C6 position, whilst 5-hydroxytryptophan (**3**) was halogenated at C4, suggesting a different mode of substrate binding.[50] The non-indolic substrates (**4–6**) are all halogenated at the more chemically activated C5 position. Overall, this reveals that Th-Hal shows similar substrate specificity and regioselectivity to the SttH halogenase which we explored previously.[31]

As anticipated, tryptophan is the best substrate for Th-Hal affording quantitative chlorination and bromination (Fig. 2). Activity of Th-Hal was reduced upon introduction of a methyl group on to the substrate indole nitrogen (**2**), presumably through the loss of a hydrogen bonding contact to indole moiety. Activity is further decreased when an hydroxyl group is introduced into the substrate adjacent to the halogenation position (**3**). Of the non-indolic substrates tested kynurenine (**4**) is preferred over anthranilic acid (**5**) and anthranilamide (**6**), which gave only modest conversions over the same time period. In all cases, conversions for bromination reactions were comparable with those for chlorination (Fig. 2).

Activity and stability of the thermophilic halogenase and reductase

The kinetic parameters of the thermophilic halogenase enzyme, with the natural tryptophan substrate, were compared with related Fl-Hal including the tryptophan 5-halogenase PyrH, tryptophan 6-halogenases SttH and KtzR along with tryptophan 7-halogenases PrnA and RebH which are all derived from mesophilic bacteria (Table 1). By selecting halogenases with different regioselectivities, a better comparison of kinetic parameters for Fl-Hal in this class could be obtained. The k_{cat} values of Th-Hal with tryptophan was found to be 4.3 min^{-1} indicating that this enzyme is more active towards tryptophan than any of the other halogenases tested at 30 °C. Th-Hal was found to be *ca*. 2 times faster than PyrH, which is the most efficient of the other Fl-Hal tested here, with similar substrate binding (K_m) values. Th-Hal exhibits both improved catalytic turnover and binding parameters, compared with the remaining Fl-Hal tested, resulting in between 5- to 35-fold higher catalytic efficiency (k_{cat}/K_m). Though Th_hal showed a higher catalytic turnover value at 45 °C, it was found that catalytic efficiency decreases with increase in temperature mainly due to poor binding of tryptophan at higher temperature (Table 1).

Table 1 A kinetic comparison of the thermophilic halogenase and other tryptophan halogenase enzymes with tryptophan at 30 °C

Enzyme	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M
Th-Hal (30 °C)	4.3 ± 0.5	12.2 ± 1.8	0.35 ± 0.07
Th-Hal (45 °C)	5.1 ± 0.4	20.4 ± 1.3	0.25 ± 0.03
PyrH	2.4 ± 0.4	15.2 ± 4.2	0.16 ± 0.05
SttH	1.7 ± 0.1	25.3 ± 3.2	0.07 ± 0.01
PrnA	1.1 ± 0.1	20.7 ± 0.1	0.05 ± 0.005
KtzR	0.4 ± 0.1	34.1 ± 2.1	0.01 ± 0.003
RebH	0.6 ± 0.1	28.7 ± 1.3	0.02 ± 0.004

In order to explore the feasibility of establishing a thermophilic biocatalytic halogenation system, melting temperatures of individual halogenase enzymes and flavin reductase enzymes were determined using Circular Dichroism (CD) spectroscopy (Fig. 3, S2 & S3†). Comparison of the Th-Hal melting temperature with that of PyrH, SttH, PrnA, KtzR and RebH revealed that the Th-Hal has a higher thermal stability, with a melting temperature (47.8 °C) that is approximately 10 °C above that of the other halogenase enzymes. The melting temperature of the thermophilic flavin reductase (Th-Fre) from *Bacillus subtilis* WU-S2 was also compared with two reductase partners from *E. coli* which are commonly used with Fl-Hal *in vitro*, the NAD(P)H-flavin reductase Fre and the NAD(P)H-dependent FMN reductase SsuE.[51,52] This showed that the Th-Fre has a melting temperature of 49.0 °C which was 9 °C above that of the *E. coli* Fre and 15.4 °C above that for SsuE, indicating that Th-Fre has a higher thermostability, as previously reported,[46,47] and should be more compatible with Th-Hal in high temperature biocatalytic halogenation reactions.

The kinetic parameters and co-factor specificities of the flavin reductase enzymes were determined, showing that Th-Fre exhibited a higher flavin reduction rate and reactivity towards both FMN and FAD in the presence of NADH compared to Fre and SsuE (Tables 2, S2 & S3†). The K_M values indicated that Th-Fre has a higher preference and binding affinity towards FMN (K_M = 2.5 μM) over FAD (16.3 μM). The kinetic parameters for flavin reductase enzymes with FMN and NADH were also obtained at 30 °C and 45 °C (Table 2), showing that Th-Fre has a higher activity and stability over the *E. coli* flavin reductase enzymes and confirmed Th-Fre to be suitable

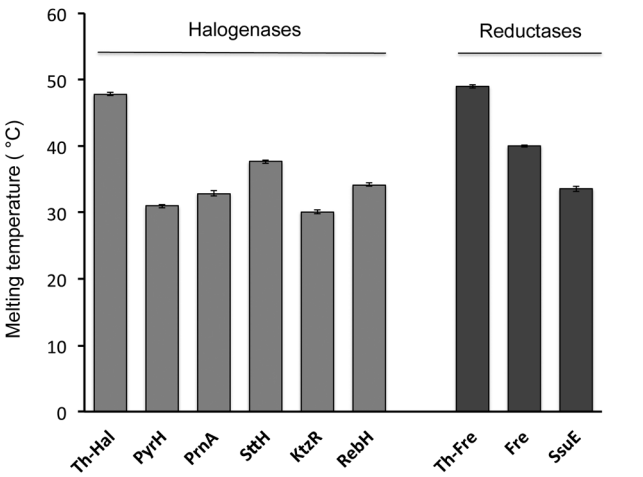


Fig. 3 Thermal stability of halogenase and flavin reductase enzymes determined by CD. The CD spectrum was measured every 2 °C, from 20 °C up to 80 °C with 10 μM enzyme in a 0.5 mm path length cuvette at fixed temperature for 2 minute with the temperature ramp of 0.2 °C per minute. The absorbance changes at 222 nm was plotted against temperature, and fitted with a nonlinear curve fitting function to determine protein melting temperatures.

Table 2 The kinetic parameters for the FMN reduction catalysed by flavin reductase enzyme with NADH at 30 °C and 45 °C

Enzyme	Temp.	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M
Th-Fre	30 °C	16.0 ± 1.0	2.5 ± 0.3	6.4 ± 0.9
	45 °C	26.0 ± 2.0	2.8 ± 0.2	9.3 ± 1.0
SsuE	30 °C	8.1 ± 0.20	8.3 ± 0.1	0.98 ± 0.03
	45 °C	1.9 ± 0.02	10.2 ± 0.4	0.02 ± 0.002
Fre	30 °C	4.3 ± 0.4	14.7 ± 0.3	0.29 ± 0.03
	45 °C	0.3 ± 0.03	16.2 ± 0.5	0.02 ± 0.002

partner for Th-Hal or other Fl-Hal in higher temperature reactions.

The activity of the Th-Hal, was next compared the other mesophilic tryptophan halogenases (PyrH, SttH, PrnA, KtzR and RebH), in combination with the thermostable reductase, Th-Fre, at two different temperatures. This showed (Fig. 4) that Th-Hal/Th-Fre affords slightly higher conversion of tryptophan to 6-chlorotryptophan at 45 °C (74%) than at 30 °C (69%) over 30 min time period of the assay (Fig. 4). In contrast the mesophilic halogenases all gave significantly lower conversions to the corresponding chlorinated tryptophan derivatives, particularly at the higher temperature. In addition, the other substrates (**2–6**) were tested with Th-Hal/Th-Fre at different temperatures, revealing that improved conversions to chlorinated products (**2a–6a**) could be obtained at 45 °C compared with the normal operating temperature (Fig. S4†). The halogenation efficiency of Th_hal was then tested with the addition of different solvents to understand the solvent tolerance (Fig. S5†). Th_hal retained significant activity in different polar solvents and the activity pattern showed that Th_hal is more tolerant and stable towards the solvents tested than its mesophilic counterpart SttH. Both of these enzymes however showed a reduced activity in non-polar solvents such as dichloromethane and acetonitrile. Taken together these results indicate that Th-Hal in combination with Th-Fre, provides a more efficient, stable and robust halogenase system, than the previously characterised tryptophan halogenase, for use *in vitro* with both the natural and non-native substrates.

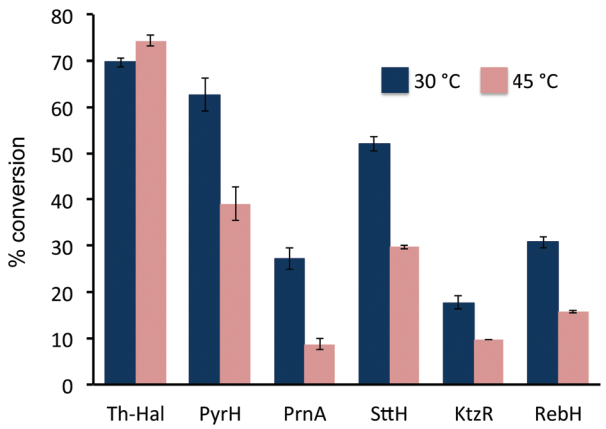


Fig. 4 Effects of temperature on % chlorination of tryptophan with halogenase enzymes and Th-Fre. Reaction conditions: halogenase enzyme (2.5 μM), Th-Fre (1 μM), FAD (1 μM), NADH (2.5 mM), $MgCl_2$ (10 mM) and tryptophan (0.5 mM) in 10 mM potassium phosphate buffer, pH 7.4, Reactions were measured by HPLC after 30 minutes incubation at 30 °C (or 45 °C) and 200 rpm.

Crystal structure and amino acid variability of Th-Hal

An X-ray crystal structure of apo Th-Hal was obtained at 2.33 Å using tryptophan 7-halogenase (PrnA, PDB 2AQJ) as the molecular replacement model (Fig. 5, Table S4†). Th-Hal, shares 76.2% sequence identity with SttH the structure of which we had determined previously.[31] The crystal structure indicates that Th-Hal forms a dimer, similar to other Fl-Hal, which was also observed in gel filtration chromatography.[48,49] The overall structure of Th-Hal is similar to other tryptophan halogenases – a single-domain protein in which several helices and two large β-sheets form a "box" and a triangular pyramid dominated by helixes. The secondary structural arrangements of other tryptophan halogenase enzymes are conserved in Th-Hal, despite the fact that Th-Hal sequences has a divergence from PyrH and PrnA sequences reflecting the individual regioselectivity (Fig. 5B). The catalytic lysine (K75) and glutamate

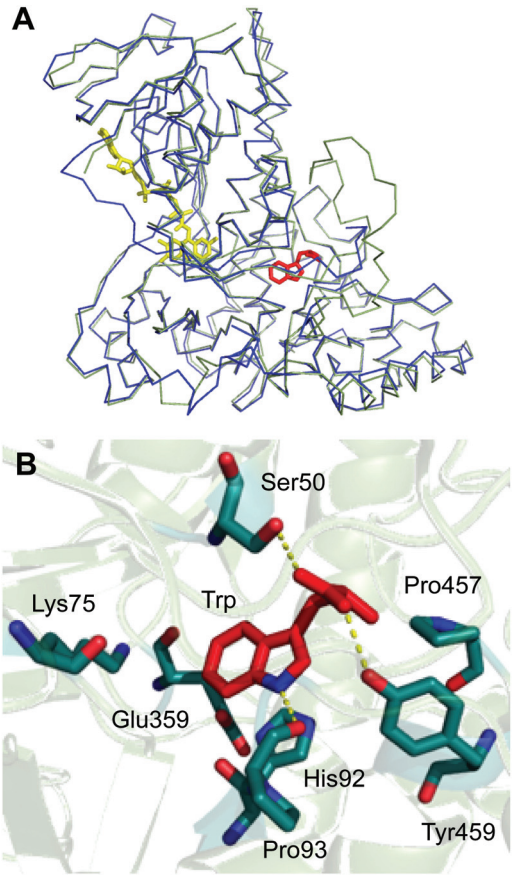


Fig. 5 Crystal structure of thermophilic tryptophan halogenase (Th-Hal) enzyme. (A) Structural alignment of Th-Hal (green ribbon) with SttH (blue ribbon). (B) Active site residues and the predicted tryptophan binding mode in the active site of Th-Hal.

(E359) residues of Th-Hal align well with those in SttH (K79, E363), PyrH (K75, E354) and PrnA (K79, E346), indicating that the alterations in tryptophan-binding mode leads to different regioselectivity of these enzymes (Fig. 5B).[22,31,48,49] The crystal structures of other tryptophan halogenase enzymes shows that substrate binding in the active site is such that the chlorination site (C6 of tryptophan in Th-Hal) is always directed towards the reactive chlorinating species controlled by the catalytic lysine.[22,31,48,49,52] A possible tryptophan bound model for the resolved Th-Hal structure could therefore be predicted (Fig. 5B) using this fact.

Similar to the PyrH, PrnA and SttH structures, the benzene portion of the tryptophan indole ring is sandwiched between the π systems of H92 and F94 residues. The indole nitrogen is hydrogen bonded to the carbonyl of P93, as in PyrH, whereas the hydroxyl of Y454 and S50 could form hydrogen bonding interactions with the tryptophan carboxyl group. Residues L455, P456 and P457 in Th-Hal, which are comparable to residues L460, P461 and P462 in SttH, are also in a close proximity to the tryptophan-binding site and are oriented similar to that in SttH. The loop region involving these residues has been previously shown to be important in contributing towards the regioselectivity of SttH and PrnA.[31,48,53,54] The residues near to and part of the FAD binding strap region (A41 to E46 in Th-Hal, T45 to E50 in SttH) is completely disordered in apo Th-Hal structure as the monomer structures are without FAD bound.[31] The inherent structural flexibility of this region was reported previously as the FAD binding region is solvent exposed in tryptophan halogenases.[53,54] Other than this, the secondary structural arrangements of FAD binding modules of Th-Hal are very similar to that of SttH. The signature motifs of the flavin dependent halogenases GxGxxG (residues 9–14, FAD consensus sequence) and WxWxIP (residues 284–289, to prevent monoxygenase activity) that are located in the FAD module are also structurally conserved in Th-Hal and in SttH.

The lower thermal stability of SttH means that the structural and sequence comparison between Th-Hal with SttH could shed more light on the structural features that are responsible for the increased thermal stability and reactivity in Th-Hal. Conversely, this information could also provide insight into the poor catalytic efficiency and stability of the tryptophan halogenase family. The amino acid variability in Th-Hal was analysed against SttH. Sequence alignment with SttH showed that the sequence variations are present in the entire length of protein (Fig. S8†). However, when these changes were marked in the crystal structure of Th-Hal; the identified pattern showed that most of the variations are at surface exposed residues (Fig. 6). Supporting this observation, there were no mutations at the dimeric interface region. About 40% of these changes were to introduce polar residues on the Th-Hal surface regions. It was noted before that introducing surface polar residues could influence surface hydrogen bonds and thus increase the stability of these proteins by inhibiting aggregation.[55–58] As more diverse chemical substitutions are easily incorporated in surface residues rather than the catalytic inner core residues of proteins during evolution, it is possible that nature has adopted this strategy to produce a more active and thermally stable halogenase enzyme.[55–58]

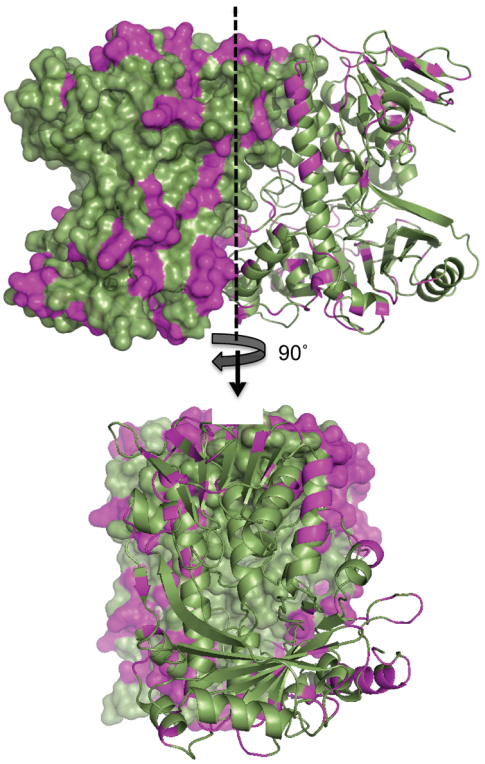


Fig. 6 Surface map structure of Th-Hal (PDB 5LV9) showing its amino acid variation from SttH structure. One monomeric unit is shown as surface map format while the other unit as a cartoon. The sequence variation is shown in magenta colour. The Th-Hal dimeric interface (after 90° rotation) is also shown for clarity.

Crystal structure of Th-Fre

The crystal structure of Th-Fre was resolved at 2.53 Å using the general stress protein 14 (PDB 3F2 V) as the molecular replacement model (Fig. 7, Table S4†). Th-Fre is a homodimer in the

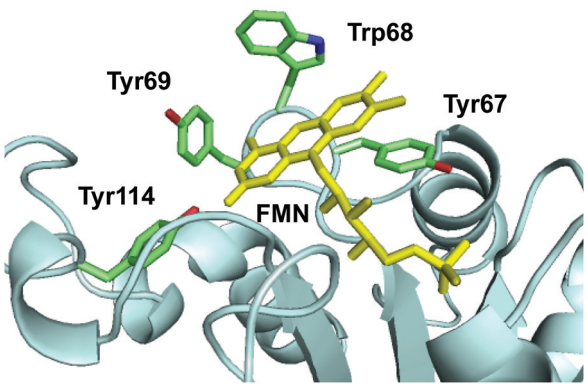


Fig. 7 The active site structure of thermophilic flavin reductase (Th-Fre) enzyme (PDB 5LVA). The aromatic core formed by residues Tyr67, Trp68, Tyr69, Tyr114 in the active site around FMN cofactor of Th-Fre is shown. FMN is shown as yellow sticks.

asymmetric unit and both monomeric units contained FMN that was incorporated into the protein from *E. coli* cells during recombinant expression. Each monomeric unit comprises of a classical flavodoxin_2 domain formed from five β-sheets that connected with loops and helixes. FMN is bound at the apex of this flavodoxin_2 domain and in the asymmetric unit monomers are arranged in antiparallel directions. The planar stacking interactions with the FMN isoalloxazine ring are mainly from the phenyl ring of Tyr67. Along with this two additional aromatic residues Trp68 and Tyr69 are thought to be involved in substrate stacking and associated dynamics. Tyr114 forms a hydrogen bonding interaction with isoalloxazine ketonic group at position 2. His9, Val16, Val16 and Asn18 residues are in direct contact with phosphate tail of FMN (Fig. 7).

The resolved structure showed that Th-Fre is structurally similar to quinone reductase 2 enzymes (Fig. 8). In this category the closest structurally identical protein is NADPH-quinone reductase from *P. pentosaceus* (PDB 3HA2). Quinone reductase 2 enzymes are one of known human targets for anti-malarial drugs, primaquine and chloroquine.[59,60] The most well-characterised quinone reductase 2 enzyme is human quinone reductase 2 (Nqo2) (PDB 2FGL, 2FGI), which is unique in that it uses dihydronicotinamide riboside (NRH) as a reducing coenzyme rather than NADH or NADPH.[59,60] The structural and sequence alignment of Th-Fre with human quinone reductase 2 shows that Th-Fre is a shorter protein and there is a 30 amino acid insertion at the N terminal of human quinone reductase 2 that forms an extended helix above the flavin domain of opposite monomer (Fig. 8, S9 and S10†). Whether this extended region in human quinone reductase 2 provide specific interactions for NRH binding and π stacking over flavin isoalloxazine ring is not clear. The mesophilic flavin reductase enzymes, Fre and SsuE from *E. coli*, are not homologous to Th-Fre (25.2% and 22.4% sequence identitity with Fre and SsuE respectively). Crystal structure of Fre (PDB 1QFJ) also showed that Th-fre is structurally different to Th-Fre making it difficult to plot the sequence variability of other known flavin reductase enzymes on to Th-Fre for comparison.

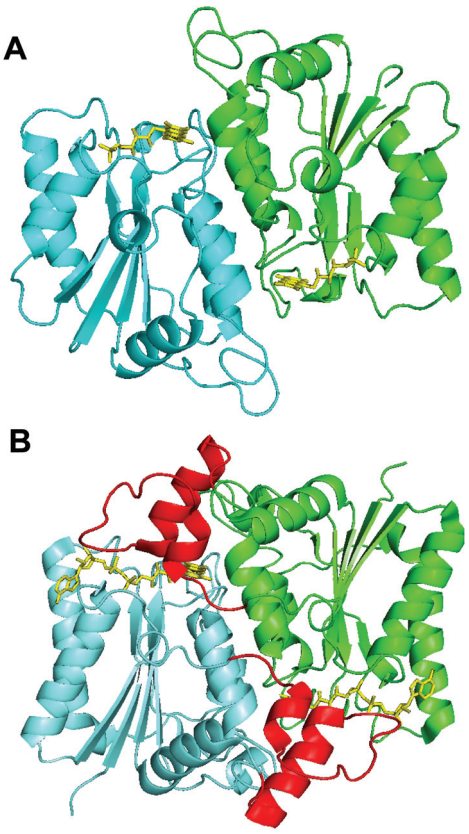


Fig. 8 Crystal structures of Th-Fre and a quinone reductase 2 enzyme. The structural comparison between Th-Fre (A) and human quinone reductase 2 enzyme Nqo2 (B) are also shown. The monomeric units in each protein are shown in different colours (Cyan and Green) while flavin cofactor as yellow sticks. The extra amino acid region at the N terminal of human quinone reductase 2 that forms an extended helix above the flavin domain of opposite monomer is shown in red colour (PDB 2FGL, 2FGI).

Conclusions

In nature, there are many flavin dependent halogenases (Fl-Hal) encoded within various microbial biosynthetic gene clusters, which are involved in the biosynthesis of medicinally important halogenated secondary metabolites. Whilst biosynthetic pathways incorporating these Fl-Hal have been used to produce chlorinated antibiotics on an industrial scale, through fermentation,[35–38] the deployment of these enzymes *in vitro* for chemoenzymatic synthesis remains a major challenge. The exploitation of Fl-Hal for large-scale chemo-enzymatic synthesis of pharmaceuticals and other valuable chemicals is encumbered by the low activity, stability and conversion rates that these enzymes exhibit.

In order to generate more robust and efficient Fl-Hal biocatalysts, we set out to find and characterise thermally stable variants of Fl-Hal. Our search led to the identification of a thermostable tryptophan halogenase (Th-Hal) from a thermophilic and halotolerant strain of *Streptomyces*. In order to develop this enzyme for direct biocatalytic applications, we have combined it with a thermophilic flavin reductase (Th-Fre) enzyme from *Bacillus*. Kinetic characterisation and thermal stability assays showed that these enzymes have a higher stability and reaction rate compared to the other halogenases and reductases that have previously been utilised for *in vitro* biohalogenation. The crystal structures of both of these enzymes are also reported here. Th-Hal showed a similar overall structure to SttH, a mesophilic Fl-Hal with similar substrate scope and regioselectivity. Notably, however, the structural comparisons revealed that there are significant variations of amino acids residues on the surface of the proteins, with Th-Hal possessing many more polar surface residues than SttH, which could influence surface hydrogen bonding, reduce aggregation, and contribute to the higher stability of Th-Hal. These insights could be used in genome mining for more stable, naturally evolved, Fl-Hal and could also provide guidance for

the rational engineering of more effective halogenase enzymes. Establishing structure–activity relationships for Fl-Hal from thermophilic *vs*. mesophilic environments, could enable targeted mutagenesis approaches to improve the properties of the many other less productive Fl-Hal, from mesophilic microorganisms. Finally, the availability of more stable enzymes could allow for the more effective integration of halogenase biocatalysis with chemocatalytic cross-coupling reactions for scaffold diversification and synthesis of high value targets.

Acknowledgements

We acknowledge BBSRC (grant BB/K00199X/1), GlaxoSmithKline and CoEBio3 for financial support. We acknowledge Diamond Light Source for time on beamlines.

Notes and references

1 M. C. Ford and P. S. Ho, *J. Med. Chem.*, 2016, **59**, 1655.
2 P. Jeschke, *Pest Manage. Sci.*, 2010, **66**, 10.
3 M. Hernandes, S. Cavalcanti, D. Moreira, W. Azevedo and C. Leite, *Curr. Drug Targets*, 2010, **11**, 303.
4 K. C. Nicolaou, P. G. Bulger and D. Sarlah, *Angew. Chem., Int. Ed.*, 2005, **44**, 4442.
5 C. Valente, S. Çalimsiz, K. H. Hoi, D. Mallik, M. Sayah and M. G. Organ, *Angew. Chem., Int. Ed.*, 2012, **51**, 3314.
6 F.-S. Han, *Chem. Soc. Rev.*, 2013, **42**, 5270.
7 S. Z. Tasker, E. A. Standley and T. F. Jamison, *Nature*, 2014, **509**, 299.
8 Z. Zou, D. T. Ahneman, L. Chi, J. A. Terrett, A. G. Doyle and D. W. C. MacMillan, *Science*, 2014, **345**, 437.
9 M. C. White, *Synlett*, 2012, 2746.
10 G. K. Dewkar, S. V. Narina and A. Sudalai, *Org. Lett.*, 2003, **5**, 4501.
11 D. C. Braddock, G. Cansell and S. A. Hermitage, *Synlett*, 2004, 461.
12 G. K. S. Prakash, T. Matthew, D. Hoole, P. M. Esteeves, Q. Wang, G. Rasul and G. A. Olah, *J. Am. Chem. Soc.*, 2004, **126**, 15770.
13 Y. Zhang, K. Shibatomi and H. Yamamoto, *Synlett*, 2005, 2837.
14 S. Brown and S. E. O'Connor, *ChemBioChem*, 2015, **16**, 2129.
15 V. Weichold, D. Milbredt and K.-H. Van Pée, *Angew. Chem., Int. Ed.*, 2016, **55**, 6374.
16 J. T. Payne and J. C. Lewis, *Synpacts*, 2014, **25**, 1345.
17 F. H. E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3364.
18 C. S. Neumann, D. G. Fujimori and C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99.
19 A. Butler and M. Sandy, *Nature*, 2009, **460**, 848.
20 K.-H. Van Pée, C. Dong, S. Flecks, J. Naismith, E. P. Patallo and T. Wage, *Adv. Appl. Microbiol.*, 2006, **59**, 127.
21 E. Yeh, L. J. Cole, E. W. Barr, J. M. Bollinger, D. P. Ballou and C. T. Walsh, *Biochemistry*, 2006, **45**, 7904.
22 C. J. Dong, S. Flecks, S. Unversucht, C. Haupt, K.-H. Van Pée and J. H. Naismith, *Science*, 2005, **309**, 2216.
23 X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K.-H. Van Pée and J. H. Naismith, *J. Mol. Biol.*, 2009, **391**, 74.
24 E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan and C. T. Walsh, *Biochemistry*, 2007, **46**, 1284.
25 A. D. Roy, S. Grüschow, N. Cairns and R. J. M. Goss, *J. Am. Chem. Soc.*, 2010, **132**, 12243.
26 L. J. Durak, J. T. Payne and J. C. Lewis, *ACS Catal.*, 2016, **6**, 1451.
27 M. Frese, C. Schepel, H. Minges, H. Voβ, R. Feiner and N. Sewald, *ChemCatChem*, 2016, **8**, 1799.
28 A. Lang, S. Polnick, T. Nicke, P. William, E. P. Patallo, J. H. Naismith and K.-H. Van Pée, *Angew. Chem., Int. Ed.*, 2011, **50**, 2951.
29 S. A. Shepherd, C. Karthikeyan, J. Latham, A.-W. Struck, M. L. Thompson, B. R. K. Menon, M. Q. Styles, C. Levy, D. Leys and J. Micklefield, *Chem. Sci.*, 2015, **6**, 3454.
30 J. T. Payne, C. B. Poor and J. C. Lewis, *Angew. Chem., Int. Ed.*, 2015, **54**, 4226.
31 S. A. Shepherd, B. R. K. Menon, H. Fisk, A.-W. Struck, C. Levy, D. Leys and J. Micklefield, *ChemBioChem*, 2016, **17**, 821.
32 J. Latham, J.-M. Henry, H. H. Sharif, B. R. K. Menon, S. A. Shepherd, M. F. Greaney and J. Micklefield, *Nat. Commun.*, 2016, **7**, 11873.
33 C. B. Poor, M. C. Andorfer and J. C. Lewis, *ChemBioChem*, 2014, **15**, 1286.
34 A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfix and R. Milo, *Biochemistry*, 2011, **50**, 4402.
35 J. J. McIntyre, A. T. Bull and A. W. Bunch, *Biotechnol. Bioeng.*, 1996, **49**, 412.
36 V. V. Dasu and T. Panda, *Bioproc. Bioeng.*, 1999, **21**, 489.
37 H.-M. Jung, S.-Y. Kim, H.-J. Moon, D.-K. Oh and J.-K. Lee, *Appl. Microbiol. Biotechnol.*, 2007, **77**, 789.
38 L. B. Pickens, Y. Tang and Y.-H. Chooi, *Annu. Rev. Chem. Biomol. Eng.*, 2011, **2**, 211.
39 M. Frese and N. Sewald, *Angew. Chem., Int. Ed.*, 2014, **54**, 298.
40 H. H. Liao, *Enzyme Microb. Technol.*, 1993, **15**, 286.
41 H. Zhao and F. H. Arnold, *Prot. Eng.*, 1999, **12**, 47.
42 G. D. Haki and S. K. Rakshit, *Bioresour. Technol.*, 2003, **89**, 17.
43 I. Wu and F. H. Arnold, *Biotechnol. Bioeng.*, 2013, **110**, 1874.
44 X. Chen, B. Zhang, W. Zhang, X. Wu, M. Zhang, T. Chen, G. Liu and P. Dyson, *Genome Announcements*, 2013, **1**, 4.
45 J. Harrison and D. J. Studholme, *Microb. Biotechnol.*, 2014, **7**, 373.
46 T. Ohshiro, Y. Ishii, T. Matsubara, K. Ueda, Y. Izumi, K. Kino and K. Kirimura, *J. Biosci. Bioeng.*, 2005, **100**, 266.
47 S. Takahashi, T. Furuya, Y. Ishii, K. Kino and K. Kirimura, *J. Biosci. Bioeng.*, 2009, **107**, 38.
48 J. Zeng and J. Zhan, *Biotechnol. Lett.*, 2011, **33**, 1607.

49 J. R. Heemstra and C. T. Walsh, *J. Am. Chem. Soc.*, 2008, **130**, 14024.
50 M. Frese, P. H. Guzowska, H. Voβ and N. Sewald, *ChemCatChem*, 2014, **6**, 1270.
51 Z. T. Campbell and T. O. Baldwin, *J. Biol. Chem.*, 2009, **284**, 8322.
52 T. Matsubara, T. Oshiro, Y. Nishina and Y. Izumi, *Appl. Environ. Microbiol.*, 2001, **67**, 1179.
53 X. Zhu, W. D. Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K.-H. Van Pée and J. H. Naismith, *J. Mol. Biol.*, 2010, **391**, 74.
54 E. Bitto, Y. Huang, C. A. Bingman, S. Singh, J. S. Thorson and G. N. Phillis, *Proteins*, 2007, **70**, 289.
55 P. R. Pokkuluri, R. Raffen, L. Dieckman, C. Boogaard, F. J. Stevens and M. Schiffer, *Biophys. J.*, 2002, **82**, 391.
56 S. S. Strickler, A. V. Gribenko, A. V. Gribenko, T. R. Keiffer, J. Tomlinson, T. Reihle, V. V. Loladze and G. I. Makhatadze, *Biochemistry*, 2006, **45**, 2761.
57 M. D. Smith, M. A. Rosenow, M. Wang, J. P. Allen, J. W. Szostak and J. C. Chaput, *PLoS One*, 2007, **2**, 467.
58 C.-H. Chan, C. C. Wilbanks, G. I. Makhatadze and K.-B. Wong, *PLoS One*, 2012, **7**, 30296.
59 K. K. K. Leung and B. H. Shilton, *J. Biol. Chem.*, 2013, **288**, 11242.
60 C. F. Megarity, J. R. E. Gill, M. C. Caraher, I. J. Stratford, K. A. Nolan and D. J. Timson, *FEBS Lett.*, 2014, **588**, 1666.

ChemPubSoc Europe

DOI: 10.1002/cbic.201600471

CHEMBIOCHEM Communications

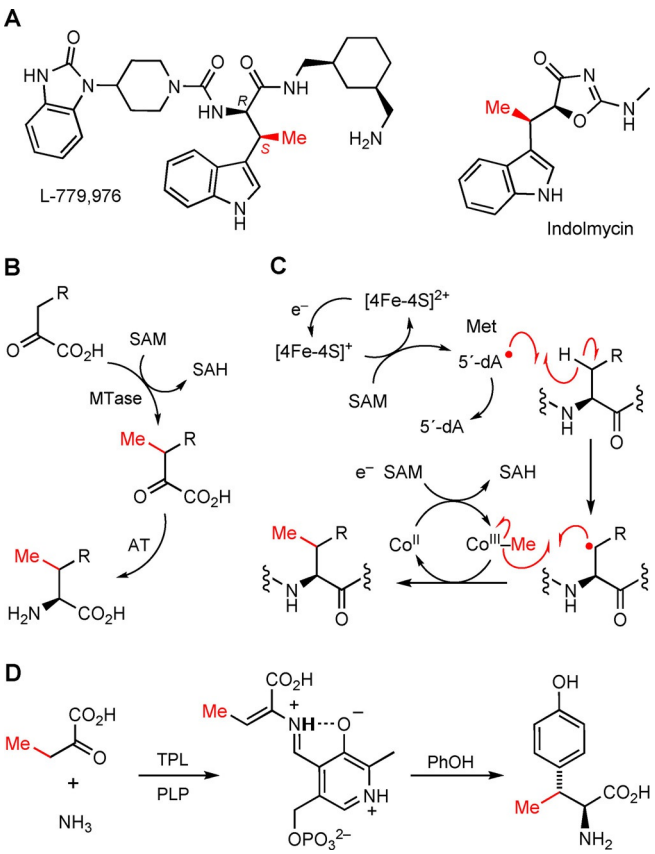
An Engineered Tryptophan Synthase Opens New Enzymatic Pathways to β-Methyltryptophan and Derivatives

Daniel Francis, Michael Winn, Jonathan Latham, Michael F. Greaney, and Jason Micklefield*[a]

β-Methyltryptophans (β-mTrp) are precursors in the biosynthesis of bioactive natural products and are used in the synthesis of peptidomimetic-based therapeutics. Currently β-mTrp is produced by inefficient multistep synthetic methods. Here we demonstrate how an engineered variant of tryptophan synthase from *Salmonella* (*St*TrpS) can catalyse the efficient condensation of L-threonine and various indoles to generate β-mTrp and derivatives in a single step. Although L-serine is the natural substrate for TrpS, targeted mutagenesis of the *St*TrpS active site provided a variant (βL166V) that can better accommodate L-Thr as a substrate. The condensation of L-Thr and indole proceeds with retention of configuration at both α- and β-positions to give (2*S*,3*S*)-β-mTrp. The integration of *St*TrpS (βL166V) with L-amino acid oxidase, halogenase enzymes and palladium chemocatalysts provides access to further D-configured and regioselectively halogenated or arylated β-mTrp derivatives.

β-Methyl-α-amino acids are important building blocks in the synthesis of peptidomimetics and other pharmaceuticals (for examples see Scheme 1 A and Scheme S1 in the Supporting Information).[1] The introduction of a β-methyl substituent into synthetically modified peptides decreases the conformational freedom of the amino acid side chain, reducing the entropic penalty of binding to a biological receptor, which can result in higher affinity and improved biological activity.[2] In light of this, β-branched synthetic peptides have been used to develop improved δ opioid agonists, somatostatin receptor agonists, cholecystokinin B receptor agonists, glucagon receptor antagonists and AT4 receptor antagonists.[1,3] Nature has also adopted the strategy of introducing β-methyl-α-amino acids into peptides, alkaloids and other bioactive natural products (for examples see Schemes 1 A and S2);[4] this presumably also serves to preorganise the amino acid side chain for tighter binding to molecular targets in vivo.

The prominence of β-methyl-α-amino acids in synthetic peptides, natural products and other valuable compounds has resulted in the development of many methods for the synthesis of β-methyl amino acids, including β-lactone ring opening, alkylation of imines using secondary sulfonates, [3,3] sigmatropic rearrangements, the use of chiral auxiliaries and/or kinetic resolutions.[5] Many of these and other approaches require laborious multistep synthetic procedures, use deleterious reagents and often fail to provide enantiomerically pure products. In contrast, there have been relatively few enzymatic approaches used in the preparation of β-methyl-α-amino acids, despite the inherent advantages of enzymes, which include high stereoselectivity and cleaner, more benign aqueous reaction conditions. Previously, we elucidated the biosynthesis of β-methylglutamic acid, which involved the *S*-adenosylmethionine (SAM)-dependent methyltransferase (MTase)-catalysed methylation of α-ke-



Scheme 1. A) Synthetic peptidomimetic L-779,976, a somatostatin agonist and diabetes drug candidate from Merck,[1c,d] and the antibiotic indolmycin from *Streptomyces griseus*.[4g] B) α-Ketoacid MTase and aminotransferase (AT) enzymes produce β-methyl-α-amino acids[6,7] including β-mTrp.[7a] C) Radical SAM MTases are predicted to catalyse β-methylation of amino acid residues in peptide natural products with a 5′-deoxyadenosine (5′-dA) radical likely to abstract the β-hydrogen atoms thereby facillitating subsequent β-methylation with methylcobalamin (Co^{III}-Me).[8] D) TPL-catalysed preparation of (2*S*,3*R*)-β-methyltyrosine.[9]

[a] *D. Francis, Dr. M. Winn, J. Latham, M. F. Greaney, Prof. J. Micklefield*
School of Chemistry and Manchester Institute of Biotechnology
The University of Manchester
131 Princess Street, Manchester M1 7DN (UK)
E-mail: jason.micklefield@manchester.ac.uk

Supporting information for this article can be found under: http://dx.doi.org/10.1002/cbic.201600471.

toglutarate and subsequent transamination by an aminotransferase (AT).[6] Following this, a number of other α-ketoacid MTases were identified, and it transpired that this is a common strategy used in nature for the biosynthesis of β-methyl-α-amino acids (Scheme 1 B).[7] More recently, radical SAM-dependent MTase enzymes have been characterised,[8] and several members of this family have been predicted to β-methylate amino acid residues within peptides including bottromycin and polytheonamide precursors (Scheme 1 C and S2).[4h,i] Despite these insights, neither α-ketoacid MTases or the radical SAM-dependent MTases offer a viable means for the biocatalytic preparation of β-methylated amino acids or peptides. Firstly, SAM is an expensive cofactor, and there are no effective ways available for recycling SAM in vitro. Also, to date, only a few radical SAM MTases have been characterised and these enzymes also require a cobalamin cofactor as well as a [4Fe–4S] cluster that necessitates in vitro reconstitution under anaerobic conditions.[8] An alternative approach to producing β-methyl-α-amino acids is to exploit the promiscuity of enzymes from proteinogenic amino acid metabolism. For example, tyrosine phenol-lyase (TPL) has been used to catalyse the reverse synthesis of β-methyltyrosine from α-ketobutyrate, ammonia and phenol (Scheme 1 D).[9]

In this paper, we demonstrate how an engineered variant of tryptophan synthase (TrpS) can efficiently produce a range of enantiomerically pure (2*S*,3*S*)-L-β-methyltryptophan (β-mTrp) derivatives. Enzyme cascades utilising the TrpS variant with L-amino acid oxidase (L-AAO) and halogenase enzymes provide access to more diverse L- and D-β-mTrp analogues. TrpS is a pyridoxal phosphate (PLP)-dependent enzyme comprising two subunits in an $\alpha_2\beta_2$ tetramer. The α subunit catalyses a retro-aldol cleavage of indole 3-glycerol phosphate to liberate indole, which is then channelled to the β subunit where an aldimine is formed between PLP and L-Ser facilitating dehydration and subsequent condensation with indole to give L-tryptophan (Figure 1 A).[10] As PLP is regenerated during the catalytic cycle, addition of an expensive stoichiometric cofactor is not required, unlike for the SAM-dependent methyltransferases. Moreover, PLP-dependent enzymes have already been demonstrated to have great potential for pharmaceutical synthesis.[11] Previous studies have shown that TrpS from *Salmonella typhimurium* (*St*TrpS ATCC 37845) is promiscuous and will accept a number of indole derivatives.[12] We wished to explore the possibility of utilising threonine as a substrate for *St*TrpS instead of serine to generate β-mTrp **1** (Figure 1 A and Scheme 2). Accordingly *St*TrpS was overproduced in *Escherichia coli* BL21(DE3), and the resulting cell-free extract was incubated with indole and a tenfold excess of L-Thr. The progress of the reaction was followed by HPLC (Figure S3), which revealed the formation of a new product peak, β-mTrp **1**, which reached a maximum conversion of 60(±4)% after 5 hours (Figure S4). No production of β-mTrp **1** was evident in control assays when indole and L-Thr were incubated with standard BL21 lysate lacking *St*TrpS (Figure S3). Subsequent scale up, followed by solvent and solid-phase (C_{18}) extractions led to the isolation of β-mTrp in a 54% yield, thus demonstrating for the first time that the wild-type *St*TrpS can be harnessed to generate a β-methyl amino acid.

Although L-Thr is accepted as a substrate by *St*TrpS, the activity and isolated yields of L-β-mTrp are low compared to the wild-type reaction with L-Ser, which gives over 90% of L-Trp under the same conditions. The published crystal structure(s) of *St*TrpS (PDB ID: 4HPX) show that a leucine residue (βL166) is in close proximity to the β-position of the α-aminoacrylate species.[13] With L-Thr as a substrate, the introduction of a methyl group into the corresponding (*E*)-2-aminobut-2-enoate intermediate could lead to steric hindrance with the side chain of βL166 (Figure 1 B). With this in mind, a *St*TrpS mutant was generated that replaced βL166 with the less bulky valine. Cell-lysate experiments carried out with this new mutant (βL166V) showed improved activity compared to the wild-type, with the conversion of indole to β-mTrp reaching 98(±0.2)% within

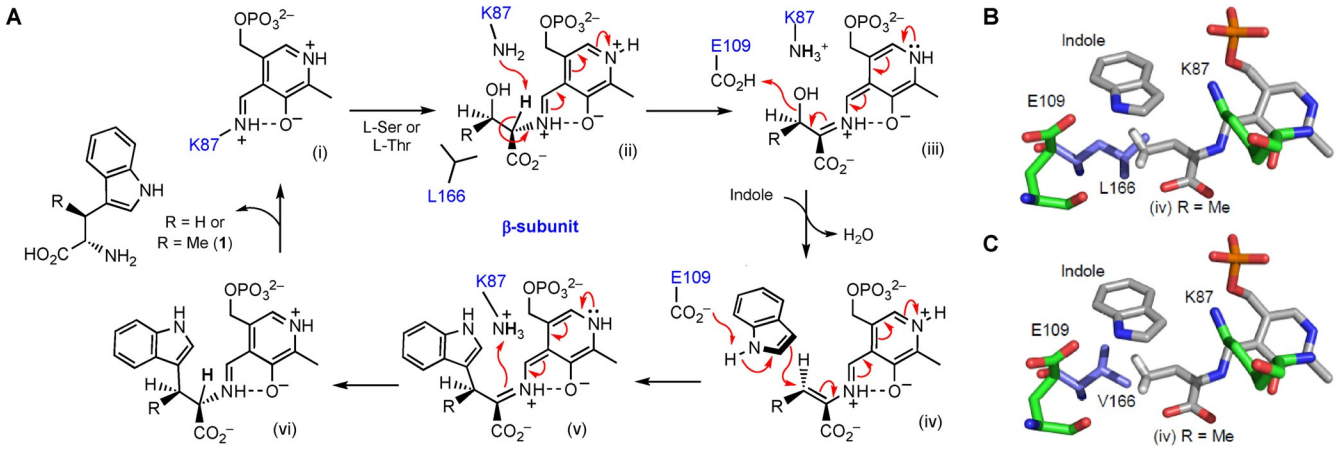
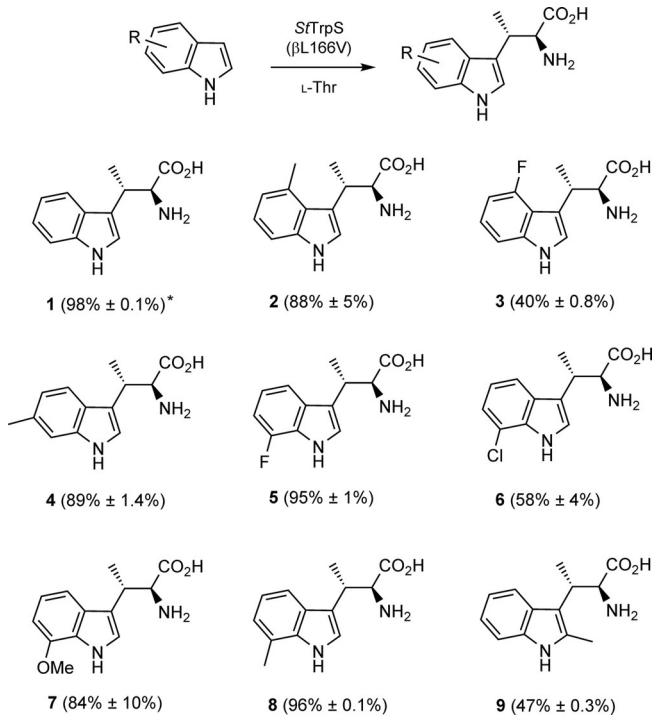


Figure 1. A) TrpS β-subunit reaction mechanism and stereochemical course. L-Ser reacts with the internal aldimine (i) to generate an external aldimine (ii), which is deprotonated by βLys87 to give quinone (iii), which facilitates elimination of the β-OH leading to the α-aminoacrylate species (iv). It is suggested that βGlu109 then deprotonates the indole facilitating conjugate addition with (iv) leading to quinone (v) and the L-Trp external aldimine (vi). B) X-ray crystal structure (PDB ID: 4HPX) of the *St*TrpS β subunit[13] with an (*E*)-2-aminobut-2-enoate intermediate (iv). C) Model of intermediate iv bound to the βL166V TrpS mutant.



Scheme 2. The single-step enzymatic synthesis of β-mTrp derivatives by using the mutant *St*TrpS βL166V (% conversions after 4 h, or *3 h incubation).

3 hours (Figure S4). SDS-PAGE analysis suggests that both wild-type and βL166V lysates contain a similar amount of *St*TrpS protein (Figure S5). However, to facilitate direct comparison of the wild-type *St*TrpS and the βL166V mutant, the β-subunits were expressed as hexahistidine fusion proteins and purified by metal affinity chromatography. The β-subunit was expressed individually to prevent the formation of heterogeneous mixtures of $\alpha_2\beta_2$ and β_2 complexes. As the $\alpha_2\beta_2$ is more active than the β_2[14] such unquantified mixtures could bias comparative activity assays between the wild-type and the mutant. Subsequent kinetic analysis revealed that the activity of the βL166V mutant was ten times higher than that of the wild type (Table 1). Presumably the increase in the size of the active site in the βL166V mutant allows it to better accommodate L-Thr as a substrate (Figure 1C). An additional mutant was generated to further expand the space within the active site, βL166A, but, although this showed a fivefold improvement over the wild-type, it was only half as active as βL166V.

Table 1. Kinetics of *St*TrpS (β_2) enzymes with L-threonine and L-serine.

β_2 (substrate)	K_m [mM]	k_{cat} [min^{-1}]	k_{cat}/K_m [$M^{-1} min^{-1}$]
WT (L-Thr)	210(±22)	0.041(±0.002)	0.19(±0.02)
L166A (L-Thr)	340(±36)	0.21(±0.01)	0.61(±0.07)
L166V (L-Thr)	310(±25)	0.42(±0.02)	1.40(±0.1)
WT (L-Ser)	1.6(±0.2)	21(±0.5)	13000(±6700)
L166A (L-Ser)	52(±6.6)	11(±0.8)	220(±180)
L166V (L-Ser)	1.4(±0.2)	4.9(±0.08)	3600(±600)

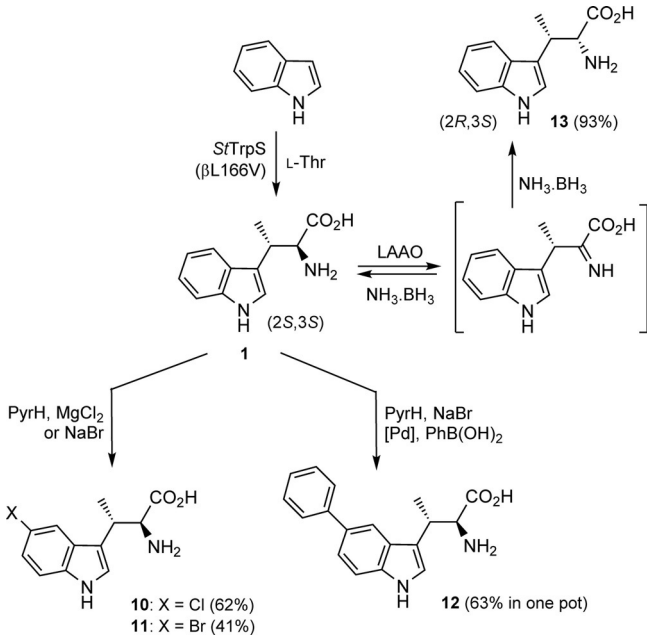
Replacing βL166 with alanine might provide too much space within the active site, possibly allowing either the indole or the (*E*)-2-aminobut-2-enoate intermediate to adopt a suboptimal conformation within the active site, thus leading to less efficient conjugate addition. It is worth noting that, although the measured k_{cat} of βL166V with L-Thr is low, the enzyme functions in a slow but steady fashion, remaining stable and active within a crude cell lysate for several days; this means that good quantities of product can still be isolated. Also the observed rate of the $\alpha_2\beta_2$ tetramer present in the cell lysate will be higher than that of the purified β_2 complex observed here. Analysis of the kinetic parameters with the natural substrate L-serine shows that activity towards serine is reduced in both mutants. These results indicate that the βL166V mutant is a promising biocatalyst for the formation of enantiomerically pure β-mTrp derivatives.

Previous stereochemical studies with TrpS using labelled (2*S*,3*R*)- and (2*S*,3*S*)-[3-^{3}H]serine show that dehydration and the addition of indole to form L-Trp proceed with retention of configuration at C3 as well as at C2 (Figure 1A).[15] Based on this information, combined with analysis of X-ray crystal structures of *St*TrpS[13] (Figure 1B), we predicted that the condensation of L-Thr and indole catalysed by *St*TrpS would proceed with indole addition to the *Re* face of the β-carbon in an (*E*)-2-aminobut-2-enoate intermediate (iv) resulting in (2*S*,3*S*)-β-mTrp (**1**). As anticipated, the configuration of the β-mTrp generated by both the wild-type *St*TrpS and the βL166V mutant was confirmed to be (2*S*,3*S*) by comparison of optical rotation and NMR data (Figure S6), including a $^3J_{H\alpha H\beta}$ coupling constant of 7.4 Hz, with data from ref. [7a]. The enantiomeric purity of **1** was further assessed by using L-amino acid oxidase (LAAO). Incubation of **1** with LAAO resulted in >98% conversion of **1** to the corresponding α-keto acid as determined by HPLC (Figure S7). Conversely, when **1** was incubated with D-amino acid oxidase (DAAO) under identical conditions, no α-keto acid was formed (Figure S7).

To further explore the scope of the βL166V mutant, biotransformations were carried out with a range of halogenated, methylated and methoxylated indoles (Scheme 2). They showed that the mutant could be used to generate β-mTrp derivatives with substituents in either the 2-, 4-, 6- or 7-positions of the indole ring (**1**–**9**). The percentage conversions varied from 40–96% and could be further improved by recycling the unreacted indole for further biotransformation.[12a] The preparation of enantiomerically pure L-β-mTrp derivatives in a single biotransformation offers significant advantages over the existing synthetic routes to these compounds. At least six different synthetic routes to β-mTrp derivatives have been reported (Figure S8)—this emphasises the importance of these compounds—and each synthesis requires between five and eight steps, the use of deleterious reagents, expensive chiral auxiliaries and chiral catalysts, or resolution steps with low overall yields.

Although 2-, 4-, 6- or 7-substituted indoles were accepted by *St*TrpS (βL166V), along with L-Thr, 5-subsituted indoles proved to be very poor substrates for the enzyme. To address this, we explored the possibility of using the flavin-dependent

tryptophan-5-halogenase PyrH to derivatise C5 of the indole moiety of **1**. In addition to regioselectively halogenating L-Trp,[16] PyrH has been shown to be promiscuous and can halogenate a number of other aromatic compounds.[17] The tryptophan halogenases are relatively unstable and inefficient biocatalysts; however, recent studies have shown that the productivity of a tryptophan-7-halogenase (RebH) can be considerably improved through the generation of crosslinked enzyme aggregates (CLEAs).[18] Accordingly a CLEA was prepared that contained PyrH, the flavin reductase Fre (for recycling the $FADH_2$ cofactor) and alcohol dehydrogenase (for NADH regeneration). With this CLEA, **1** could be converted into 5-chloro-β-mTrp (**10**) in a 62% yield by using only $MgCl_2$, O_2 (from air) and isopropanol as stoichiometric reagents (Scheme 3). This



Scheme 3. Diversification of β-mTrp **1** by using halogenase (PyrH)-catalysed C5-chlorination and bromination; C5-arylation by using an integrated one-pot halogenase-Suzuki—Miyaura crosscoupling procedure;[17] and biocatalytic stereoinversion (% isolated yields).

represents a 40% overall yield for the two enzymatic reactions from the indole starting material. Substituting $MgCl_2$ with NaBr as inorganic halide donor allowed the similar preparation of 5-bromo-β-mTrp (**11**) in 41% yield. Recently our laboratory and the Sewald group both showed how halogenase enzymes can be integrated with palladium-catalysed crosscoupling chemistry in one-pot reactions so as to affect the regioselective arylation or alkenylation of CH positions on aromatic scaffolds.[19] With this approach, we were able to directly C5-arylate **1** to give 5-phenyl-β-mTrp (**12**) in a one-pot reaction with 63% yield. PyrH-CLEA was used to generate the intermediate aryl bromide (**11**), and sodium 2′-dicyclohexylphosphino-2,6-dimethoxy-1,1′-biphenyl-3-sulfonate (sSPhoS) and Na_2PdCl_4 were used to catalyse crosscoupling with phenyl boronic acid. These results, coupled with our previous studies,[19a] indicate that combining *St*TrpS (βL166V) with halogenases and transition-metal catalysis can open the way to more highly modified β-mTrp derivatives that would be difficult to prepare directly from a functionalised indole precursor using TrpS.

Although the exquisite stereoselectivity of TrpS is a major advantage, access to other diastereoisomers would be desirable. For example the D-configured epimer of **1**, (2*R*,3*S*)-β-mTrp (**13**), is present in peptidomimetic drug candidates for diabetes, such as L-779,976 (Scheme 1 A).[1c,d] Given that we found (2*S*,3*S*)-β-mTrp to be a substrate for LAAO, we envisaged effecting stereoinversion at the α-position to give (2*R*,3*S*)-β-mTrp (**12**) by nonselective reduction of the imine intermediate formed from the LAAO oxidation of **1**. Such cyclic oxidation–reduction procedures have been used successfully in the deracemisation or epimerisation of other α-amino acid substrates.[20] Accordingly, **1** was incubated with LAAO in the presence of an excess of ammonia-borane under the established conditions[20] to give the 2*R*,3*S* epimer **13** in 93% yield. The ^{1}H NMR spectrum of **13** clearly indicates the expected change in chemical shifts of and coupling constant between the α and β protons (Figure S6). The NMR and other analytical data are also in agreement with the literature.[7a] Furthermore, incubation of **13** with DAAO led to complete oxidation, whereas no reaction was observed with LAAO (Figure S7). The overall yield for the two-step enzymatic preparation of (2*R*,3*S*)-β-mTrp from indole is about 66%.

In order to generate β-mTrp with the opposite configuration at C3, we envisaged utilising L-*allo*-threonine (2*S*,3*S*-Thr) as a substrate for TrpS. Based on the structure and mechanism of *St*TrpS (Figure 1), L-*allo*-Thr would, if accepted, be predicted to generate the (*Z*)- rather than (*E*)-2-aminobut-2-enoate intermediate (iv), leading to (2*S*,3*R*)-β-mTrp. However, we found that L-*allo*-Thr is not a substrate for *St*TrpS or the mutants we prepared. Given that L-*allo*-threonine aldolase can be used to produce L-*allo*-threonine from glycine and acetaldehyde,[21] engineering TrpS to accept L-*allo*-threonine as an alternative amino acid substrate is an attractive future goal.

In summary we have demonstrated that tryptophan synthase (*St*TrpS) can utilise threonine, along with indole, in the preparation of enantiopure (2*S*,3*S*)-β-mTrp. Rational mutagenesis of the *St*TrpS β-subunit (Leu166 to valine), which is likely to better accommodate the larger L-Thr substrate, provided an enzyme with tenfold improved activity. The improved variant (βL166V) also accepts a range of substituted indoles and can be used to prepare 2-, 4-, 6- and 7-functionalised (2*S*,3*S*)-β-mTrp. Although 5-substitued indoles proved to be poor substrates for *St*TrpS (βL166V), the halogenase PyrH can be used to generate 5-chloro or 5-bromo-(2*S*,3*S*)-β-mTrp, with the more reactive bromo derivative enabling further derivatisation by using crosscoupling chemistry in a one-pot, integrated approach.[19] Finally, the use of LAAO in the presence of a nonselective reductant enables complete stereoinversion of (2*S*,3*S*)-β-mTrp to the 2*R*,3*S* diastereomer in near quantitative yields. *St*TrpS (βL166V) alone or in combination with a halogenase or LAAO provides a convenient biocatalytic route, which offers significant advantages over synthetic procedures, to a range of functionalised β-mTrp derivatives, which have proven to be useful building blocks for drug synthesis.[1–4]

During revision of this manuscript, we became aware of a recent publication describing an evolved mutant of the tryptophan synthase β-subunit from the thermophile *Pyrococcus furiosus* (*Pf*TrpB) that can also be used to produce β-mTrp.[22] Kinetic parameters were not determined in this other study; this precludes a direct quantitative comparison between the *Pf*TrpB and *St*TrpS variants. However, unlike the *Pf*TrpB variant, which requires high temperature (75 °C) for optimal catalytic activity, the *St*TrpS mutant described here efficiently catalyses condensation of indoles with L-Thr at ambient temperature, which is preferred for biocatalytic processes.

Acknowledgements

We thank the BBSRC for a Ph.D. studentship award to D.F. (allied to project BB/K002341/1 cofunded by Syngenta) and the University of Manchester for a Ph.D. scholarship award to J.L.

Keywords: biocatalysis · enzyme cascades · methylamino acids · peptidomimetics · tryptophan synthase

[1] a) C. Haskell-Luevano, L. W. Boteju, H. Miwa, C. Dickinson, I. Gantz, T. Yamada, M. E. Hadley, V. J. Hruby, *J. Med. Chem.* **1995**, *38*, 4720 – 4729; b) Z. Huang, Y. B. He, K. Raynor, M. Tallent, T. Reisine, M. Goodman, *J. Am. Chem. Soc.* **1992**, *114*, 9390 – 9401; c) S. P. Rohrer, E. T. Birzin, R. T. Mosley, S. C. Berk, S. M. Hutchins, D. M. Shen, Y. Xiong, E. C. Hayes, R. M. Parmar, F. Foor, S. W. Mitra, S. J. Degrado, M. Shu, J. M. Klopp, S. J. Cai, A. Blake, W. W. Chan, A. Pasternak, L. Yang, A. A. Patchett, R. G. Smith, K. T. Chapman, J. M. Schaeffer, *Science* **1998**, *282*, 737 – 740; d) L. Yang, S. C. Berk, S. P. Rohrer, R. T. Mosley, L. Guo, D. J. Underwood, B. H. Arison, E. T. Birzin, E. C. Hayes, S. W. Mitra, R. M. Parmar, K. Cheng, T. J. Wu, B. S. Butler, F. Foor, A. Pasternak, Y. Pan, M. Silva, R. M. Freidinger, R. G. Smith, et al., *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10836 – 10841.

[2] a) V. W. Cornish, M. I. Kaplan, D. L. Veenstra, P. A. Kollman, P. G. Schultz, *Biochemistry* **1994**, *33*, 12022 – 12031; b) V. J. Hruby, *J. Med. Chem.* **2003**, *46*, 4215 – 4231; c) C. Haskell-Luevano, K. Toth, L. Boteju, C. Job, A. M. de L. Castrucci, M. E. Hadley, V. J. Hruby, *J. Med. Chem.* **1997**, *40*, 2740 – 2749.

[3] a) D. Tourwé, E. Mannekens, T. N. T. Diem, P. Verheyden, H. Jaspers, G. Tóth, A. Péter, I. Kertész, G. Török, N. N. Chung, P. W. Schiller, *J. Med. Chem.* **1998**, *41*, 5167 – 5176; b) L. W. Boteju, G. V. Nikiforovich, C. Haskell-Luevano, S.-N. Fang, T. Zalewska, D. Stropova, H. I. Yamamura, V. J. Hruby, *J. Med. Chem.* **1996**, *39*, 4120 – 4124; c) B. Y. Azizeh, M. D. Shenderovich, D. Trivedi, G. Li, N. S. Sturm, V. J. Hruby, *J. Med. Chem.* **1996**, *39*, 2449 – 2455; d) A. Lukaszuk, H. Demaegdt, I. Van den Eynde, P. Vanderheyden, G. Vauquelin, D. Tourwé, *J. Pept. Sci.* **2011**, *17*, 545 – 553.

[4] a) J. C. Sheehan, D. Mania, S. Nakamura, J. A. Stock, K. Maeda, *J. Am. Chem. Soc.* **1968**, *90*, 462 – 470; b) T. W. Doyle, D. M. Balitz, R. E. Grulich, D. E. Nettleton, S. J. Gould, C.-h. Tann, A. E. Moews, *Tetrahedron Lett.* **1981**, *22*, 4595 – 4598; c) W. Balk-Bindseil, E. Helmke, H. Weyland, H. Laatsch, *Liebigs Ann.* **1995**, *1995*, 1291 – 1294; d) Z. Hojati, C. Milne, B. Harvey, L. Gordon, M. Borg, F. Flett, B. Wilkinson, P. J. Sidebottom, B. A. M. Rudd, M. A. Hayes, C. P. Smith, J. Micklefield, *Chem. Biol.* **2002**, *9*, 1175 – 1187; e) H. He, R. T. Williamson, B. Shen, E. I. Graziani, H. Y. Yang, S. M. Sakya, P. J. Petersen, G. T. Carter, *J. Am. Chem. Soc.* **2002**, *124*, 9729 – 9736; f) E. Rössner, A. Zeeck, W. A. König, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 64 – 65; *Angew. Chem.* **1990**, *102*, 84 – 85; g) Y. Hou, M. D. B. Tianero, J. C. Kwan, T. P. Wyche, C. R. Michel, G. A. Ellis, E. Vazquez-Rivera, D. R. Braun, W. E. Rose, E. W. Schmidt, T. S. Bugni, *Org. Lett.* **2012**, *14*, 5050 – 5053; h) M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H.-G. Sahl, S. Matsunaga, J. Piel, *Science* **2012**, *338*, 387 – 390; i) Y.-L. Du, L. M. Alkhalaf, K. S. Ryan, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 2717 – 2722.

[5] a) A. Armstrong, S. P. Geldart, C. R. Jenner, J. N. Scutt, *J. Org. Chem.* **2007**, *72*, 8091 – 8094; b) S. Lou, G. M. McKenna, S. A. Tymonko, A. Ramirez, T. Benkovics, D. A. Conlon, F. González-Bobes, *Org. Lett.* **2015**, *17*, 5000 – 5003; c) K. Sakaguchi, M. Yamamoto, T. Kawamoto, T. Yamada, T. Shinada, K. Shimamoto, Y. Ohfune, *Tetrahedron Lett.* **2004**, *45*, 5869 – 5872; d) R. Dharanipragada, K. VanHulle, A. Bannister, S. Bear, L. Kennedy, V. J. Hruby, *Tetrahedron* **1992**, *48*, 4733 – 4748; e) T. Ooi, D. Kato, K. Inamura, K. Ohmatsu, K. Maruoka, *Org. Lett.* **2007**, *9*, 3945 – 3948; f) M. Alías, M. P. López, C. Cativiela, *Tetrahedron* **2004**, *60*, 885 – 891.

[6] a) C. Milne, A. Powell, J. Jim, M. Al Nakeeb, C. P. Smith, J. Micklefield, *J. Am. Chem. Soc.* **2006**, *128*, 11250 – 11259; b) C. Mahlert, F. Kopp, J. Thirlway, J. Micklefield, M. A. Marahiel, *J. Am. Chem. Soc.* **2007**, *129*, 12011 – 12018.

[7] a) Y. Zou, Q. Fang, H. Yin, Z. Liang, D. Kong, L. Bai, Z. Deng, S. Lin, *Angew. Chem. Int. Ed.* **2013**, *52*, 12951 – 12955; *Angew. Chem.* **2013**, *125*, 13189 – 13193; b) S. D. Braun, J. Hofmann, A. Wensing, M. S. Ullrich, H. Weingart, B. Volksch, D. Spiteller, *Appl. Environ. Microbiol.* **2010**, *76*, 2500 – 2508; c) Y.-T. Huang, S.-Y. Lyu, P.-H. Chuang, N.-S. Hsu, Y.-S. Li, H.-C. Chan, C.-J. Huang, Y.-C. Liu, C.-J. Wu, W.-B. Yang, T.-L. Li, *ChemBioChem* **2009**, *10*, 2480 – 2487.

[8] a) M. R. Bauerle, E. L. Schwalm, S. J. Booker, *J. Biol. Chem.* **2015**, *290*, 3995 – 4002; b) S. Pierre, A. Guillot, A. Benjdia, C. Sandström, P. Langella, O. Berteau, *Nat. Chem. Biol.* **2012**, *8*, 957 – 959; c) H. J. Kim, R. M. McCarty, Y. Ogasawara, Y.-n. Liu, S. O. Mansoorabadi, J. LeVieux, H.-w. Liu, *J. Am. Chem. Soc.* **2013**, *135*, 8093 – 8096.

[9] K. Kim, P. A. Cole, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1205 – 1208.

[10] M. F. Dunn, D. Niks, H. Ngo, T. R. M. Barends, I. Schlichting, *Trends Biochem. Sci.* **2008**, *33*, 254 – 264.

[11] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305 – 309.

[12] a) R. J. M. Goss, P. L. A. Newill, *Chem. Commun.* **2006**, 4924 – 4925; b) R. S. Phillips, *Tetrahedron: Asymmetry* **2004**, *15*, 2787 – 2792.

[13] D. Niks, E. Hilario, A. Dierkers, H. Ngo, D. Borchardt, T. J. Neubauer, L. Fan, L. J. Mueller, M. F. Dunn, *Biochemistry* **2013**, *52*, 6396 – 6411.

[14] a) E. J. Faeder, G. G. Hammes, *Biochemistry* **1971**, *10*, 1041 – 1045; b) A. N. Lane, K. Kirschner, *Eur. J. Biochem.* **1983**, *129*, 571 – 582.

[15] M. D. Tsai, E. Schleicher, R. Potts, G. E. Skye, H. G. Floss, *J. Biol. Chem.* **1978**, *253*, 5344 – 5349.

[16] S. Zehner, A. Kotzsch, B. Bister, R. D. Süssmuth, C. Méndez, J. A. Salas, K.-H. van Pée, *Chem. Biol.* **2005**, *12*, 445 – 452.

[17] S. A. Shepherd, C. Karthikeyan, J. Latham, A.-W. Struck, M. L. Thompson, B. R. K. Menon, M. Q. Styles, C. Levy, D. Leys, J. Micklefield, *Chem. Sci.* **2015**, *6*, 3454 – 3460.

[18] M. Frese, N. Sewald, *Angew. Chem. Int. Ed.* **2015**, *54*, 298 – 301; *Angew. Chem.* **2015**, *127*, 302 – 305.

[19] a) J. Latham, J.-M. Henry, H. H. Sharif, B. R. K. Menon, S. A. Shepherd, M. F. Greaney, J. Micklefield, *Nat. Commun.* **2016**, *7*, 11873; b) M. Frese, C. Schnepel, H. Minges, H. Voß, R. Feiner, N. Sewald, *ChemCatChem* **2016**, *8*, 1799 – 1803.

[20] G. J. Roff, R. C. Lloyd, N. J. Turner, *J. Am. Chem. Soc.* **2004**, *126*, 4098 – 4099.

[21] a) J. Q. Liu, T. Dairi, M. Kataoka, S. Shimizu, H. Yamada, *J. Bacteriol.* **1997**, *179*, 3555 – 3560; b) N. Dückers, K. Baer, S. Simon, H. Gröger, W. Hummel, *Appl. Microbiol. Biotechnol.* **2010**, *88*, 409 – 424.

[22] M. Herger, P. van Roye, D. K. Romney, S. Brinkmann-Chen, A. R. Buller, F. H. Arnold, *J. Am. Chem. Soc.* **2016**, *138*, 8388 – 8391.

Manuscript received: November 22, 2016

Accepted Article published: December 22, 2016

Final Article published: January 20, 2017

CHEMICAL REVIEWS

Review

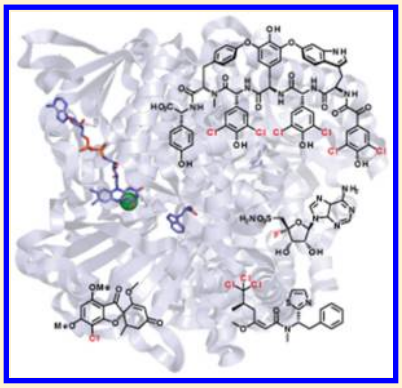
pubs.acs.org/CR

Development of Halogenase Enzymes for Use in Synthesis

Jonathan Latham, Eileen Brandenburger, Sarah A. Shepherd, Binuraj R. K. Menon, and Jason Micklefield*

School of Chemistry and Manchester Institute of Biotechnology, The University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom

ABSTRACT: Nature has evolved halogenase enzymes to regioselectively halogenate a diverse range of biosynthetic precursors, with the halogens introduced often having a profound effect on the biological activity of the resulting natural products. Synthetic endeavors to create non-natural bioactive small molecules for pharmaceutical and agrochemical applications have also arrived at a similar conclusion: halogens can dramatically improve the properties of organic molecules for selective modulation of biological targets in vivo. Consequently, a high proportion of pharmaceuticals and agrochemicals on the market today possess halogens. Halogenated organic compounds are also common intermediates in synthesis and are particularly valuable in metal-catalyzed cross-coupling reactions. Despite the potential utility of organohalogens, traditional nonenzymatic halogenation chemistry utilizes deleterious reagents and often lacks regiocontrol. Reliable, facile, and cleaner methods for the regioselective halogenation of organic compounds are therefore essential in the development of economical and environmentally friendly industrial processes. A potential avenue toward such methods is the use of halogenase enzymes, responsible for the biosynthesis of halogenated natural products, as biocatalysts. This Review will discuss advances in developing halogenases for biocatalysis, potential untapped sources of such biocatalysts and how further optimization of these enzymes is required to achieve the goal of industrial scale biohalogenation.



CONTENTS

Special Issue: Biocatalysis in Industry

Received: January 16, 2017

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

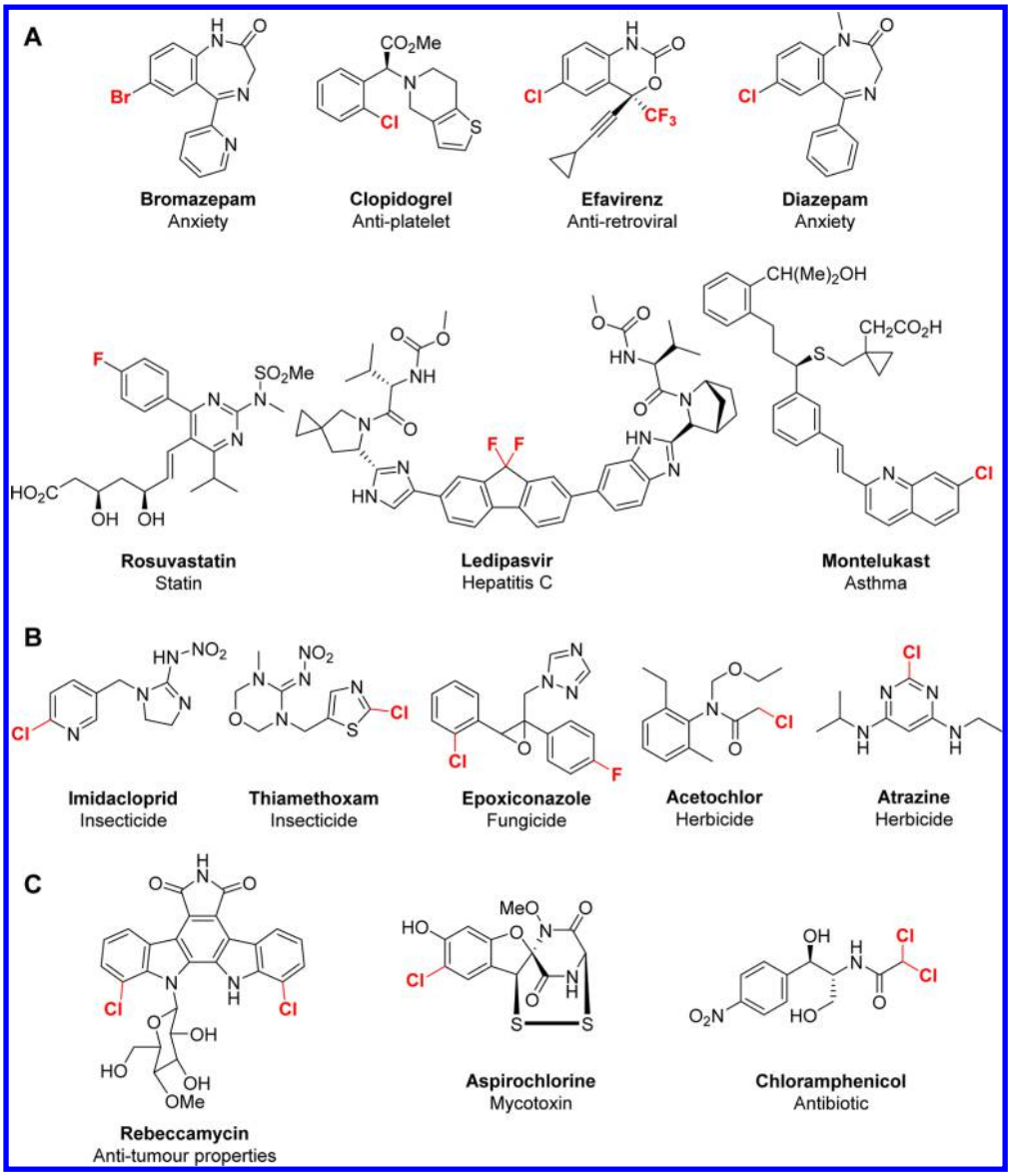


Figure 1. (A) Examples of top-selling halogen-containing pharmaceuticals. (B) Halogenated agrochemicals. (C) Halogenated natural products with known potent bioactivity.

1. INTRODUCTION

Organohalogen moieties are present in many pharmaceutical and agrochemical products, as well as other valuable materials, and are also widely used in all sectors of the chemical industry as synthetic intermediates. Transition metal-catalyzed cross-coupling reactions have become indispensable tools for the synthesis of complex molecules because of the multitude of C−C, C−F, C−N, and other C−heteroatom couplings that are possible.[1−9] Many of these reactions utilize organohalogens because of the ability to metalate C−X bonds, and consequently halogenated compounds are now ubiquitous intermediates in organic synthesis.[3,6] Additionally, the introduction of a halogen atom can have a profound effect on the bioactivity and physicochemical properties of small molecules. This effect has been exploited in medicinal chemistry, with a large proportion of all drugs in clinical trials or on the market containing halogen atoms (Figure 1A).[10−14] In the case of the antibiotic vancomycin, for example, it has been demonstrated that the halogen substituents are important for antimicrobial activity, with dechlorovancomycin derivatives exhibiting significantly reduced binding affinity for the peptidoglycan biological target.[15] The privileged effect of the halogen upon biological activity has also transcended to the design of agrochemicals, where many of the bestselling herbicides, pesticides, and insecticides are halogen-containing (Figure 1B).[16,17] Organohalogen compounds have also been found to have desirable properties in polymers and therefore are receiving increased attention for the next generations of materials.[18−20]

The impact of halogens upon bioactivity and bioavailability was thought to be due solely to modulation of lipophilicity and nonspecific hydrophobic interactions with protein targets. More recently, however, it has been shown that carbon−halogen bonds can form directional intermolecular interactions with proteins, called halogen bonds.[10,11,21−25] These come about because of the electron-deficient "sigma-hole" of the halogen in a C−X bond, which allows interaction of halogens with the lone pairs of heteroatoms like N, O and S in protein targets in a manner analogous to hydrogen bonding.[23−28] As such, the

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

incorporation of halogen atoms during medicinal chemistry efforts is a well-established practice as it can allow the introduction of additional ligand−target interactions without the need to significantly alter other interactions with the target. The C−F bond in particular is of long-standing importance in the development of pharmaceuticals; indeed fluorine is the most prevalent halogen found in drugs, and a number of the best-selling small molecule pharmaceutical compounds contain fluorine.[12−14] This effect is due in part to fluorine's similar size to hydrogen but significantly increased electronegativity. This combination facilitates the replacement of metabolically labile C−H bonds with C−F, without significantly disturbing interactions with the biological target, in addition to the formation of strong halogen bonds and significant changes in lipophilicity.

As with many synthetic endeavors, the development and application of organohalogen compounds has to a large extent been inspired by nature. A plethora of halogenated natural products have been isolated from a diverse range of microorganisms, many of which possess potent antimicrobial and antitumor activities, among others (Figure 1C).[29−31] In marine organisms, organobromine compounds prevail, while chlorinated compounds are found mostly from terrestrial sources and naturally occurring organoiodine is relatively rare. Elucidation of the biosynthetic pathways responsible for the production of these compounds has revealed a number of enzymes capable of halogenating both aliphatic and aromatic carbons on either protein-tethered or free-standing substrates.[32−34] These enzymes have received much interest since their first discoveries, and now a number of fundamentally different classes of halogenases are known, categorized on the basis of the mechanism by which they generate and utilize activated halide.

Given the critical importance of organohalogens, methods for the facile and selective installation of halogen substituents are necessary. Many of the traditional methods, especially for aromatic halogenation, require electrophilic halogen sources that are typically toxic and harmful to the environment.[35] The selectivity of these methods is usually poor, affording mixtures of halogenated regioisomers, limiting material efficiency and requiring careful disposal of persistent environmental pollutants.[36] More recent synthetic methods address some of these issues,[37] but often confer selectivity through subtle differences in the acidity of certain C−H bonds or require directing groups to control the position of functionalization, which can limit the substrate scope and number of regiochemistries accessible.[38−42] The use of halogenase enzymes from secondary metabolism for the installation of halogens into both natural and synthetic scaffolds is therefore an attractive prospect that has potential to develop greener and more selective processes for the production of important halogenated compounds. The use of enzymes in aqueous reaction media at ambient temperatures, in addition to reducing the waste generated from production of unwanted regioisomers, could have a profound effect on all sectors of the chemical industry. This Review will focus on the work, reported to date, that aims to develop halogenase enzymes toward reliable biocatalysts for industrial biotransformations, in addition to work on closely related monooxygenase enzymes that may aid their development.

2. HALOPEROXIDASES

Haloperoxidases were among the first halogenase enzymes to be discovered and characterized in vitro. Enzymes of this class generate hypohalous acid as the halogenating agent from hydrogen peroxide and halide ions. Chlorination, bromination, and iodination can be observed with this type of enzyme; however, fluorination has yet to be detected, most likely due to the high redox potential and tight aqueous solvation of fluoride. These enzymes are classified by the most electronegative halide they can activate. For example, a chloroperoxidase can activate chloride, bromide, and iodide, whereas a bromoperoxidase can activate bromide and iodide but not chloride, although this does not always correlate with the rate of halogenation.[43] Haloperoxidases can be further classified on the basis of the cofactor that they utilize, and this will be discussed in the sections that follow.

Generally, haloperoxidase activity can be screened for using substrates whose spectral properties change upon halogenation. Monochlorodimedone (**1**) has been used in numerous assays to assess both bromination and chlorination activity following the change in UV absorbance maxima from 290 to 270 nm (Figure 2).[43−46] This assay does have its drawbacks with both false

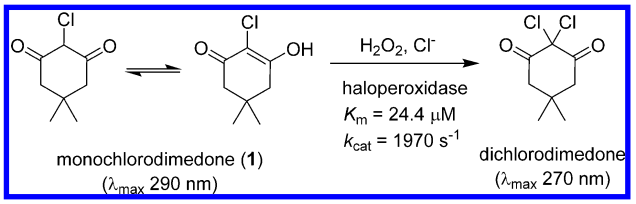


Figure 2. Monochlorodimedone-based assay for detecting haloperoxidase activity. The kinetic parameters displayed are for the chlorination of **1** by CPO from *Caldariomyces fumago*.[55]

negatives and false positives reported for specific vanadium-dependent chloroperoxidases.[47−49] Other assays utilize phenol red and form bromophenol blue[50−53] or monitor the formation of triiodide.[43,54]

2.1. Heme-Iron-Dependent Haloperoxidases

The first halogenating enzyme to be identified was the chloroperoxidase from the fungus *Caldariomyces fumago* (*Leptoxyphium fumago*), which was discovered during the investigation of caldariomycin biosynthesis.[46,56,57] The *C. fumago* chloroperoxidase dihalogenates 1,3-cyclopentanedione leading to caldariomycin; however, it is also capable of halogenating other electron-rich substrates.[58]

Crystallography revealed the *C. fumago* chloroperoxidase to be a 42 kDa glycoprotein with a heme cofactor and unusual tertiary structure dominated by 8 helices (Figure 3A).[59−62] The enzyme shares features of both cytochrome P450s and peroxidases.[59] The proximal heme ligand is cysteine (Cys-29), as in cytochrome P450s, but the distal side of the heme is polar like peroxidases (Figure 3B). A further difference between the traditional peroxidases and chloroperoxidase can be noted in the general acid−base catalytic group, which is a glutamate in chloroperoxidase, rather than a histidine, which is found in the peroxidases. Glutamate is thought to participate in peroxide O−O bond cleavage. When this glutamate residue is mutated to histidine, chlorination activity is severely reduced.[63]

The catalytic cycle of chloroperoxidase is initiated through the binding of peroxide to the axial position of the Fe^{III} complex (**2**), followed by the heterolytic O−O bond cleavage leading to the formation of the Fe^{IV}−oxo species (**3**), which is known as compound I. The halide ion is then oxidized by compound I, generating hypohalous acid and regenerating the

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

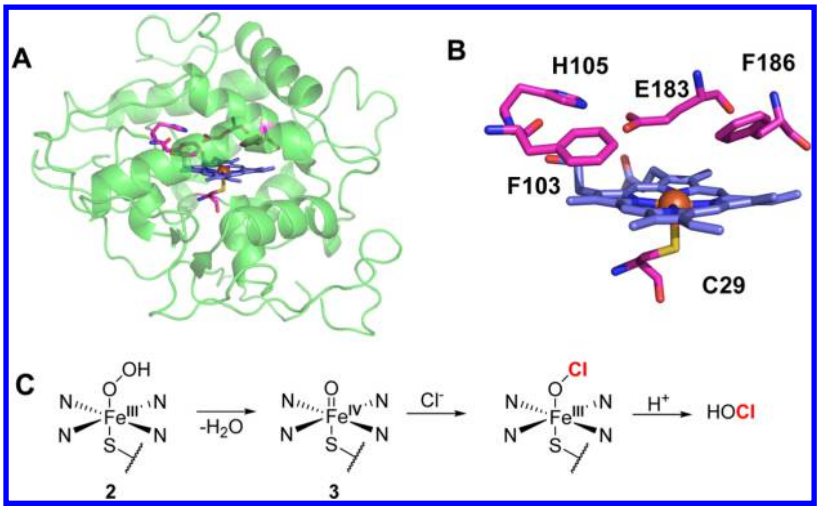


Figure 3. (A) Overall structure of chloroperoxidase from *C. fumago* (PDB 1CPO). (B) Active site of CPO from *C. fumago* (PDB 1CPO). Blue sticks represent heme, and orange spheres represent Fe. (C) Proposed catalytic cycle of hypochlorus acid generation by Fe^{II}-heme halogenases.

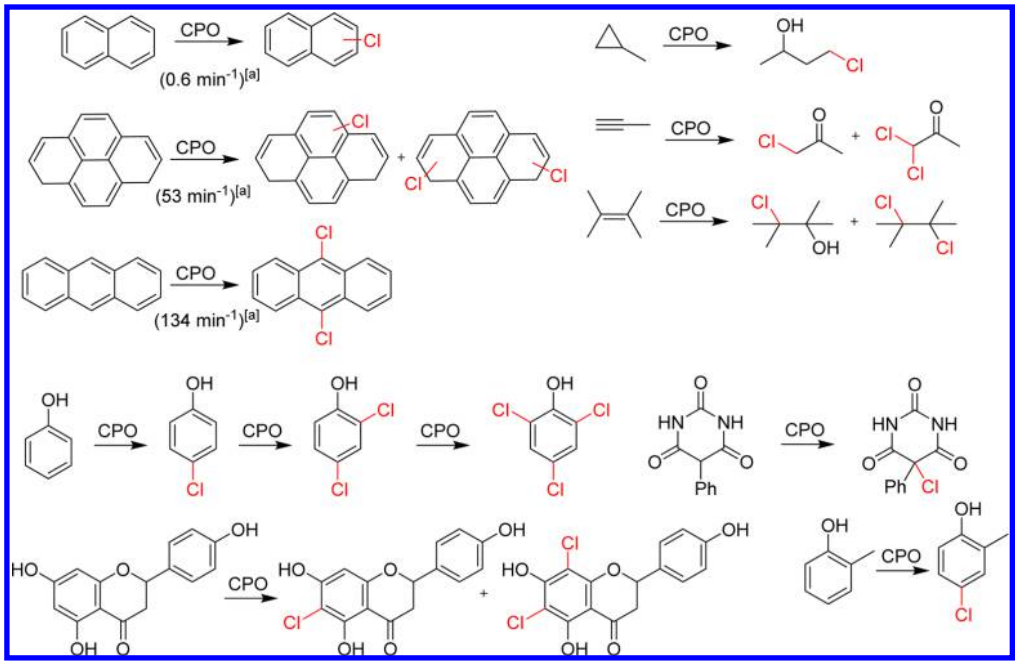


Figure 4. Fe-heme halogenase-catalyzed biotransformations. [a] Specific activity of substrate halogenation by CPO.[66]

Fe^{III} resting state (Figure 3C).[60,64,65] There has been considerable debate about how hypohalous acid then reacts with substrates; however, due to the lack of selectivity conferred by the enzyme, it is likely that hypohalous acid is freely diffusible and can therefore react with many substrates that are susceptible to electrophilic attack (Figure 4).[66−72]

2.2. Vanadium-Dependent Haloperoxidases

A number of vanadium-dependent haloperoxidases are also known, the first of which to be identified was from the brown algae *Ascophyllum nodosum*.[73,74] Vanadium-dependent haloperoxidases are mainly found in marine organisms with bromoperoxidases predominating.[45,75−79] Vanadium-dependent haloperoxidases have been found in algae, fungi, lichens, and more recently bacteria.[48,73,80−83] These enzymes can be recombinantly expressed in *E. coli* and yeast, and the bromoperoxidase from *A. nodosum* is commercially available.[50,77,84−86] Vanadium-dependent haloperoxidases have been subject to mutagenesis to further improve their thermal stability as well as solvent and pH tolerance.[50,87,88]

Structurally, the vanadium-dependent haloperoxidases are similar to the acid phosphatases; therefore, it is not surprising that these enzymes can bind phosphate in place of vanadate and exhibit phosphatase activity.[89] Phosphate buffers can be therefore useful for crystallization of the enzymes, but should be avoided in assays as phosphate can act as a competitive inhibitor.[90,91] The vanadate ion is bound at the bottom of a wide funnel, which is between 15 and 20 Å in length at the core of two four helix bundles. A conserved histidine bonds to the vanadium, while vanadate oxygens are coordinated to basic residues (one lysine and two arginines) (Figure 5B).[92,93] While the active site is highly conserved among the vanadium-dependent haloperoxidases, their quaternary state is highly variable. The vanadium-dependent chloroperoxidase from *Curvularia inaequalis* is monomeric, whereas the vanadium-dependent bromoperoxidases from *A. nodosum* and *Corallina*

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

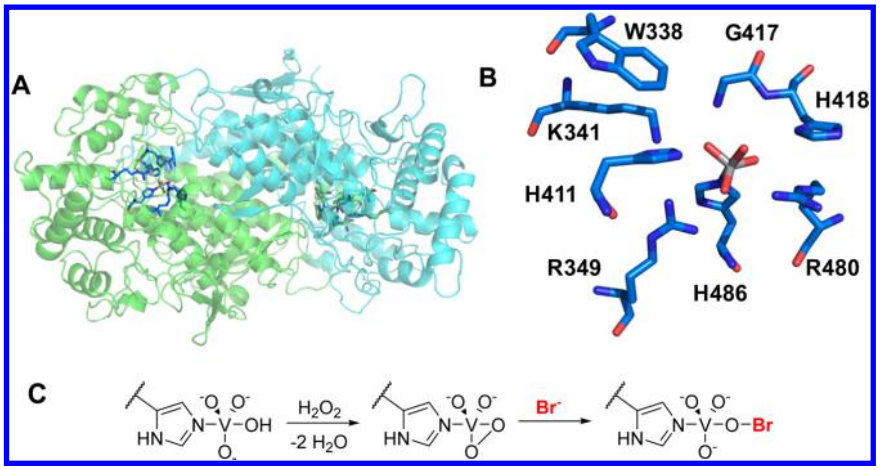


Figure 5. (A) Overall structure of C-BPO from *A. nodosum* (PDB 1QI9). (B) Active site of C-BPO from *A. nodosum* (PDB 1QI9). (C) Proposed catalytic cycle of hypobromous acid generation by vanadium-dependent haloperoxidases.

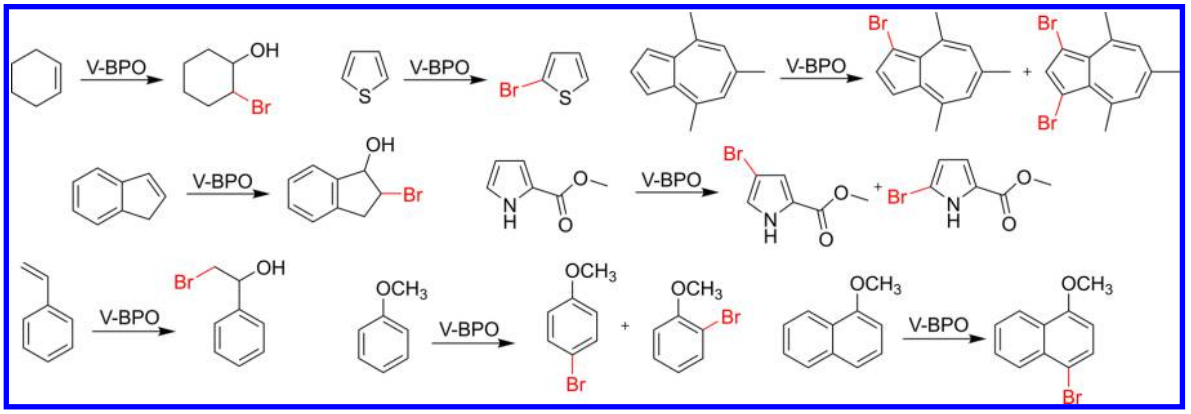


Figure 6. Typical substrates from vanadium-dependent haloperoxidases.

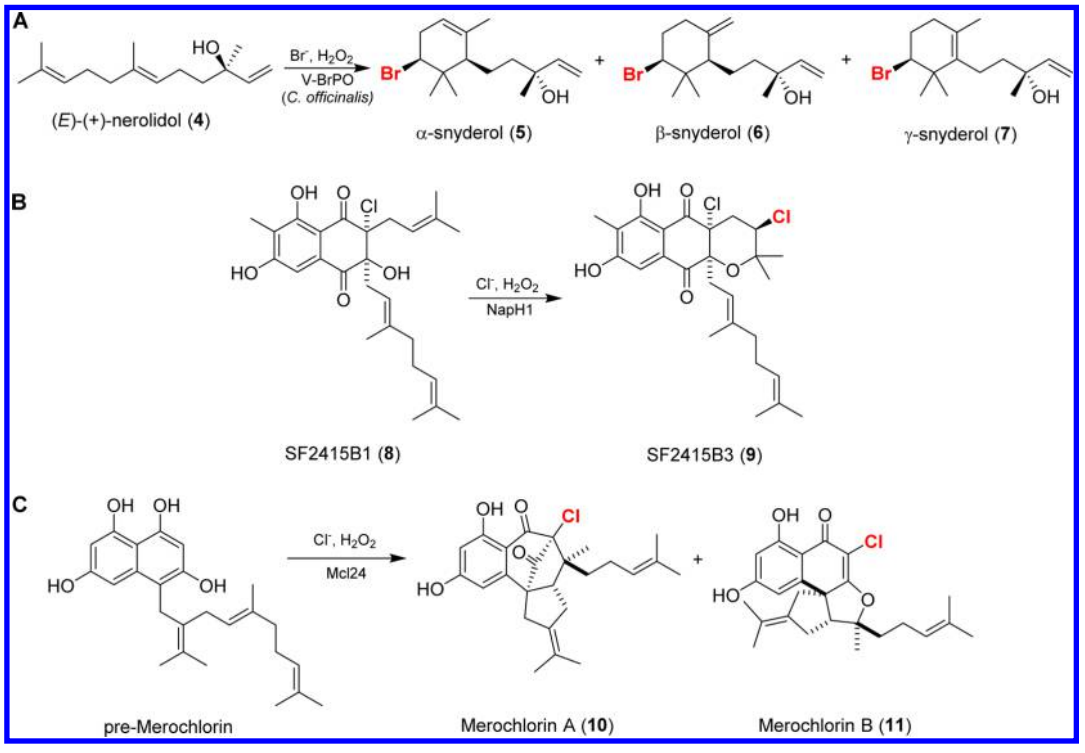


Figure 7. Examples of specific halogenations by vanadium-dependent haloperoxidases using (A) V-BrPO from *C. officinalis*, (B) NapH1, and (C) Mcl24.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

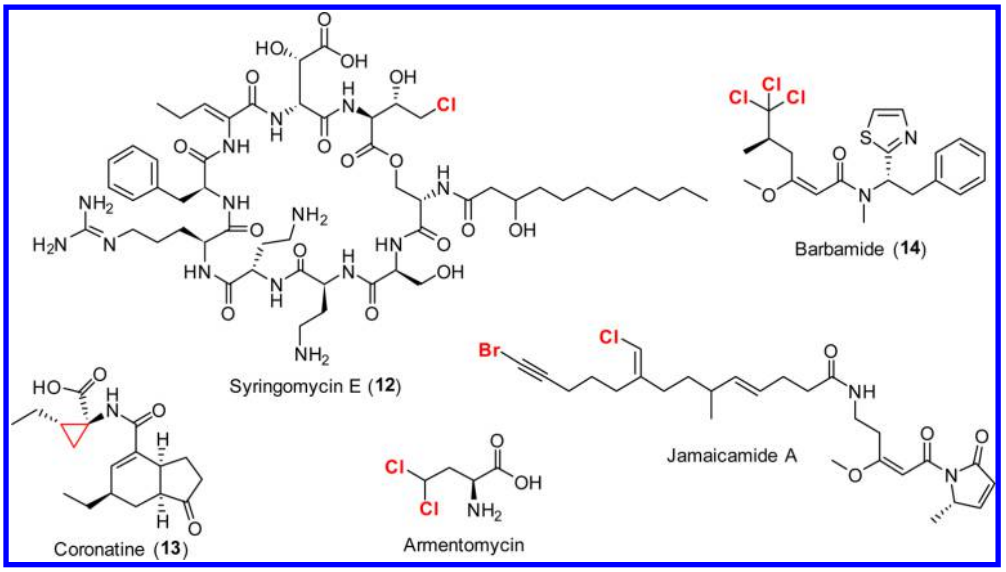


Figure 8. Halogenated natural products produced using nonheme Fe(II)-α-KG-dependent halogenases. The coronatine cyclopropyl moiety is formed after γ-halogenation of L-allo-isoleucine followed by intramolecular ring closure.

pilulifera are a homodimer and a homododecamer, respectively.[43,93−95]

Unlike the heme-dependent haloperoxidases, the oxidation state of the catalytic vanadium center is maintained throughout the catalytic cycle, offering the advantage that these enzymes do not suffer from oxidative inactivation. Initially, hydrogen peroxide binds to the distal position of the vanadate complex, leading to a loss of water and the formation of peroxo-vanadate intermediate. Halide ion is then oxidized, leading to the formation of hypohalous acid (Figure 5C). Examples in the literature suggest that, like the heme-dependent haloperoxidases, the vanadium-dependent haloperoxidases can release free hypohalous acid, allowing halogenation of a variety of substrates (Figure 6),[76,96−98] although some examples of substrate specificity have been observed.[92,99,100]

The vanadium-dependent bromoperoxidase from *Corallina officinalis* was found to be selective and utilized to asymmetrically brominate and cyclize the sesquiterpene (*E*)-(+)-nerolidol (**4**) to the marine natural products α-, β-, and γ-snyderol (**5**−**7**, Figure 7A). Nonenzymatic bromination of *E*-(+)-nerolidol (**4**) produced a mixture of diastereomers, whereas the bromoperoxidase only formed a single diastereomer. This study demonstrated the stereo- and regioselectivity of these enzymes for the first time, thereby suggesting that the substrate binds to the active site in a specific manner (Figure 7).[76]

More recently, the first example of a bacterial vanadium-dependent chloroperoxidase was reported from *Streptomyces* sp. CNQ-525. Three genes were identified, *napH1*, *H3*, and *H4*, with homology to other known vanadium-dependent haloperoxidases. It was hypothesized that at least one of these genes encoded for an enzyme that was involved in the halogenation and cyclization of SF2415B1 (**8**) to SF2415B3 (**9**) in a manner similar to that of (*E*)-(+)-nerolidol (**4**). All three enzymes were heterologously expressed in *E. coli*; however, NapH4 proved to be insoluble. NapH3 was found to be unable to catalyze the formation of SF2415B3 (**9**); however, NapH1 demonstrated the ability to catalyze the halocyclization in the presence of hydrogen peroxide and chloride at pH 6 (Figure 7B). Interestingly, NapH1 was not able to chlorinate non-native substrates, (±)-nerolidol or lapachol; however, nonspecific activity was exhibited with (±)-nerolidol and lapachol in the presence of bromide ions.[48] Another bacterial vanadium-dependent haloperoxidase (Mcl24) has been identified from the marine bacterium *Streptomyces* CNH-189. Mcl24 is a vanadium-dependent chloroperoxidase involved in the late stages of merochlorin (**10** and **11**) biosynthesis catalyzing a site-selective napthol chlorination, followed by a sequence of oxidative dearomatization and terpene cyclization reactions (Figure 7C). Like NapH1, Mcl24 is specific for chloride; however, in the presence of bromide, monochlorodimedone was also brominated, suggesting that halogenation by freely diffusing hypohalous acid may also occur if substrate is not bound to the active site.[49,83,101]

While haloperoxidases can halogenate a wide range of substrates under mild conditions, as shown in Figures 4 and 6, they generally either lack specificity or, as in the case of the bacterial vanadium-dependent haloperoxidases, are highly specific and therefore are limited in biocatalytic application.

3. α-KETOGLUTARATE-DEPENDENT HALOGENASES

3.1. Discovery of α-Ketoglutarate-Dependent Halogenases

Nonheme Fe(II)-α-ketoglutarate-dependent (α-KG) halogenases are involved in the halogenation of unactivated carbon centers on aliphatic moieties such as terminal methyl groups and predominantly utilize substrates tethered to acyl or peptidyl carrier proteins (PCP). The enzyme class was first reported in 2005, with the first example forming part of the biosynthetic assembly line required for production of the nonribosomal peptide Syringomycin E (**12**) in *Pseudomonas syringae* (Figure 8). The carrier protein tethered L-threonine (L-Thr-*S*-SyrB1) from Syringomycin biosynthesis was demonstrated to act as a substrate for the SyrB2-mediated chlorination.[102]

The selectivity of SyrB1 for L-Thr was confirmed by kinetic comparison with other amino acids showing a 60-fold preference for threonine over serine (L-Thr, K_m = 3.1 ± 0.2 mM, k_{cat} = 29.1 ± 0.9 min^{-1}; L-Ser, K_m 7.7 ± 0.5 mM, k_{cat} 1.23 ± 0.05 min^{-1}). Isolation of the final product from SyrB2-mediated chlorinations confirmed the product to be 4-Cl-L-threonine. SyrB2 was found to not accept free L-threonine at all, meaning that substrate must be tethered to the phosphopante-

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

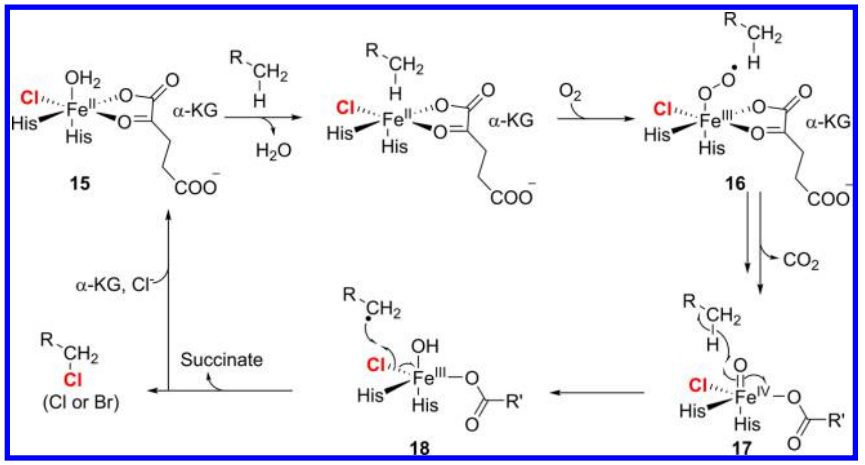


Figure 9. Proposed mechanism for Fe(II)/α-KG-dependent halogenases.

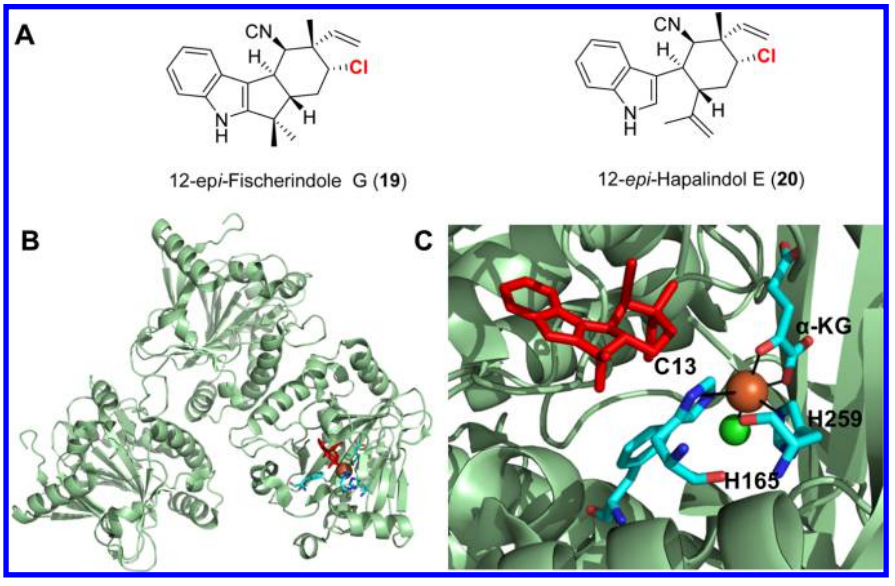


Figure 10. (A) Selectively chlorinated products from WelO5-catalyzed halogenation. (B) Crystal structure of Wel05 showing overall structure (PDB 5IQV). (C) Active site coordination of Wel05 (PDB 5IQV). Fe and Cl are represented by orange and green spheres, respectively.

theine arm of the halo-SyrB1 thiolation domain for chlorination to occur. Tethering of the substrate to SyrB1 was confirmed by the detection of [^{36}Cl]-L-Thr-*S*-SyrB1 using the radioactive ^{36}Cl and gel dosimetry. In vitro studies revealed a requirement for Fe(II), O_2, and Cl^- or halogenase activity. In addition, deactivation of halogenase activity was rapid such that no more than seven turnovers could be detected.[102]

On the basis of the observed cofactor and cosubstrate requirement, and by analogy to the α-KG hydroxylases, a mechanism for the α-KG halogenases was proposed involving radical abstraction of the methyl hydrogen by a high-valent Fe^{IV}-oxido species. Recombination of this radical with Fe-coordinated Cl^- was then thought to afford 4-chloro-L-theronine. At a similar time, investigation of the biosynthetic pathway leading to coronatine (**13**) in *P. syringae* identified CmaB and CmaC to be involved in halogenation of the methyl group of L-allo-isolecine, which undergoes subsequent cyclization to generate the cyclopropyl ring of the side chain amino acid (corronamic acid).[103] The trichloromethyl group of barbamide (**14**) derivatives from the marine cyanobacterium *Lyngbya majuscule* was also found to be derived from the pro-*R* methyl group of leucine.[104,105] Subsequent in vitro studies have confirmed that two α-KG-dependent halogenases (BarB1 and BarB2) are required to trichlorinate a carrier protein tethered L-Leu precursor in barbamide (**14**) biosynthesis (Figure 8).[106]

Crystallography of SyrB2 subsequently shed further light on the mechanistic aspects of the Fe(II)/α-KG-dependent halogenases.[107] Analogous to the hydroxylation mechanism of TauD, a working mechanism was proposed for these terminal halogenations. SyrB2 was the first structurally characterized mononuclear iron protein which did not display the characteristic 2-His, 1-carboxylate iron coordination. Instead, two histidine ligands are present (His116 and His235), and the iron coordinates to a chloride ion rather than carboxylate. It was also observed that water and α-KG coordinate to iron in the resting state (**15**) (Figure 9). Upon L-Thr-*S*-SyrB1 binding, water is displaced by dioxygen (**16**), and decarboxylation of α-KG would then lead to a high valent Fe^{IV}-oxido intermediate (**17**). This highly reactive species is then able to radically abstract a hydrogen atom from substrate (**18**), which can then recombine with a chloride ligand producing chlorinated L-Thr-*S*-SyrB1 and regenerated Fe(II). Although there is clearly the possibility of hydroxylation during the course of the reaction, chlorination seems to be greatly favored. One possible

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

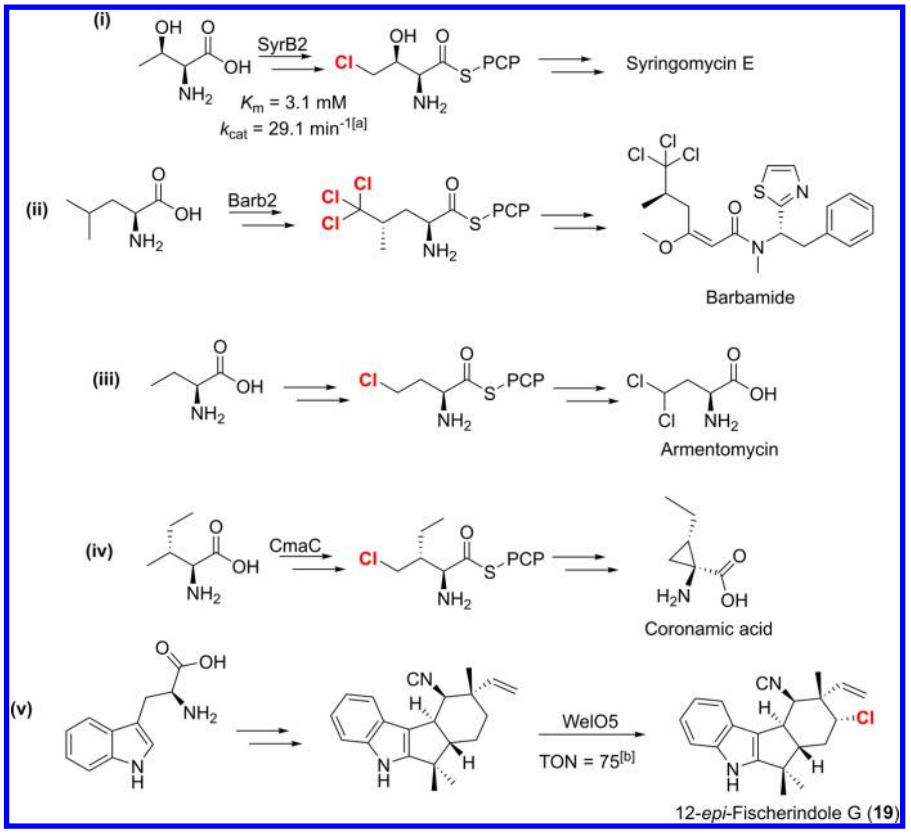


Figure 11. Involvement of Fe(II)/α-KG-dependent halogenases in the biosynthesis of halogenated natural products. [a] Refers to the loading of L-Thr onto SyrB1 as those for SyrB2 could not be obtained due to enzyme deactivation. TON for chlorination of L-Thr-*S*-PCP by SyrB2 is <7.[102] [b] Refers to the chlorination of free-standing substrate by WelO5.[108] PCP = peptidyl carrier protein. TON = turnover number.

explanation is the relative positioning of the chloride ligand toward the substrate, exclusively affording halogenation.

Recently, the Fe(II)/α-KG-dependent halogenase WelO5 was discovered from the welwitindoline biosynthetic pathway in *Hapalosiphon welwitschii*.[108−110] It was shown that this enzyme can regioselectively monochlorinate an aliphatic carbons to afford 12-*epi*-Fischerindole G (**19**) and 12-*epi*-Hapalindol E (**20**, Figure 10A).[111] WelO5 is particularly notable as the first example of an Fe(II)/α-KG-dependent halogenase that functions independently of a carrier protein. It is also the first example of a halogenase installing a halogen atom stereoselectively on a nonactivated position, C13 of 12-epi-fischerindole G, to create a new stereogenic center. When the WelO5 sequence was compared to other members of the α-KG-dependent oxygenase superfamily, the D/E residue in the characteristic HX(D/E)XnH motif for iron binding was found to be absent and replaced by the glycine 166 residue. The two histidine residues H164 and H259 were found crucial for iron coordination. Site-directed mutagenesis of H259 to phenylalanine completely abolished activity toward the 12-*epi*-fischerindole U. The recent publication of the WelO5 crystal structure revealed the close proximity of the C13 halogenation site to the putative *cis*-halo-oxo-Fe^{IV} complex, which confirms its preferred activity toward halogenation (Figure 10C).[111] Unlike the standard α-KG hydroxylases, WelO5 contains a glycine (Gly166) at the sequence position of an aspartate or glutamate ligand. Accordingly, it was shown that the G166D variant exclusively gives C13 hydroxylation, as predicted. It was also demonstrated that the second-sphere mutant S189A produced a mixture of hydroxylated and halogenated product, showing how outer sphere hydrogen bonding can influence chemoselectivity.[111]

3.2. Application of Fe(II)/α-KG-Dependent Halogenases as Biocatalysts

The proteinaceous nature of substrates for the majority of Fe(II)/α-KG halogenases forms a barrier for their application as biocatalysts. These enzymes also have to be purified under inert conditions, and their turnover numbers are generally low. In fact, kinetic parameters are difficult to determine due to auto-oxidative inactivation of the biocatalyst and limited substrate availability (maximum of 75 turnovers), which restricts their application in industrial biocatalysis. Moreover, the substrate scope identified to date seems to be largely limited to natural amino acids (Figure 11). However, the recent discovery of WelO5 allowing halogenations without the requirement of a carrier protein opens new opportunities for the application of those enzymes in synthesis. Moreover, given that mutagenesis studies have shown that the Fe(II)/α-KG halogenases can be switched to hydroxylases,[107,111,112] it seems plausible that halogenases can be engineered from Fe(II)/α-KG-dependent hydroxyases. Indeed, a recent report has demonstrated that it is possible to engineer a new halogenase enzyme from an Fe(II)/α-KG-dependent hydroxylase using structure guided mutagenesis.[113] This approach[113] is attractive given that there are more known Fe(II)/α-KG-dependent hydroxylases than halogenases, and the range of substrates that are processed by the hydroxylases is greater and more structurally diverse. Moreover, unlike halogenases, which, with the exception of WelO5, halogenate carrier protein-tethered substrates, the majority of hydroxylases utilize simple

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

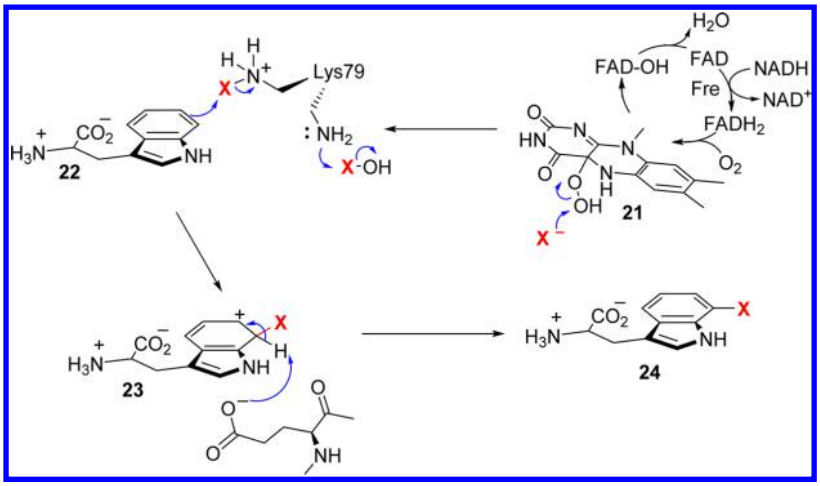


Figure 12. Proposed mechanism of flavin-dependent tryptophan halogenase-catalyzed halogenation.

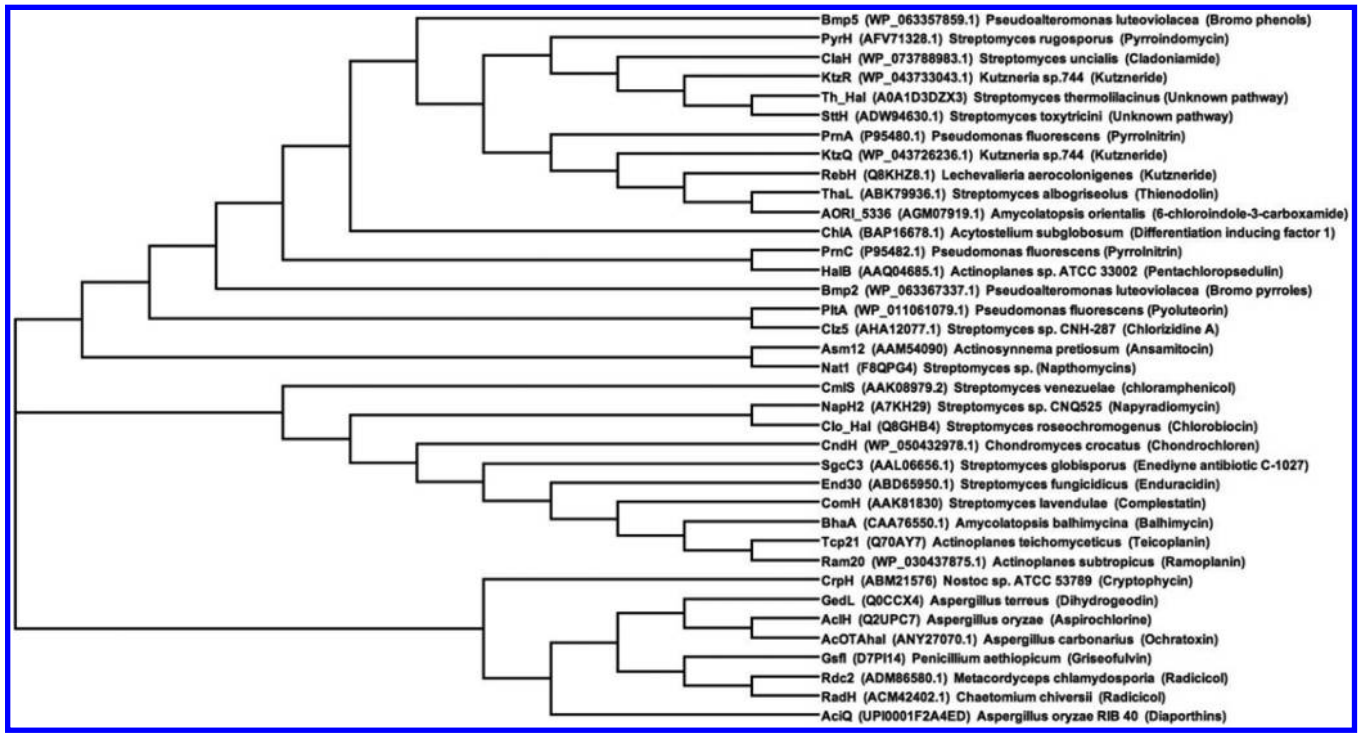


Figure 13. Phylogenetic tree of the Fl-Hals discussed herein. Generated using the neighbor-joining tree method without distance corrections via Dendroscope 3.

nontethered substrates. Despite their low catalytic activity, the fact that Fe(II)/α-KG enzymes halogenate, or hydroxylate, unactivated aliphatic substrates with regio- and stereoselectivity is attractive for further investigation given that there are few viable synthetic alternatives.

4. FLAVIN-DEPENDENT HALOGENASES

4.1. Flavin-Dependent Halogenases in Nature

Flavin-dependent halogenase (Fl-Hal) enzymes belong to the superfamily of flavin-dependent monooxygenases.[114,115] A key feature of this class of enzymes is their activation of molecular oxygen using reduced flavin ($FADH_2$) to generate C4a-hydroperoxy flavin (**21**),[116−118] which allows diverse reactions such as hydroxylation, epoxidation, Baeyer−Villager oxidation, and heteroatom oxidations.[119−121] The flavin-dependent halogenases are typically classified as two-component monooxygenases, meaning that they utilize freely diffusing reduced flavin cofactors produced by an additional flavin-reductase enzyme.[115] In addition to the flavin reductase, several Fl-Hals from NRPS and PKS pathways also require enzymes for substrate activation or tethering.[32] In the last two decades, a number of flavin-dependent halogenases from both bacterial and fungal biosynthetic pathways have been identified.[114,122−126]

The flavin-dependent halogenases were originally thought to operate via a mechanism analogous to those of other flavin-dependent monooxygenases through direct reaction of substrate with C4a-hydroperoxy-flavin (FAD-OOH, **21**), followed by subsequent reaction with halide anion.[123,127−129] However, crystallography of the tryptophan (**22**) 7-halogenase PrnA revealed that the substrate and flavin binding sites were spatially distinct, separated by a 10 Å tunnel, and therefore

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

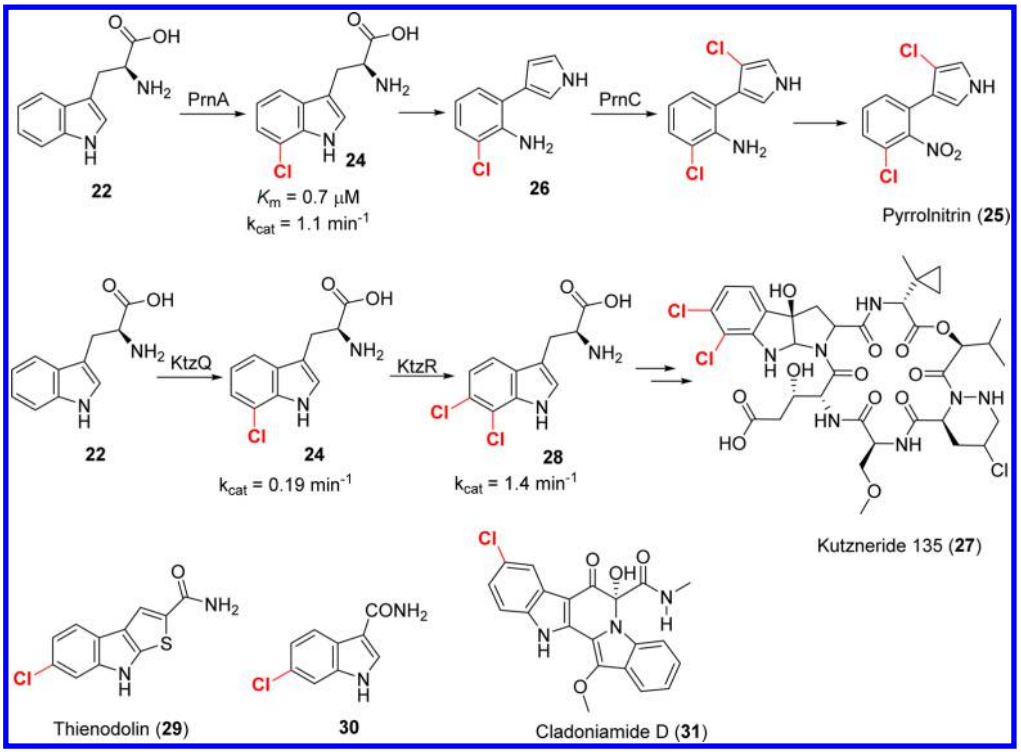


Figure 14. Halogenated natural products biosynthesized using flavin-dependent tryptophan halogenases.[141,155]

direct interaction of the substrate with the cofactor is not possible.[130,131] Mechanistic studies subsequently demonstrated that FAD-OOH (**21**) was generated prior to halogenation, even in the absence of substrate, suggesting that nucleophilic attack of chloride on **21** may result in hypochlorous acid generation in the flavin-binding domain.[132] Further work demonstrated the formation of a long-lived enzyme-chloride adduct, which, after identifying the vital importance of an active site lysine residue, is believed to be a covalent chloramine adduct or hydrogen-bonded lysine-hypochlorous acid species.[133] This electrophilic chlorine species is believed to be ultimately responsible for aromatic substitution of the substrate to generate the Wheland intermediate (**23**), which is then deprotonated by a conserved glutamate residue to afford chlorinated product (**24**, Figure 12).[114,131,133−135] Positioning of this active site lysine relative to substrate is therefore believed to control which position of the substrate is halogenated.[130] Direct monooxygenase-type activity is thought to be prevented in the tryptophan Fl-Hals by a conserved structural motif (WxWxIP), which blocks direct contact between FAD-OOH (**21**) and substrate.[130,136]

This highly conserved motif, along with the nucleoside-binding motif GxGxxG responsible for binding $FADH_2$, are therefore considered to be the signature motifs for identifying putative Fl-Hal genes. These signature motifs have been used to identify flavin-dependent halogenases from diverse biosynthetic pathways that are responsible for the halogenation of pyroles, phenols, and other aromatic natural product precursors (Figure 13). Similar to other flavin-dependent monoxygenases, Fl-Hals of different substrate scope are thought to conserve the core flavin-binding domain with most structural changes coming about from the recruitment of different substrate binding domains.[115] Fl-Hals can therefore be further subdivided on the basis of their natural substrate.

4.1.1. Flavin-Dependent Tryptophan Halogenases. The flavin-dependent tryptophan halogenases are the most extensively studied and characterized Fl-Hals.[32−34] The first tryptophan Fl-Hal to be identified was PrnA, which is required for the biosynthesis of pyrrolnitrin (**25**), a broad spectrum antifungal compound produced by *Pseudomonas fluorescens*.[127,137,138] In the pyrrolnitrin pathway, PrnA chlorinates the 7-position of tryptophan, while a second Fl-Hal, PrnC, chlorinates the pyrrolic intermediate **26** (Figure 14). Another tryptophan 7-halogenase RebH, with 55% sequence identity to PrnA, was subsequently identified from the rebeccamycin biosynthetic pathway in *Lechevalieria aerocolonigenes*.[123,124,139] X-ray crystal structures of these two Fl-Hals were subsequently determined which, along with further biochemical studies, provided the basis of our understanding of the mechanism and reactivity of Fl-Hals.[114,123,127,132,134,135]

Studies into the biosynthesis of the dichlorinated non-ribosomal peptide kutzneride (**27**), an antifungal from actinobacteria *Kutzneria* sp. 744, revealed that two tryptophan Fl-Hals (KtzQ and KtzR) and a cryptic Fe(II)/α-KG-dependent halogenase (KtzD) were required for kutzneride biosynthesis.[140−142] KtzQ and KtzR work sequentially, with KtzQ catalyzing the formation of 7-chlorotryptophan prior to halogenation at the 6-position by KtzR to produce 6,7-dichlorotryptophan (**28**, Figure 14).[141] In vitro studies showed that KtzR exhibits a 120-fold substrate preference for 7-chlorotryptophan (**24**) over nonchlorinated tryptophan.[141] A number of independent tryptophan 6-halogenases have also been identified recently.[143−147] For example, ThaL from the indole alkaloid thienodolin (**29**) pathway in *Streptomyces albogriseolus* is responsible for the 6-chlorination of tryptophan prior to thiophene ring formation.[144,146,148,149] SttH from *Streptomyces toxytricini* and Th_Hal from the thermophile *Streptomyces violaceusniger* SPC6 have also been shown to function as tryptophan 6-halogenases.[143,145] Although both have been characterized in vitro and their structures determined,[143,145,150] the biosynthetic pathways in which they

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

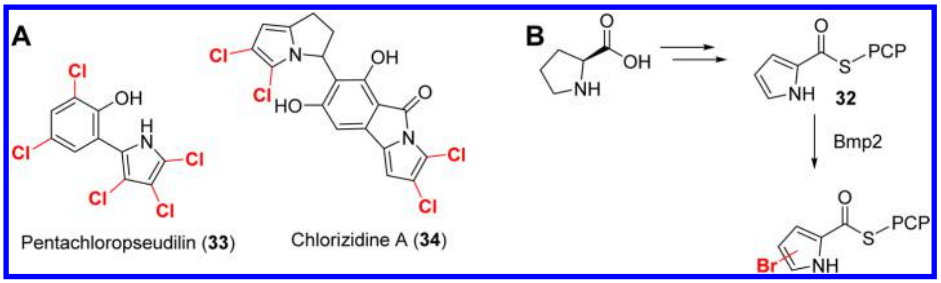


Figure 15. Flavin-dependent pyrrole halogenases in natural product biosynthesis. (A) Natural products containing chlorinated pyrrole moieties. (B) Halogenation of pyrrole-*S*-PCP by Bmp2.

function are yet to be explored. Recently, a tryptophan 6-halogenase (AORI_5336) was identified in a hybrid NRPS/PKS hybrid gene cluster of *Amycolatopsis orientalis*, which appears to be involved in the biosynthesis of the plant growth-regulating compound **30**.[147]

The only fully characterized tryptophan 5-halogenase is PyrH from the pyrroindomycin producing strain *Streptomyces rugosporus*.[130,151] Although a tryptophan 5-halogenase is known to be involved in the biosynthesis of the bis-indole alkaloid cladoniamide (**31**) in *Streptomyces uncialis*, details of in vitro activity are still lacking, although in vivo studies suggest it may be of low activity compared to other tryptophan halogenases.[152,153] Additionally, a recent report suggests that a tryptophan 5-halogenase may be responsible for halogenation of a tryptophan moiety in a peptide chain, rather than freely diffusing tryptophan, in the biosynthesis of the lantibiotic NAI-107.[154] Structural and phylogenetic analyses (Figure 13) provide details of the similarities between halogenases of different regioselectivity. Many of the structural differences are present in a single loop region, with the tryptophan 7-halogenase PrnA containing a larger loop region (L443-W455) as compared to the 5- and 6-halogenases, suggesting that the 5-, 6-, and 7-tryptophan halogenases may have evolved from the same ancestral enzyme via loop insertions and deletions to give rise to different orientations of substrate in the active site.[130,150]

4.1.2. Flavin-Dependent Pyrrole Halogenases. In addition to the pyrrolnitrilin halogenase, PrnC, which halogenates a free-standing pyrrolic substrate, there are a number of Fl-Hals that process pyrrol-2-carboxy thioester substrates (**32**) tethered to peptidyl carrier proteins in NRPS or hybrid NRPS-PKS assembly lines (Figure 15).[122,156] This includes PltA from the pyoluteorin biosynthetic gene cluster.[122,137,157] The crystal structure of PltA revealed a unique helical region at the C-terminus, which blocks the substrate binding cleft. Binding of PltL-tethered substrate induces a conformation change, however, which allows access to the substrate-binding cleft.[122] Interestingly, PltA does not contain a residue analogous to the catalytic glutamate found with the tryptophan halogenases, suggesting a water molecule, or an as yet unidentified active site residue, might function to deprotonate the Wheland intermediate. Both PrnC and PltA appear to have relaxed regioselectivity as compared to the tryptophan halogenases, and mixtures of mono- and dichlorinated products are observed more frequently in vitro.[122,158] HalB is another pyrrole halogenase with high sequence similarity to PrnC and PltA, involved in the biosynthesis of pentachloropseudilin (**33**) from *Actinoplanes* sp. ATCC3302.[159,160] Clz5 from *Streptomyces* sp. CNQ-418 is also similar to PltA and has been identified as the pyrrole Fl-Hal in the biosynthesis of chlorizidine A (**34**, Figure 15A).[161−163] The peptidyl carrier protein (PCP) tethered substrates of Clz5, PltA, and HalB are all derived from oxidation of proline-*S*-PCP intermediates in NRPS or NRPS-PKS assembly lines.[162,163]

Genome mining of the marine organisms responsible for the biosynthesis of brominated marine natural products has led to the identification of an interesting pyrrole Fl-Hal, Bmp2, from *Pseudoalteromonas luteoviolacea*.[156,164,165] Similar to the gene cluster containing PltA, Bmp2 is associated with a thioesterase (TE, Bmp1) and proline adenyltransferase (Bmp4).[156,164] Reconstitution of these enzymes in vitro showed Bmp2 to brominate pyrrole-2-carboxy-*S*-Bmp1 (**32**) to mono-, di-, and tribromo products (Figure 15B).[164] Unlike the other Fl-Hals studied to date, Bmp2 appeared capable of iodination of **32** in addition to bromination, but not chlorination.[164] Investigation of the structural features which allow Bmp2 to iodinate may be of interest in the engineering of other Fl-Hals, as the tryptophan Fl-Hals have been demonstrated to be inhibited by the presence of I^-.[114,166] Biocatalytic iodination is attractive for synthetic applications, as aryl iodides are typically more reactive, and thus more readily derivatized, than the corresponding aryl chlorides or bromides. Comparison of Bmp2 to its closest analogue Myp16 and point mutations of Bmp2 to mimic this chlorinase did not afford a Bmp2 mutant capable of chlorination, suggesting that other, more subtle, features may be responsible for the observed halide preference.[156]

4.1.3. Flavin-Dependent Phenolic Halogenases. Three classes of phenolic flavin-dependent halogenase enzymes have been identified to date: those that act on free-standing substrates (class A); those that require substrates tethered to a carrier protein (class B); and halogenases that follow a decarboxylative halogenation mechanism (class C).

4.1.3.1. Predicted Class A Phenolic Halogenases. The fungal halogenase Rdc2 from *Pochonia chlamydosporia* and the similar enzyme RadH from *Chaetonium chiversi* are both involved in the biosynthesis of the resorcinylic acid lactone (RAL) radicicol (**35**) and act on free-standing phenolic substrates.[125,167−170] Radicicol is a potent inhibitor of heat shock protein 90 (Hsp90), a medically important target due to its involvement in many cancer-causing pathways, and also exhibits antifungal activity.[170] Rdc2 was heterologously expressed in *E. coli* and demonstrated to be capable of halogenating the macrocyclic natural products zearaleone, dihydrosorcylide, curvularin, and curcumin in vitro.[169−172] Both chlorination and bromination reactions were found to form dihalogenated compounds after prolonged periods when multiple OH groups were present on the aromatic ring.[169−172] Rdc2 and RadH are both post-PKS tailoring enzymes, which is a plausible explanation for why these enzymes act on free-standing substrates. A number of other class A phenolic Fl-Hals

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

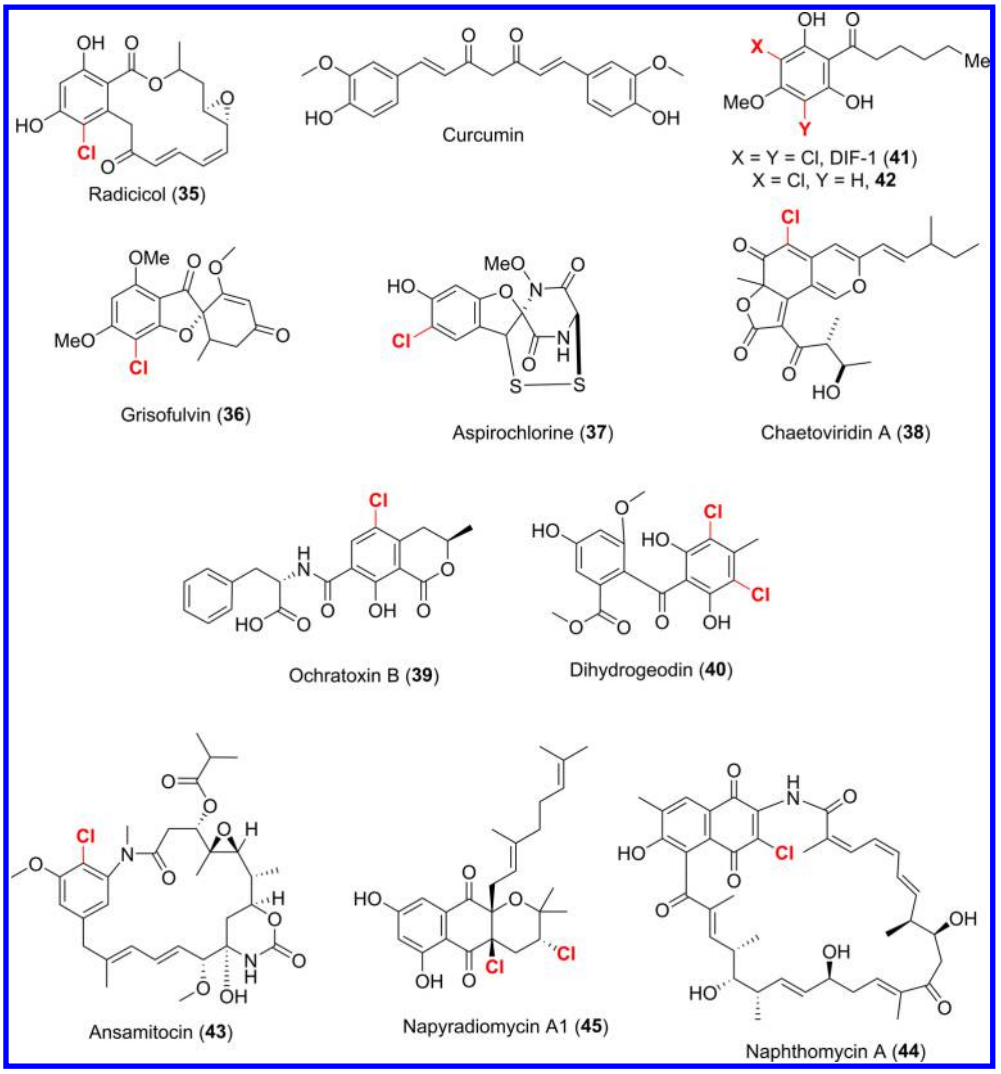


Figure 16. Natural products biosynthesized using phenolic flavin-dependent halogenases.

have been identified that are involved in the biosynthesis of a range of fungal natural products including griseofulvin (**36**, Gsfl),[173−175] aspirochlorine (**37**, AclH),[176−178] chaetoviridin (**38**, CazI),[179−181] ochratoxin (**39**, AcOTAhal),[126,182−184] and dihydrogeodin (**40**, GedL) (Figure 16).[185,186] Like Rdc2 and RadH, these related fungal enzymes (Gfsl, AclH, CazI, and GedL) halogenate *ortho* to a phenolic hydroxyl group of late-stage biosynthetic intermediates, except for the ochratoxin (**39**) halogenase (AcOTAhal) which installs a chloro substituent *para* to the hydroxyl.

Although not directly phylogenetically linked to Rdc2 or RadH, the differentiation-inducing factor 1 (DIF-1, **41**) halogenase (ChlA) from *Dictyostelium discoideum* is a class A flavin-dependent phenolic halogenase.[187−190] ChlA was over-expressed and purified in *E. coli* and demonstrated in vitro to produce both mono- and dihalogenated DIF-1 (**41** and **42**) in the presence of an external flavin reductase.[187] The observation that many of the phenolic halogenases appear to require substrates with a hydroxyl group *ortho* to the halogenation site (vide infra) suggests that electrophilic aromatic substitution may be mediated through deprotonation of the *ortho*-hydroxyl.

Another halogenase, predicted to process a free-standing phenolic substrate, is Asm12 from ansamitocin (**43**) biosynthesis.[191−193] Asm12 shares 73% sequence identity with Nat1, a halogenase involved in the biosynthesis of napthomycins (**44**, NATs) by *Streptomyces* sp. CS.[194] Napthomycins are 29-membered napthalenic ansamacrolactam antibiotics with similarities to the ansamitocin type scaffold. Inactivation of the *nat1* gene in the napthomycin biosynthetic gene cluster abolished production of the chlorinated napthomycin A, which was restored on complementation with Asm12, demonstrating the functional similarity between the two enzymes.[194] Finally, NapH2 is the halogenase from the napyradiomycin (**45**) biosynthetic cluster (nap) in *Streptomyces* sp. CNQ-525.[195] In vivo studies suggest that NapH2 may halogenate the C2 position of a napthoquinone precursor, facilitating subsequent prenylation during napyradiomycin biosynthesis.[195,196]

4.1.3.2. Predicted Class B Phenolic Halogenases. A number of putative phenolic halogenases such as Clo-Hal (involved in the biosynthesis of chlorobiocin (**46**) in *Streptomyces roseochromogenes*),[197,198] BhaA (from the balhimycin (**47**) biosynthetic pathway in *Amycolatopsis mediterranei*),[199−201] and ComH (responsible for chlorination en route to complestatin (**46**) in *Streptomyces lavendulae*)[202−204] have been identified (Figure 17). While further in vitro studies are required to elucidate the true substrates of these enzymes, it is likely that Clo-Hal, BhaA, and ComH are all class B, carrier protein-dependent, phenolic halogenases. In chlorobiocin (**59**) biosyn-

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

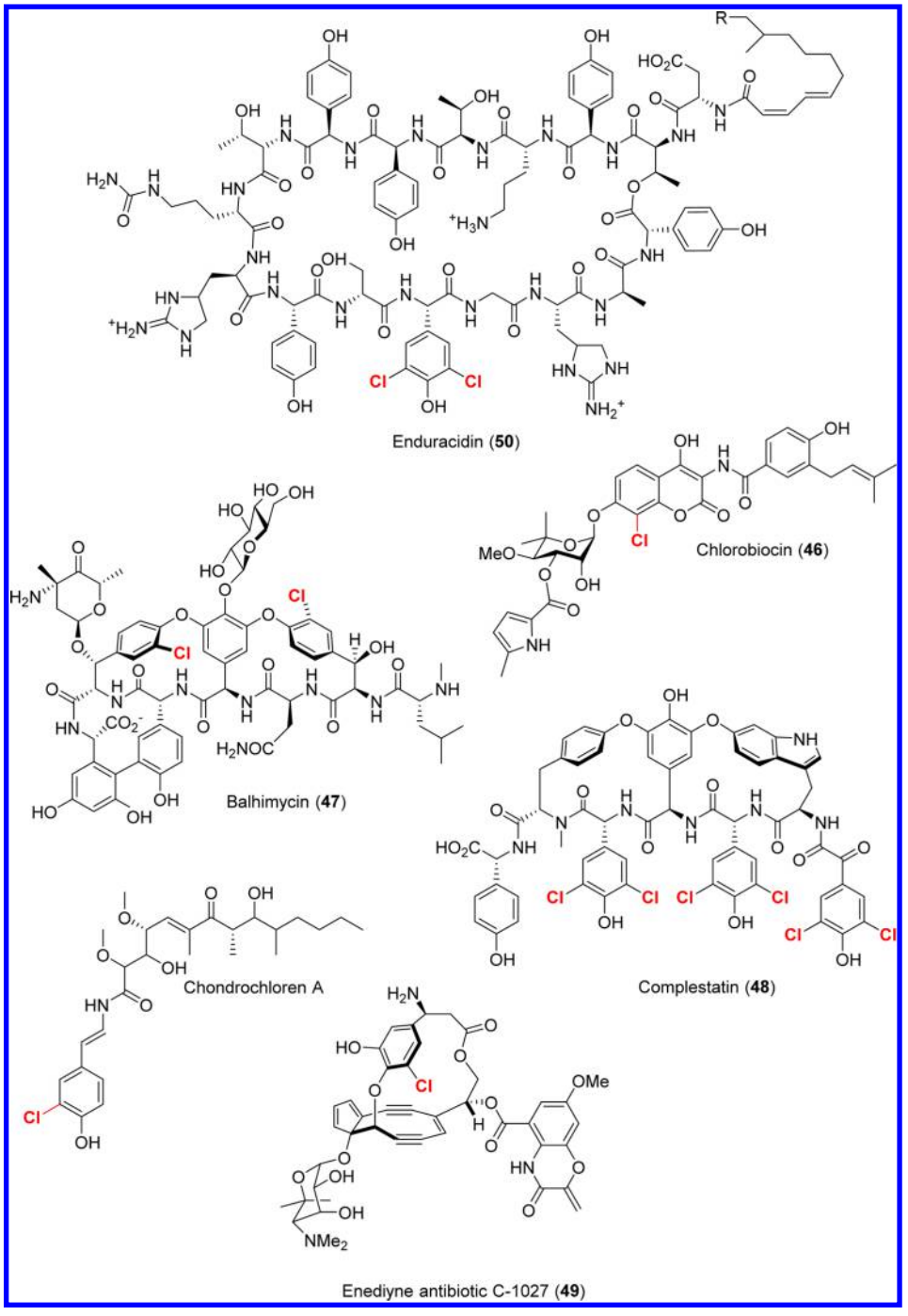


Figure 17. Natural products biosynthesized using phenolic flavin-dependent halogenases.

thesis, it is suggested that Clo-Hal halogenates a PCP tethered tyrosine or β-hydroxy tyrosine (β-HT) intermediate prior to formation of the aminocoumarin ring system.[197] The location of the BhaA encoding gene in the balhimycin gene cluster alongside NRPS encoding genes, and in vitro experiments also suggest that BhaA halogenates a PCP tethered tyrosine or β-HT intermediate.[199−205] The halogenase (Tcp21) in the teicoplanin gene cluster of *Actinoplanes teichomyceticus* is similarly predicted to process a Tyr- or β-HT-*S*-PCP substrate.[205−207]

Crystallography of the related chondrochloren halogenase CndH revealed that the substrate for CndH may also be a tyrosine-related intermediate that is most likely PCP tethered.[208] CndH has a large nonpolar surface patch that could accommodate the putative carrier protein, and it was also proposed that the catalytic base required for halogenation (Glu in tryptophan halogenases) might be supplied by C-terminal domain of the carrier protein.[208]

Other class B phenolic halogenases that require a carrier protein include SgcC3 involved in the biosynthesis of C-1027 (**49**), a chromoprotein antitumor antibiotic isolated from *Streptomyces globisporus*.[209] In vitro activity of SgcC3, overexpressed in *E. coli*, was observed only in the presence of SgcC2 peptidyl carrier protein-tethered substrates. SgcC3 was found to utilize both (*S*)- and (*R*)-β-tyrosyl-*S*-SgcC2, but not 3-hydroxy-β-tyrosyl-*S*-SgcC2.[209−211]

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

The halogenases Ram20 and End30 have also been described and are suggested to halogenate hydroxyphenylglycine (Hpg) residues during the biosynthesis of related lipopeptide antibiotics ramoplanin and enuracidin (**50**).[212−215] Despite exhibiting high sequence similarity, End30 is responsible for dichlorination of Hpg13 in enduracidin, while Ram20 monohalogenates the Hpg17 residue of ramoplanin. Interestingly, deletion of *ram20* from the ramoplanin biosynthetic gene cluster and complementation with *end30* resulted in monohalogenation of Hpg17, while the same complementation experiment in a strain lacking the ramoplanin mannosyltransferase resulted in halogenation of Hpg13.[212] This suggests that halogenation occurs after peptide assembly and mannosylation of Hpg11, with the bulky mannosyl groups blocking halogenation at the more proximal Hpg13 residue. Further in vitro experiments are required to fully evaluate the regioselectivities of End30 and Ram20.

4.1.3.3. Class C Phenolic Halogenases. A decarboxylating phenol brominase enzyme, Bmp5, was discovered from *Pseudoalteromonas luteoviolacea* which was the first example of a class C phenolic Fl-Hal.[156,164,165] This enzyme lacks sequence similarity with other canonical flavin-dependent halogenases and has sequence homology to known single-component flavin-dependent monooxygenases.[156,164,165] Bmp5 does not require an external flavin reductase enzyme for in vitro activity, and activity was abolished in the absence of either bromide or NADPH, indicating Bmp5 alone is capable of flavin reduction and bromination. Reaction of Bmp5 with 4-hydroxybenzoic acid (**51**) in the presence of bromide, NADPH, and FAD led to the formation of 3-bromo-4-hydroxybenzoic acid **52**. A second bromination is then proceeded by decarboxylation to afford dibromo phenol **53** (Figure 18A). Bmp5 showed no

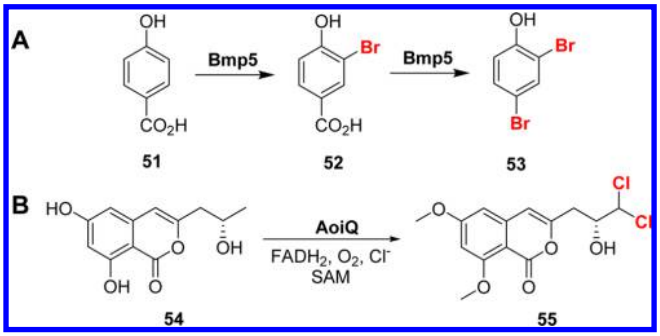


Figure 18. (A) Decarboxylative bromination catalyzed by Bmp5. (B) Halogenation and methylation activity of the proposed Fl-Hal-methyl transferase fusion protein AoiQ.

chlorination activity, while it forms iodophenols, indicating a highly evolved bromide or iodide binding site in the protein which is not present in other canonical flavin-dependent halogenases.[164] The full mechanistic and structural details of Bmp5 are still lacking, and the proposed decarboxylative flavin-dependent halogenase mechanism has no precedent. Whether the two steps are consecutive or require involvement of any specific catalytic residues is unclear; however, identification of Bmp5 homologues from different genomic clusters may shed light on those questions.[164]

4.1.4. Aliphatic Flavin-Dependent Halogenases. A small number of flavin-dependent halogenases have recently been identified, which are capable of halogenating aliphatic C−H bonds. The crystal structure of the chloramphenicol halogenase CmlS from *Streptomyces venezuelae*, for example, has been solved and displays a number of structural features not found in other Fl-Hals. A unique dynamic C-terminal domain creates a T-shaped tunnel leading to the active site. Most surprisingly, FAD appears to be covalently bound to the halogenase by the 8α carbon, not found in other Fl-Hals.[136] A number of nonpolar surface patches suggest an activator or substrate-tethering protein is also involved, as with CndH.[208] ClmK, the proposed activator protein, has sequence homology to acyl-CoA synthetases leading to the postulation that ClmS may halogenate the CoA-thioester or free acyl group directly. The active site residue E44 is thought to catalyze HOCl generation, while Y350 is believed to stabilize the enolate intermediate, which can then act as a nucleophile to generate chlorinated product.[136]

A bifunctional methyl-transferase halogenase has recently been identified from the diapothins genome cluster of *Aspergillus oryzae* RIB40 (AoiQ) and found to be responsible for the geminal dichlorination of a methyl group in diaporthin (**54**) both in vitro and in vivo affording **55** (Figure 18B).[216,217] The position of halogenation suggests a mechanism different from both canonical Fl-Hals and CmlS as generation of an enolate from a secondary alcohol would likely require additional oxidoreductases.[217] Genomic analysis of AoiQ revealed that homologous enzymes are encoded in various other fungal biosynthetic gene clusters.[217] Identification of these aliphatic Fl-Hals, in addition to further in vitro structural and functional characterization, may provide exciting insight into their mechanism, as well as provide the potential for regio- and stereoselective aliphatic halogenation reactions for biocatalytic applications.

4.2. Flavin-Dependent Halogenases as Biocatalysts

4.2.1. Substrate Scope of Fl-Hals. There has been significant interest in using flavin-dependent halogenases as biocatalysts because of their potential to regioselectively halogenate aromatic substrates under benign conditions. Thus far, most of this work has been focused on the tryptophan halogenases and has revealed a number of enzymes capable of halogenating tryptophan derivatives,[218] non-natural indolic substrates,[150,155,219−221] in addition to benzamides and benzoic acids[150,155] as well as naphthols and naphthyl amines (Figure 19).[220,221] In a number of cases, halogenation occurs with good regioselectivity. The tryptophan-7-halogenase RebH has been demonstrated to solely halogenate the 7-position of a number of non-natural indoles (**56**−**58**) and *ortho*- to the $-NH_2$ and −OH groups in **59** and **60**, respectively.[220,221] Interestingly, PrnA, another tryptophan-7-halogenase, was found to halogenate some of these indolic substrates at the more electronically favored C2 position.[219]

Introduction of electron-donating substituents onto the indole ring, which activate the substrate toward electrophilic aromatic substitution, led to the formation of dichlorinated products with RebH.[220,221] The formation of dichlorinated products and regioisomers has also been observed from the halogenation of tryptophan derivatives bearing electron-donating substituents with RebH (**61**−**63**), and in the halogenation of 3-(2′-aminophenyl)pyrrole (**64**) with PrnA,[218,219] demonstrating that substrate electronics influence regioselectivity when substrate positioning relative to the active site lysine is perturbed or flexible. Indeed, both RebH and PrnA have been shown to halogenate the most nucleophilic C3 position of indole (**65** and **66**) when this position is not

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

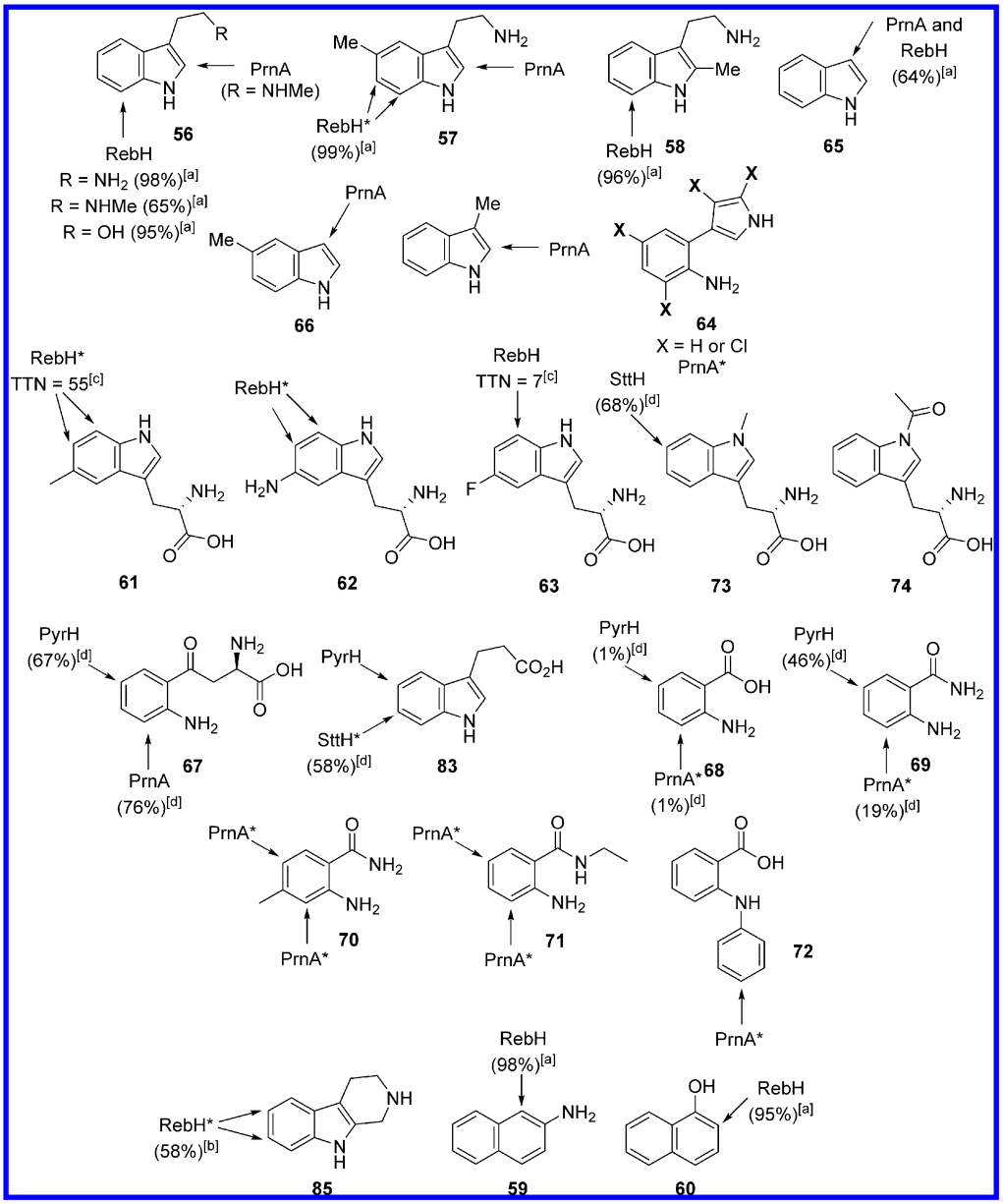


Figure 19. Known substrate scope and regioselectivity of the flavin-dependent tryptophan halogenases. *Indicates examples where halogenation at multiple positions occurs. Conversions are shown in brackets. [a] Measured using 5 mol % of RebH.[220] [b] Measured using 10 mol % of RebH.[220] [c] TTN = total turnover number determined using 0.4 mol % RebH.[218] [d] Conversions are reported relative to the conversion of tryptophan by the relevant enzyme under the same conditions.[155]

functionalized.[219−221] Moreover, with the highly activated substrate **64**, seven halogenated products were formed (including regioisomers and dichlorinated products) due to the electron-rich nature of both the pyrrole and the aniline moieties.[219] Most of the tryptophan halogenases have also been shown to chlorinate both L- and D- tryptophan, with preference for the natural L-enantiomer.[219−221]

In an attempt to broaden substrate scope of the enzymatic halogenation, the tryptophan-5- and tryptophan-6-halogenases PyrH and SttH have also been studied. With a number of 2-amino benzamides and benzoic acids (**67**−**72**), PyrH and SttH were found to solely chlorinate *para* to the $-NH_2$, while PrnA afforded mixtures of both *ortho*- and *para*-chlorinated products in most cases.[150,155] The formation of regioisomers with the smaller substrates **68**−**72** is thought to be due to their flexibility upon binding within the active site and hence undefined positioning relative to the catalytic lysine. Notably kynurenine (**67**), a larger substrate that can potentially form more hydrogen-bonding contacts with the active site, could be halogenated in a regio-divergent manner at the 3- or 5-position by PrnA or PyrH and SttH, respectively,[150,155] demonstrating the potential for these biocatalysts to offer a halogenation with catalyst-controlled regioselectivity. Additionally, SttH halogenates the 6-position of *N*-methylated tryptophan (**73**) with good efficiency,[150] while PrnA and RebH do not accept this or other *N*-functionalized tryptophans (**73** and **74**).[219−221]

Halogenases of different natural substrate scope have been explored for the halogenation of alternative aromatic scaffolds. Notably, Rdc2, a phenolic flavin-dependent halogenase, was shown to halogenate the hydroxyisoquinolines **75**−**78** in addition to a number of other macrocyclic lactones (**79**−**82**, Figure 20).[125,167,168,172] In most cases, these highly activated

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

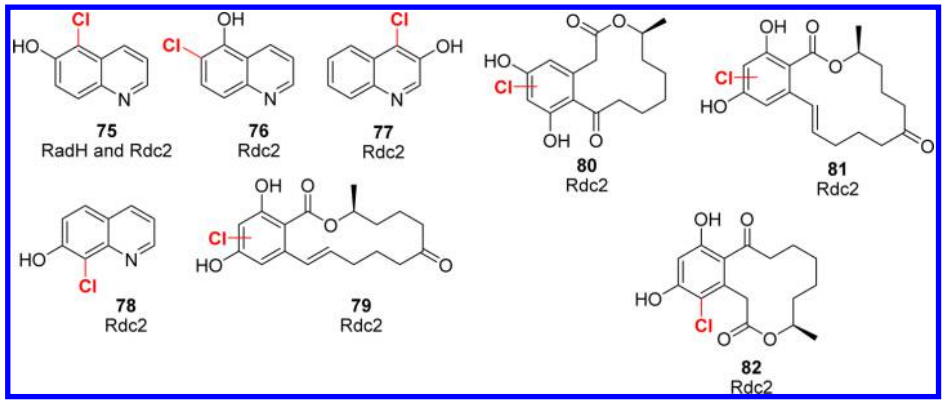


Figure 20. Substrate scope and regioselectivity of the flavin-dependent phenolic halogenases.

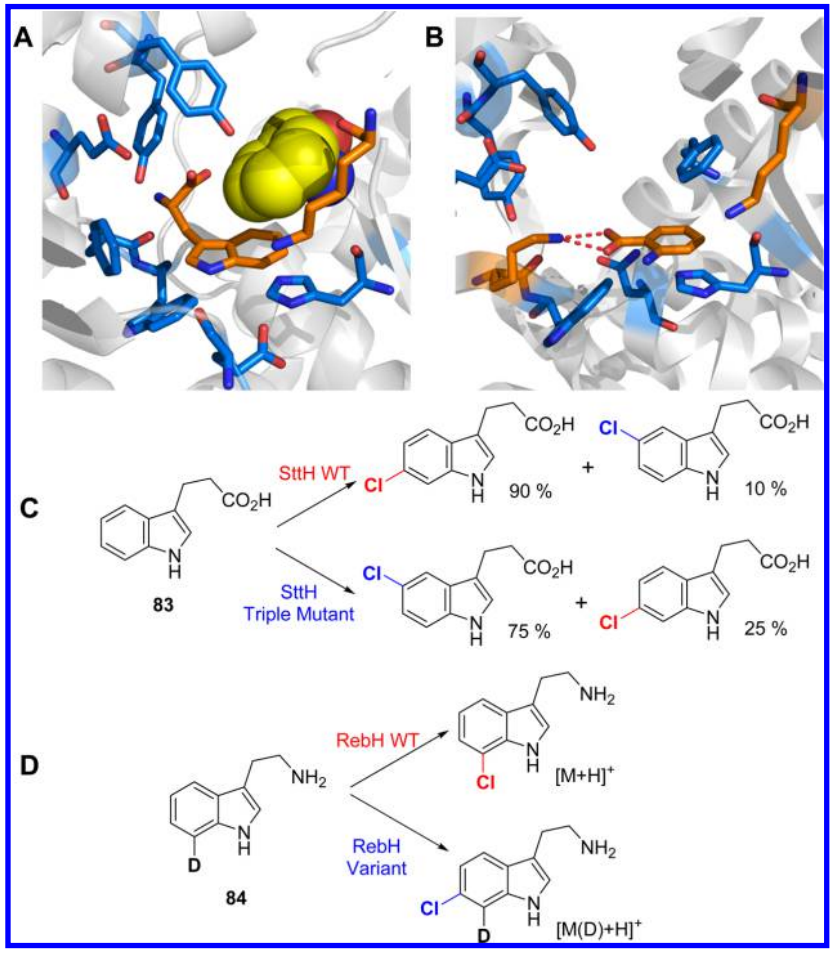


Figure 21. (A) Active site of PrnA with tryptophan bound, with F103 highlighted as yellow spheres showing how this residue prevents approach of catalytic lysine at position 5- (PDB 2AQJ). (B) Active site of PrnA E450 K mutant with anthranilic acid placed showing how the lysine mutation could increase binding affinity of anthranilic acid through hydrogen bonding (PDB 4Z43). (C) Regiodivergent halogenation of indole-3-propinoic acid by SttH and a triple mutant based upon structural differences between SttH and PyrH. (D) MALDI-ToF HTPS of halogenase variant regioselectivity using deuterated substrates.

phenolic compounds are halogenated regioselectively *ortho*- to the hydroxy functionality, suggesting that hydrogen bonding or deprotonation of this hydroxyl may be important in the control of regioselectivity and activity of Rdc2 and related phenolic Fl-Hals.

Each of the flavin-dependent halogenases mentioned above catalyzes the halogenation of an aromatic C−H. Interestingly, the bifunctional halogenase-methyl transferase fusion protein from *Aspergillus oryzae* RIB40 is believed to be responsible for the dichlorination of **54** (Figure 18B).[217] While significant work is likely to be required to understand the mechanism of this aliphatic chlorination and develop this enzyme as a biocatalyst, there is an exciting prospect of regioselective and stereoselective chlorination of alkyl C−H's using this or related enzymes.

4.2.2. Engineering Fl-Hals To Alter Substrate Scope and Regioselectivity. To deliver biocatalysts with improved or altered regioselectivity, a number of attempts have been

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

made using rational mutagenesis to change the regioselectivity of the flavin-dependent tryptophan halogenases.[150,155,222] The first example using PrnA found that mutation of the bulky active site residue F103, which would usually shield the C5 position of tryptophan from approach by active site lysine, to alanine led to the formation of 7- and 5-brominated tryptophan in a 2:1 ratio, while the wild-type afforded only 7-brominated product (Figure 21A).[222] Although a modest shift in regioselectivity, this work set the precedent that site-directed mutagenesis could be used to alter the regioselectivity of the tryptophan halogenases. Subsequent work used active site mutations to stabilize one postulated orientation of anthranilic acid (**68**) in the active site of PrnA and found a single mutant capable of increasing activity toward this substrate evidenced as a change in binding affinity by an almost 10-fold reduction in K_m as compared to wild-type (from 3161 to 384 μM).[155] The same work also identified a double mutant capable of shifting regioselectivity from predominantly 3- to 5-chlorination (Figure 21B), with further improved kinetic parameters (K_m = 205 μM, k_{cat} = 1.82 min^{-1}) as compared to wild-type (k_{cat} = 0.51 min^{-1}) and the single mutant (k_{cat} = 0.93 min^{-1}).[155] A similar approach with SttH found a triple mutant capable of shifting the regioselectivity of halogenation from the 6- to the 5-position of 3-indolepropionate (**83**). Mutations were rationalized on the basis of comparison of the crystal structure of SttH with PyrH, which revealed key differences in a region close to the active site of both.[150] This example is the first to demonstrate that regioselectivity can be switched without altering the catalytic efficiency of the enzyme.

In addition to the rational and structure-guided approaches, random mutagenesis and directed evolution have been used to alter regioselectivity.[223] In this work, substrates with a single deuterium in place of an aromatic C−H were used in a MALDI-ToF-based screen to detect mutants that halogenated at positions other than the deuterated one (Figure 21C). Starting from a RebH mutant previously found to be more thermostable[224] and using a total of six rounds of screening, two RebH mutants were found, which halogenated the 6- and 5-positions of an indolic substrate (**84**) with good selectivity and reasonable conversion.[223] This method does not require structural information about the enzyme and can be used to modulate regioselectivity in a semi high-throughput manner, in contrast to most high-throughput screening (HTPS) methods, which focus on improving stability or activity of an enzyme.

Random mutagenesis has also been used in a "substrate walking" approach to allow the late-stage halogenation of a number of large bioactive substrates such as **85**−**88** (Figure 22).[225] Over a total of four generations of error-prone PCR, RebH was mutated to a quintuple mutant capable of halogenating the large C4-functionalized substrate **87** with complete regioselectivity. Together, these methods demonstrate the potential to develop a suite of biocatalysts for regiodivergent halogenation, either from a single parent enzyme or using halogenases of different natural regioselectivity.

4.2.3. Engineering Fl-Hals To Improve Activity and Stability. While the flavin-dependent halogenases are promising biocatalysts for regioselective and regiodivergent halogenation of aromatic compounds, their applications are still limited to the analytical and semipreparative scale because of low k_{cat} values, poor stability, and susceptibility to substrate−product inhibition.[32−34,218,224,225] This is likely to be an artifact of these enzymes evolving as part of the biosynthetic pathways to nonessential secondary metabolites. Halogenases are not

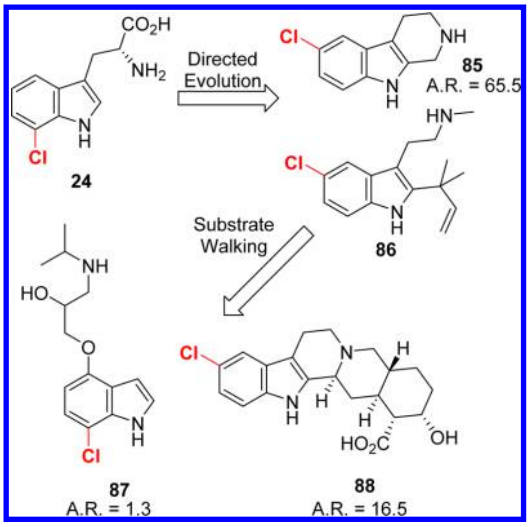


Figure 22. Products from the halogenation of large bioactive natural products by RebH variants generated through directed evolution and substrate walking. A.R. relates to the activity ratio of each substrate with the best variant RebH as compared to wild-type.[225] Substrate **86** showed negligible activity with wild-type, and therefore A.R. cannot be determined.

essential for the survival or growth of the native host, and hence there has been little evolutionary pressure for highly active halogenase enzymes.[226] Although attempts have been made to improve biocatalytic halogenations using Fl-Hals through reaction optimization, design-of-experiment approaches were required due to the large number of variables in such a system, and only modest improvements were obtained.[227] Significant work has therefore been focused on the engineering of Fl-Hals with improved biocatalytic properties.

One strategy to find more stable enzymes, which are therefore usually more amenable biocatalysts, is to look for analogous enzymes found in thermophilic organisms. Thermostable enzyme variants have been demonstrated to confer many advantages, including prolonged lifetime and increased tolerance to organic solvents and proteolysis.[228−231] This approach has been employed with the flavin-dependent halogenases. A thermophilic tryptophan 6-halogenase (Th_Hal) was identified from *Streptomyces violaceusniger* SPC6, a thermophilic and halotolerant bacterium,[232,233] with a melting temperature (T_m) almost 10 °C higher than that of the mesophilic tryptophan 6-halogenase SttH and a higher k_{cat} than a number of other tryptophan halogenases.[145] This thermophilic halogenase (Th_Hal) was partnered with a flavin-reductase from a thermophilic *Bacillus* strain,[234] allowing biocatalytic halogenation reactions to occur at 45 °C in vitro with a number of non-natural substrates. The regioselectivity of Th_Hal was the same as with SttH.[150] A crystal structure of Th_Hal found that 40% of the total differences with the nearest homologue SttH were polar residues on the surface of Th_Hal (Figure 23A), suggesting that increasing surface charge is likely to be largely responsible for the higher stability of Th_Hal.[145] In the case of the naturally evolved thermophilic Fl-Hal, this is likely to be because it is easier to modify the surface of a protein rather than the inner catalytic core whilst retaining activity.[235−237] Increasing surface charge is thought to confer stability by deterring protein aggregation and increasing hydrogen bonding to water.[235,237,238]

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

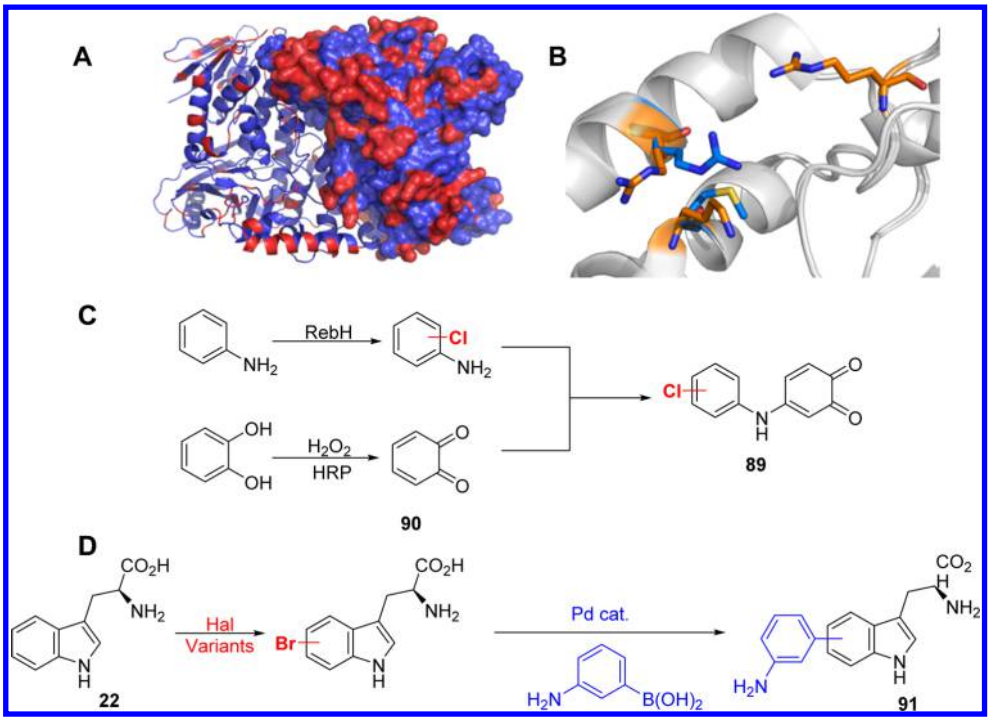


Figure 23. (A) Crystal structure of SttH with differences in amino acid sequence from T_Hal highlighted in red (PDB 5HY5). (B) Crystal structures of wild-type (orange, PDB 2OA1) and 3-SLR (blue, PDB 4LU6) variant RebH overlaid to show some of the amino acid substitutions thought to contribute to increased stability. (C) UV−visible based screen for halogenase activity using either conjugation of aryl halide with *ortho*-quinones. (D) High-throughput halogenase activity screen through combination of enzymatic halogenation with palladium-catalyzed Suzuki chemistry to generate fluorescent adducts.

Other methods have involved using error-prone PCR (ePCR) and directed evolution to mutate a mesophilic enzyme toward increased stability using temperature as the selection pressure.[224] After three rounds of ePCR, two variants with seven and eight mutations possessing T_{opt} and T_m higher than the wild-type starting point were found, both of which proved to be stable over longer reaction times than wild-type, but with decreased turnover. Crystallography of one of these variants (PDB 4LU6) revealed a number of structural features, which may have contributed to increased enzyme stability.[224] Notably, the mutation of a surface glutamine to arginine resulting in increased surface charge, similar to the thermophilic Th_Hal.[145,238] In other portions of the RebH variant, charge density was reduced to reduce repulsion with nearby residues (Figure 23B), and an N-terminal serine was mutated to proline and is believed to increase protein rigidity at this terminus and hence improve stability.

To achieve full synthetic utility, it is likely that the flavin-dependent halogenases will need significantly more engineering than the reported low-throughput chromatographic methods of screening can offer in a reasonable time frame with economical resource comsuption.[224,225] For this purpose, a number of high-throughput screens for halogenase activity have been developed (Figure 23C and D).[239−241] One such example relies upon the formation of aryl chloride-*ortho*-quinone adducts (**89**) with distinct UV/visible absorbance properties.[241] The generation of such adducts requires HRP (horseradish peroxidase) in addition to the halogenase and cofactor recycling enzymes to generate the *ortho*-quinone **90** from catechol. Michael addition of the aryl chloride and spontaneous reoxidation then generates the probe of interest (**89**, Figure 23C). Although this method does allow reliable monitoring of halogenase reactions using UV−vis in 96-well plates rather than chromatographic methods, the need to quench reactions and use additional enzymes and substrates may limit its applicability. In a similar vein, a recent report of a fluorescence-based screen involves the combination of the enzymatic halogenation with palladium-catalyzed cross-coupling chemistry to create an adduct (**91**) with fluorescence sufficiently distinct from the cofactors and proteins required for biocatalytic halogenation, thereby allowing detection in a high-throughput manner using a fluorescence plate-reader (Figure 23D).[240] After applying this method to screen variants from error-prone PCR, a double mutant of Thal was found with a T_m 10 °C higher than that of wild-type. The structural differences thought to be responsible for stabilization of the engineered and naturally occurring halogenase enzymes, as well as the methods used to discover them, could be used as the basis of engineering halogenase variants with further improved activity. The combination of these screening methods with strategies for targeting the generation of variant libraries to specific sites, as have been demonstrated to afford a minor improvement in RebH catalytic parameters by targeting mutations to the flavin-binding pocket,[242] could be used to focus such engineering efforts.

The stability and scalability of a number of Fl-Hals have been improved by their immobilization into heterogeneous cross-linked enzyme aggregates (CLEAs) of a Fl-Hal, flavin reductase, and alcohol dehydrogenase, for concurrent cofactor regeneration,[150,243,244] allowing enzymatic halogenation on a gram scale.[244] CLEAs are well-known to improve the efficiency of biocatalysts by improving catalyst lifetime, reducing the effects of substrate/product inhibition, and allowing more efficient biocatalyst removal and recycling.[245−247] In the case of the Fl-Hals, this cross-linking may help to stabilize interdomain interactions, or effectively "protect" some of the biocatalyst from high substrate concentration by shielding it within the

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

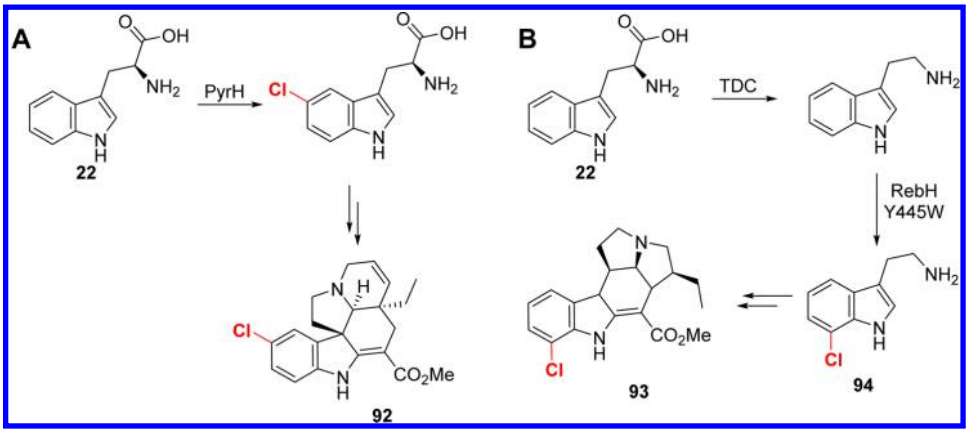


Figure 24. Integration of flavin-dependent halogenases into various biosynthetic pathways to yield chlorinated natural products. (A) Use of PyrH to afford 5-chlorinated indolocarbazoles. (B) Use of an engineered RebH variant to selectively catalyze 7-chlorination of tryptamine. TDC = tryptophan decarboxylase.

heterogeneous catalyst.[244,248] The preparation of such biocatalysts from crude cell lysates, rather than needing to purify enzymes,[244] is also advantageous.

4.2.4. Integration of Fl-Hals into Non-Native Biosynthetic Pathways. There are a number of examples where the Fl-Hals have been incorporated into natural product biosynthetic pathways, to increase the diversity of the products obtained and therefore attenuate bioactivity of the resulting compounds (Figure 24).[33,249−254] Generation of natural product analogues via this method is significantly more efficient, and therefore economical and ecologically friendly, than via laborious total synthesis.[255] A number of these reported examples rely upon the chlorination of tryptophan prior to incorporation into a biosynthetic pathway, thereby resulting in a natural product with a chlorinated tryptophan moiety.[249,251,252] The seminal work in this area involved the combinatorial reconstitution of genes from the biosynthesis of rebeccamycin and staurosporine (natural product antitumor agents) with genes encoding for tryptophan halogenases of different regioselectivity to afford different regioisomers of chlorinated indolocarbazole "non-natural" products (**92**, Figure 24A).[252] Subsequent work involved the introduction of a tryptophan-7-halogenase directly into a pacidamycin-producing *Streptomyces* strain, previously found to incorporate 7-halo tryptophans through substrate-directed biosynthesis,[256] to afford chlorinated pacidamycin.[251] Halogenase genes have also been incorporated into plant secondary metabolism to produce chlorinated monoterpene indole alkaloids (**93**) in *Catharanthus roseus* root cultures.[249] In this case, it was found that the tryptophan decarboxylase enzyme, in plants, was inefficient at turning over 7-chlorotryptophan to give the required 7-chlorotryptamine (**94**). Therefore, a RebH mutant capable of selectively catalyzing the chlorination of tryptamine was introduced to overcome this metabolic bottleneck and reduce accumulation of 7-chlorotryptophan (**24**), which has adverse effects on plant health (Figure 24B).[250] Other groups have since expressed the tryptophan halogenases SttH and RebH in the chloroplasts of tobacco plants, allowing the production of 6- and 7-halogenated tryptophans and, when partnered with a tryptophan decarboxylase, chlorinated tryptamines (**94**).[257]

The ability of Rdc2 to catalyze the late-stage chlorination of a number of macrolactones, as well as smaller phenolic compounds, suggests it would be an ideal candidate for introduction into non-indolic secondary metabolite biosynthetic pathways.[125,167,168] Rdc2, for example, has been incorporated into a reconstituted resveratrol biosynthetic pathway in *E. coli* to afford production of the 2-chlorinated resveratrol derivative, which has been shown to have increased antimicrobial and antioxidant activity as compared to the nonhalogenated parent compound.[258,259] CazI was recently demonstrated to be a flavin-dependent halogenase responsible for chlorination during the biosynthesis of chaetoviridin,[179] suggesting potential for the generation of halogenated polyketide non-natural products.

4.2.5. Integration of Fl-Hal with Transition-Metal Catalysis. The introduction of halogen atoms into complex natural products also provides a convenient moiety for further functionalization by transition-metal-catalyzed cross-coupling chemistry.[260] The number of C−C, C−N, C−F, and C−O bond-forming reactions possible using this chemistry has made it an indispensable tool in the synthesis of complex molecules of relevance in many sectors of the chemical industry.[3] As the carbon−halide bond is a key substrate for these transformations, it seems logical that the combination of the regioselective enzymatic methods for installation of halides onto aromatic substrates with this powerful cross-coupling chemistry may provide methodologies for the regioselective formation of C−C, C−N, C−F, and C−O bonds from unactivated C−H bonds. Current methods for such transformations, termed C−H activation, rely upon substrate control whereby either the difference in acidity of C−H bonds or intramolecular coordination is used to control the position of functionalization.[42] As such, this often means that certain positions of key moieties are inaccessible. Use of a halogenase, whereby the position of functionalization is controlled by coordinating effects between the substrate and a well-defined three-dimensional active site, however, represents an example of catalyst control as using different biocatalysts can functionalize different positions.

The sequential application of biosynthetic halogenation followed by a subsequent transition-metal-catalyzed cross-coupling was demonstrated on crude extracts of chloropacidamycin from *S. coeruleorubidus* expressing the tryptophan 7-halogenase PrnA to generate a number of 7-aryl pacidamycin derivatives through Suzuki−Miyarua chemistry.[251] A similar rationale, whereby halo-aryl containing crude extracts from halogenase biotransformations are used for Pd-catalyzed cross-

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

coupling chemistry, has been extended to allow the regioselective formation of C−C, C−N, and C−O bonds on a number of bioactive molecules (Figure 25).[261]

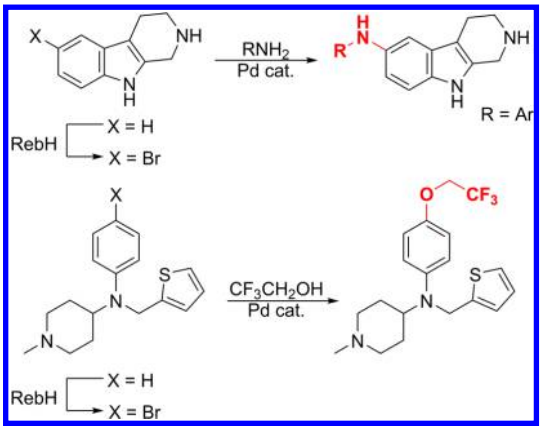


Figure 25. Integration of flavin-dependent halogenases with palladium-catalyzed cross-coupling chemistries to allow regioselective C−C and C−O cross-coupling reactions.

Although these two examples demonstrate the possibility of regioselective cross-coupling reactions using halogenases to generate an aryl halide precursor, the requirement for a two-step process including isolating the intermediate aryl halide is not ideal. The combination of multiple reactions into single-pot transformations represents a potential step-change in the efficiency of chemical synthesis by reducing solvent consumption and waste, eliminating the need for auxiliary chemicals and increasing space-time yield.[262,263] The combination of biocatalytic and chemocatalytic reactions into single pot aqueous transformations is often not straightforward, however, and requires the compartmentalization or removal of at least one of the components due to mutual deactivation of the bio- and chemocatalysts.[264−266] This has been realized to some extent by using the heterogeneous CLEAs of the flavin-dependent halogenases (Figure 26).[243,267] As the heterogeneous biocatalyst is easily removed from the reaction, compartmentalization of biocatalysts and chemocatalysts can be achieved by filtration.[243,267] Methods where such intermediary processing is not required have also been reported.[243] Here, compartmentalization is achieved by using poly dimethylsiloxane (PDMS) membranes, which, due to their hydrophobic nature, allow only the nonpolar arylhalide to diffuse freely between compartments, while the charged chemocatalysts and biocatalysts are contained separately.[264,268,269] Such one-pot transformations have been applied to the arylation, heteroarylation, and alkenylation of isoquinolines and benzamides, in addition to the 5-, 6-, and 7-positions of indoles and tryptophan derivatives (Figure 26).[243,267] Notably, the direct functionalization of the 7- and 6-positions of indoles by nonenzymatic means requires the introduction and removal of directing groups.[270,271] The methods of compartmentalization used here set the stage for the integration of other transition metal-catalyzed processes and therefore additional regioselective transformations.

4.3. Flavin-Dependent Monooxygenases

4.3.1. Natural Occurrence and Mechanism of Flavin-Dependent Monooxygenases. The flavin-dependent halogenases belong to the class of flavin-dependent monooxygenases, which operate via the activation of molecular oxygen using flavin cofactors. A number of other types of enzyme in this class have also been subject to significant interest as biocatalysts because of their ability to functionalize inert C−H bonds.[115,272−275] A number of recent reviews have extensively discussed these enzymes in terms of both structural biology and their applications as biocatalysts.[115,116,275−277] This Review will discuss how the methods employed to allow the application of some of these enzymes as industrial biocatalysts might be applied to improve the viability of using flavin-dependent halogenases in the same way.

The flavin-dependent hydroxylases and epoxidases in addition to the Bayer−Villager monooxygenases (BVMO)

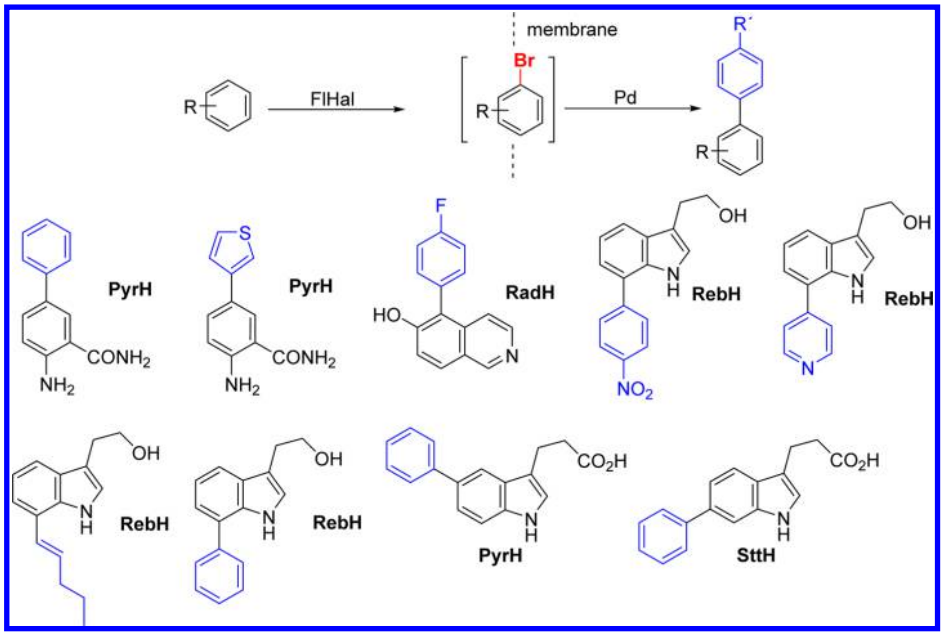


Figure 26. Integration of flavin-dependent halogenases with palladium-catalyzed cross-coupling chemistries to allow regioselective and regiodivergent arylation, vinyl-ation, and heteroarylation in one-pot reactions using membrane compartmentalization.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

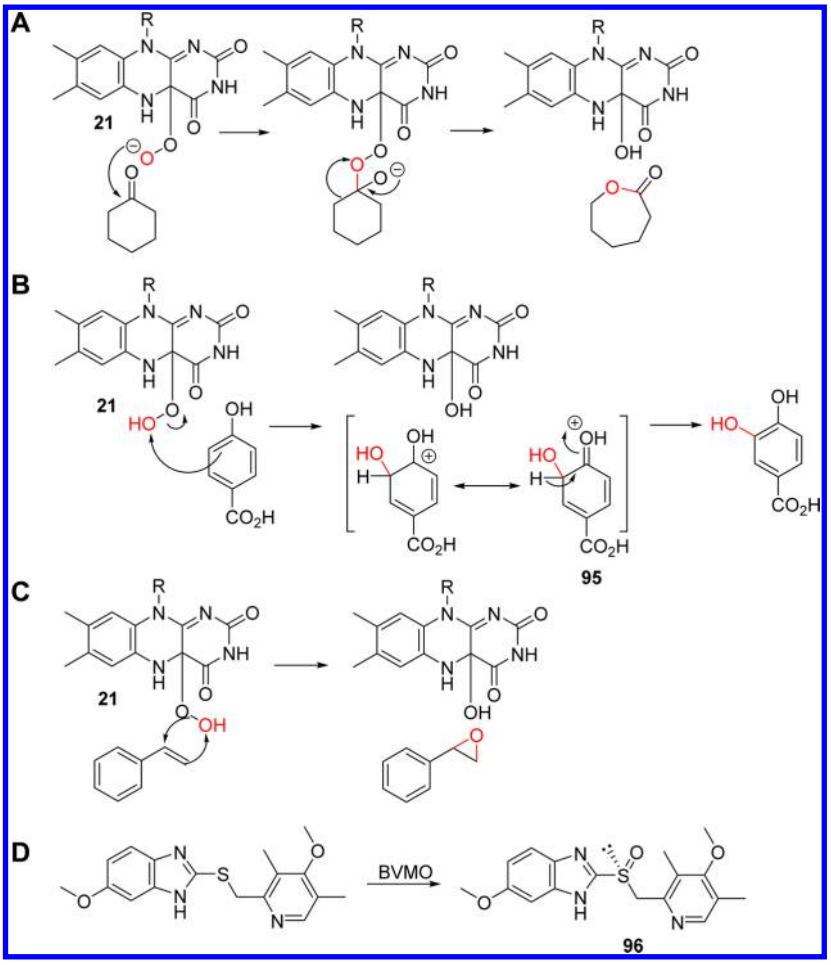


Figure 27. Reactivity of hydroperoxy-flavin (**34**) in (A) Bayer–Villager monooxygenases, (B) flavin-dependent hydroxylases, and (C) styrene monooxygenases. (D) Use of a Bayer–Villager monooxygenase for asymmetric sulfoxidation in the synthesis of esomeprazole (**96**).

operate via an analogous mechanism to Fl-Hals, except hydroperoxy-flavin (**21**) is believed to directly react with substrate to yield oxygenated product.[118,120,278] All of these enzymes therefore share the conserved flavin-binding motif GxGxxG. With the flavin-dependent hydroxylases and BVMOs, reduced flavin is found in the form of an FMN prosthetic group, which is reduced by a second domain using NADH, while the epoxidases utilize freely diffusing $FADH_2$ generated by a separate flavin-reductase in the same way as the Fl-Hals. Because of the mechanistic relationship between the Fl-Hals and the hydroxylases, epoxidases and BVMO, useful insight may be gained as to how to apply the Fl-Hals as industrial biocatalysts by studying others in the class.

The flavin-dependent hydroxylases typically hydroxylate the 2-position of phenols,[272,279] most likely due to increased stabilization of intermediate **95** and positioning relative to hydroperoxy-flavin (**21**). These enzymes are responsible for the catabolism of xenobiotics as well as the biosynthesis of fatty acids and sterols.[274,280−282] Kynurenine-3-monooxygenase, on the other hand, hydroxylases *ortho*- to a NH_2 group, due to stabilization of the analogous ketamine intermediate.[283] This particular hydroxylase has attracted much academic attention due to its involvement in tryptophan catabolism and neurodegenerative disease.[283−285] Styrene monooxygenase is similarly involved in the microbial catabolism of styrene, catalyzing epoxidation of the vinyl group prior to ultimate incorporation to the TCA cycle,[286] while BVMOs are responsible for the oxygenation of carbonyl-containing xenobiotics en route to innocuous products in addition to oxidation at heteroatoms such as S, N, and B.[287,288]

The different mechanisms of flavin-dependent monooxygenases serve to illustrate the versatility of the flavin cofactors (Figure 27). Notably, in BVMOs, the distal oxygen of hydroperxoy-flavin is deprotonated and functions as a nucleophile, in contrast to halogenases, hydroxylases, and the styrene epoxidase, where the distal oxygen is protonated and functions as an electrophile.

4.3.2. Flavin-Dependent Monooxygenases as Biocatalysts. There are a number of factors that limit the prospect of applying the flavin-dependent monooxygenases, including Fl-Hals, as industrial biocatalysts, including the requirement for stoichiometric reductants, substrate or product inhibition, the requirement for careful control of oxygenation, and poor enzyme stability.[272,275]

Reduced flavin cofactors, required by all enzymes of this class, are unstable with respect to oxygen. Stoichiometric reductants must therefore be provided in the form of nicotinamide cofactors, which are more aerobically stable. With the hydroxylases and BVMOs, reduction of the flavin prosthetic group is carried out by a reductase domain of the monooxygenase.[116,289,290] With the epoxidases and Fl-Hals, freely diffusing reduced flavin is typically produced by a separate flavin-reductase enzyme before utilization by the monooxygenase,[116] with the exception of a recently discovered

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

epoxidase-reductase fusion protein.[291,292] In the case of monooxygenases utilizing freely diffusing reduced flavin cofactors, it is possible for reduced flavin to react directly with oxygen in solution, resulting in hydroperoxy-flavin being generated outside of the active site.[293] Without stabilization through hydrogen bonding to an active site, this can break down to produce H_2O_2 in solution (Figure 28).[293,294] In

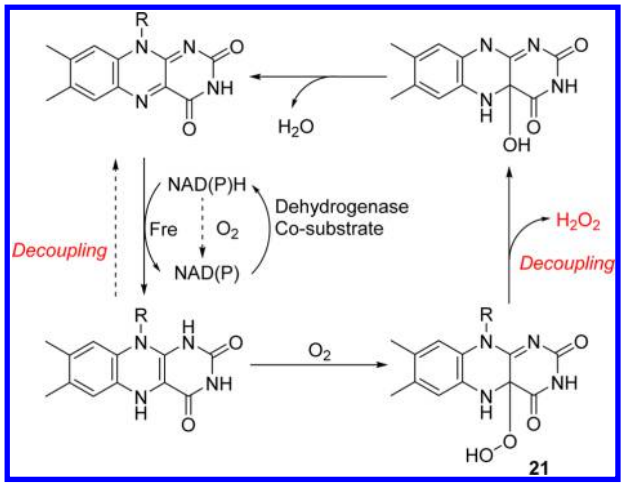


Figure 28. Cycle of peroxy-flavin (**21**) production showing potential uncoupling processes.

addition to "uncoupling" the monooxygenase reaction and wasting the reductant cosubstrate, accumulation of such reactive species can impair the stability of the biocatalysts and therefore reduce their productivity.[275] Similar uncoupling reactions can also occur within the active site, including in those of enzymes that utilize flavin prosthetic groups, by inefficient transfer of oxygen from hydroperoxy-flavin to substrate and therefore degradation to H_2O_2 and oxidized flavin. High concentrations of product are thought to promote uncoupling in this manner, by effectively blocking substrate from binding and preventing productive oxygen transfer.[279,295] Such uncoupling means that the reducing cosubstrate (NAD(P)H) must either be supplied in excess, which is undesirable due to cost, or regenerated in situ from inexpensive cosubstrates. Although oxygenation can promote uncoupling and impair productivity, it cannot simply be excluded from reaction media due to the requirement to generate hydroperoxy-flavin for productive cycles; consequently, careful control of oxygenation is often required.[296,297]

The use of additional enzymes in vitro to reduce nicotinamide cofactors using inexpensive cosubstrates such as glucose or isopropanol is widespread and has been applied to the Fl-Hals, in addition to a number of other monooxygenases, to allow the use of substoichiometric nicotinamide and flavin cofactors.[155,220,298] Methods for the direct regeneration of FAD, without using nicotinamide cofactors or additional enzymes, have also been reported and applied with the flavin-dependent monooxygenases (Figure 29).[128,299−303] These methods offer simplified electron-transfer cascades and therefore are envisioned to be a more robust alternative to enzymatic methods. A number of these examples utilize the organometallic complex $[Cp^*Rh(bpy)(H_2O)]^{2+}$ and formate to generate the reduced complex $[Cp^*Rh(bpy)H]^+$, which can then directly transfer hydride to oxidized FAD.[128,302] The use of transition metal complexes may not be ideal, however, because of their potential

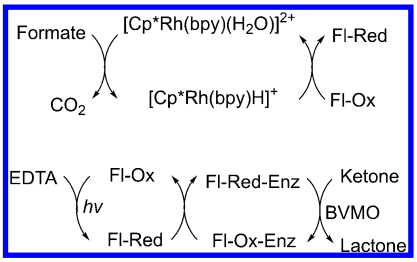


Figure 29. Nonenzymatic flavin recycling methods.

deactivation of the monooxygenase and expense. Methods of light-driven flavin reduction have therefore also been investigated, utilizing either water or EDTA as the sacrificial electron donor.[300,304] In one case, this method has been applied to the BVMO-catalyzed oxidation of cyclic ketones by using additional FAD in solution, to facilitate electron transport from EDTA to enzyme-bound FAD.[300] Further simplification has been achieved using electrochemical methods to directly reduce FAD for biocatalytic vinyl aromatic epoxidations without the need for nicotinamide cofactors.[301] Direct regeneration of $FADH_2$ using nicotinamide mimics has also been reported, which allows efficient biocatalytic oxidations using styrene monooxygenase without the flavin-reductase domain or nicotinamide cofactors.[303]

The inherent metabolism of a microbial host can also be used to regenerate the required cofactors, in addition to supplying enzymes such as superoxide dismutase or catalase, which can degrade reactive oxygen species.[305] Such an approach has the advantage of not requiring additional enzymes or cosubstrates, as well as ease-of-preparation of the biocatalyst when compared to pure protein. This was achieved to good effect with the flavin-dependent hydroxylase HBP1.[306,307] In these examples, a number of substituted phenols were successfully converted to catechols in gram amounts by using growing *E. coli* expressing the hydroxylase of interest.[306,307] The use of growing microbial host not only means that cofactors can be recycled by the inherent primary metabolism of the organism, but also that the biocatalyst is constantly being synthesized, which is useful in the case of particularly unstable enzymes.[305−307] The substrate phenols and subsequent catechols were found to be toxic to *E. coli* in fairly low concentrations, however, and therefore batch-feeding of phenol at or below the rate of bio-oxidation was used to limit accumulation of substrate. In combination with in situ product removal (ISPR) to adsorb catechol product onto a solid resin, the accumulation of substrate and product was limited to subtoxic levels, allowing constant bio-oxidation using the same whole cells.[306,307] The combination of controlling both substrate and product concentration is often termed substrate-feeding product-removal (SFPR). A similar rationale has been applied to the BVMOs, with a number of examples allowing between gram and kilogram production of enantiopure lactones from ketones.[308−314] In these examples, substrate and product are allowed to adsorb and desorb freely from a resin, therefore limiting their concentration in the reaction broth. Additionally, specially designed reactors are used to control the extent of oxygenation and therefore limit the effects of uncoupling, while allowing sufficient oxygenation for bio-oxidation.[308,314]

Biphasic reaction media can also allow the concentration of substrates and products to be controlled and therefore limit toxicity or inhibition by high concentrations of either. A water-immiscible organic cosolvent is used to create a substrate and

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

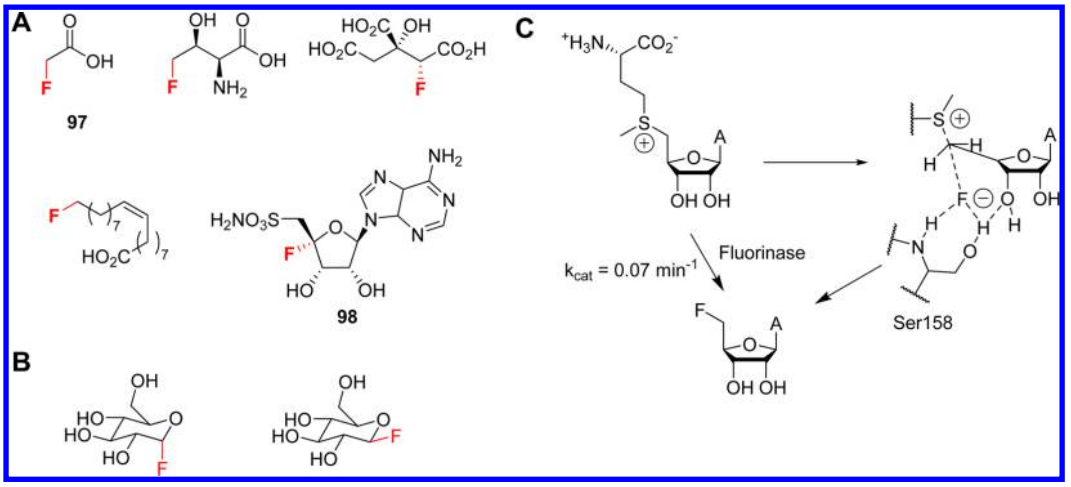


Figure 30. (A) Fluorinated natural products identified to date. (B) Fluorinated glycosides produced by mutants of glycosyl transferase enzymes. (C) Proposed mechanism of F^- binding to fluorinase and arrangement for S_N2 displacement. Kinetic parameters shown refer to the fluorination of SAM to FDA by the fluorinase from *Streptomyces cattelya*.[349]

product reservoir, away from the biocatalyst. This has been successfully achieved with both the styrene monooxygenases and the BVMOs,[315,316] in one case allowing production of up to 388 g of styrene oxide.[317]

In addition to the methods of engineering monooxygenase-mediated biotransformations to overcome substrate/product inhibition, toxicity, oxygenation, and uncoupling discussed above, significant work has focused on engineering at the protein level and is responsible for some of the highly productive biocatalytic methods known. For example, the BMVO used by Codexis for sulfoxidation in the production of esomeprazole (**96**, Figure 27D) is a cyclohexanone monooxygenase (CHMO) variant containing 41 mutations.[318] Modulating activity of the BVMO in terms of substrate scope or enantioselectivity has been achieved using both rational (structure-guided) mutagenesis and directed evolution approaches. A widely used approach for this class is the "Complete Active Site Saturation Test" (CASTing) whereby amino acid positions close to the binding pocket are selected for saturated mutagenesis, and has allowed the evolution of BVMO variants with modified substrate scope and improved enantioselectivity using homology models as a guide for targeting saturated mutagnesis.[319−323] To facilitate directed evolution by allowing rapid screening of larger variant libraries, a number of high-throughput screens for BVMO and hydroxylase activity have also been developed. The majority of the most recent methods require further reaction of the monooxygenase product, either enzymatically or nonenzymatically, to either induce a pH change[324] or afford a compound with distinct UV−vis characteristics, which can then be easily detected.[295,325−331] A number of structure-guided approaches have also been used to modulate regioselectivity and substrate scope,[332,333] or improve oxidative and thermostability.[334]

5. FLUORINASES

5.1. Discovery of Fluorinases in Nature

The enzymes responsible for biosynthesis of fluorinated natural products have remained relatively elusive, due in part to their rarity as compared to naturally occurring organochlorine and organobromine compounds.[335−338] Indeed, only five fluorinated natural products have been definitively identified to date (Figure 30A), with fluoroacetate (**97**) being the most ubiquitous.[339] A recent review has thoroughly discussed the enzymes involved in the biosynthesis of these compounds;[336] however, they will be discussed herein briefly to complete the survey of biocatalytic halogenations, which might be applied to industrial or medical processes.

Enzymatic fluorination activity was first observed with mutants of the glycosyl transferase enzymes *β*-glucosidase and *β*-mannosidase from *Agrobacterium* sp. and *Cellulosmonas fimi*, respectively.[340−342] It was found that when catalytically essential nucleophilic residues were mutated, glycosyl transferase activity was abolished. Upon addition of high concentrations of fluoride, however, such activity was restored. NMR subsequently showed the production of fluorinated glycosides (Figure 30B) as intermediates, which would allow attack by nucleophilic moieties of incoming sugar residues due to the polarized nature of the C−F bond.[340−342]

Nucleocidin (**98**), an antibiotic produced by *Streptomyces calvus*, was one of the first fluorinated metabolites to be identified, although elucidation of its biosynthetic pathway was hampered by poor fluorometabolite production under laboratory fermentation conditions.[343] *Streptomyces cattelya*, which secretes fluoroacetate and 4-fluorothreonine,[344] however proved to be sufficiently practicable to allow identification of the enzyme responsible for fluorination.[345] This particular fluorinase has been shown to catalyze fluorination of S-adenosyl methionine (SAM), to generate 5′-deoxy-5′-fluoroadenosine (FDA) in an S_N2-type reaction (Figure 30C), and has informed much of the current mechanistic and functional understanding of these fluorinases.[335,346−351] The mechanism of enzymatic fluorination is particularly interesting because the high electronegativity of fluoride means that oxidation to an electrophilic species, as is commonplace with the other classes of halogenase (vide infra), would not be possible.[336] Moreover, tight solvation of F^- in water decreases its nucleophilicity and creates a great energetic penalty for nucleophilic substitution with an aqueous fluoride source.[336,352]

Crystallography of the *S. cattelya* fluorinase revealed that fluoride ion is bound in the active site through electrostatic interactions, positioned in a S_N2-like trajectory to the ribose 5′-C,[348] consistent with the observation that introduction of fluoride occurs with inversion of stereochemistry using isotopically labeled substrates.[346,347] The energetic penalty for

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

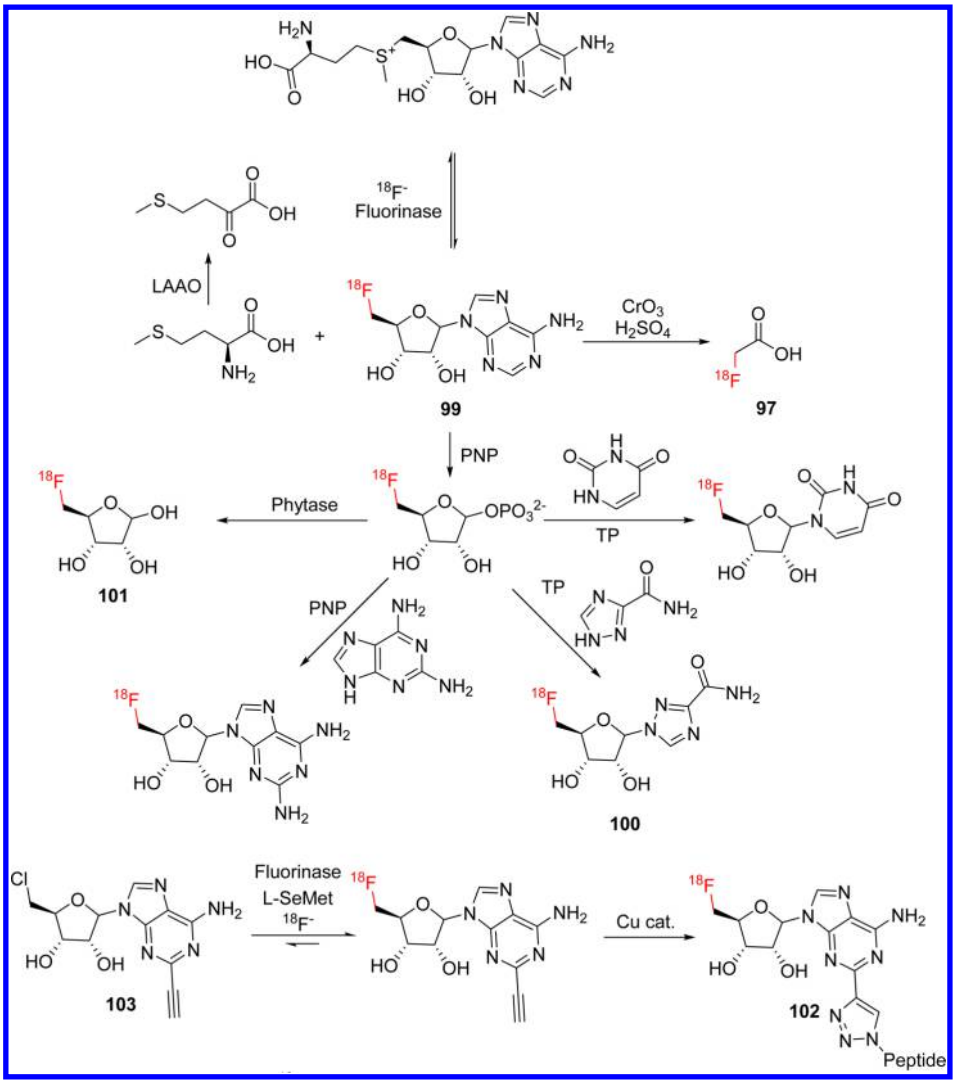


Figure 31. Preparation of ^{18}F labeled nucleotide analogues, fluoroacetate, and peptides using fluorinases in biobio and biochemo cascade reactions. LAAO = L-amino acid oxidase. PNP = purine nucleotide phosphorylase. PyNP = pyrimidine nucleotide phosphorylase.

desolvation of fluoride is thought to be compensated for by the retention of four hydrogen-bonded waters in the active site, in addition to the increased binding affinity of SAM.[336,348−350,352]

Since the seminal work on this fluorinase, a number of other enzymes have been identified with very high homology to the fluorinase from *S. cattelya*.[353−355] *Streptomyces* sp. MA37 was found to produce fluoroacetate and 4-fluorothreonine in culture, and subsequently a putitative fluorinase gene with 87% sequence identity to the fluorinase from *S. cattelya* was identified.[353,354] Expression of this enzyme in *E. coli* confirmed fluorinase activity in vitro. Potential fluorinases have also been identified in *Nocardia brasilensis* and *Actinoplanes* sp. N902-109, and similarly demonstrated to have fluorinase activity in vitro.[353,354] The marine organism *S. xinghaiensis* has been found to produce fluoroacetate, although in vitro activity of the fluorinase identified in this organism is yet to be demonstrated.[355]

A chlorinase, related mechanistically to the fluorinases, has also been identified in *Salinispora tropica*, which is responsible for chlorination during the biosynthesis of salinosporamide A (Figure 32).[356−358] Interestingly, this chlorinase cannot utilize fluoride as nucleophile in place of chloride, while the fluorinase from *S. cattelya* is capable of utilizing both halides.[359] This selectivity is thought to be due to differences in active site organization around halide binding between the fluorinase and chlorinase.[360,361] Similarly, a putative fluorinase from *Pyrococcus horikoshii* was found to utilize hydroxide (from water) as nucleophile rather than either chloride or fluoride in vitro, affording adenosine and L-methionine.[362,363] Again, crystallography revealed significant differences in the way in which the nucleophile is bound to the active site.[363]

5.2. Application of Fluorinases as Biocatalysts

5.2.1. ^{18}F Labeling. The radioisotope ^{18}F is commonly used in positron-emission tomography (PET), a key medical imaging technique.[336,364−366] The ^{18}F required to produce the compounds used in this process is obtained from bombardment of $H_2{}^{18}O$, which affords aqueous $^{18}F^-$.[336] The potential to use fluorinases to prepare ^{18}F labeled compounds, which can use aqueous fluoride directly without the need for drying or coordination to specialist ligands to allow reactivity in organic solvents, could therefore yield expeditious routes to these compounds in higher radiochemical yield (RCY).[336] A great deal of the work on applying fluorinases for biocatalysis has therefore focused on this area (Figure 31).

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

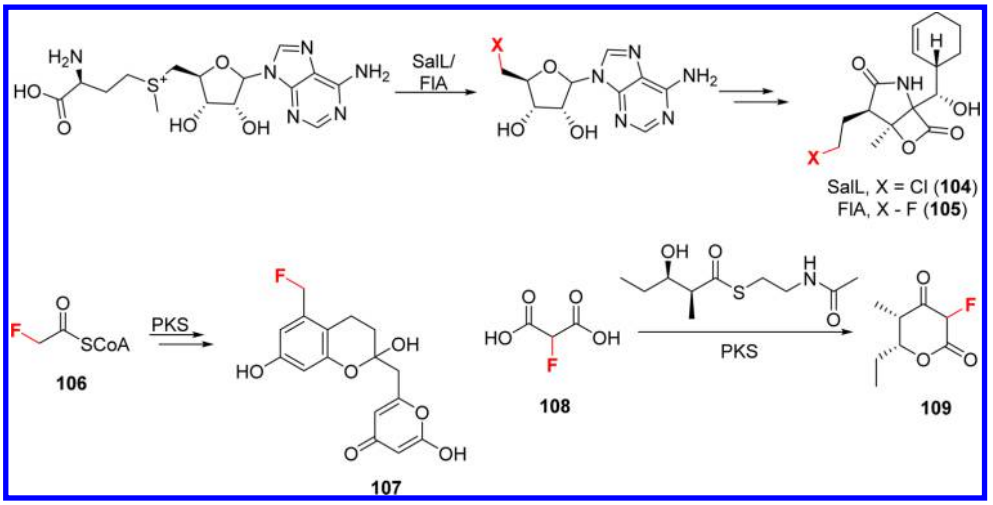


Figure 32. Preparation of fluorinated natural product analogues through either substrate-directed biosynthesis of fluorinated building blocks or heterologous expression of fluorinases.

Early work found that using $^{18}F^-$ equivalents could afford ^{18}F-FDA (**99**) using L-AAO (L-amino acid oxidase) to oxidize L-methionine and therefore suppress the reverse reaction.[367−369] Various other radionucleotides have been generated using the above fluorinase-mediated ^{18}F introduction, followed by "base-swapping" with PNP (purine nucleotide phosphorylase) or PyNP (pyrimidine nucleotide phosphorylase) enzymes to allow introduction or other purine or pyrimidine bases.[369,370] Interestingly, this method has also been applied to the one-pot synthesis of a fluorinated analogue of the antiviral agent ribavirin **100** from SAM.[369] A polymer-supported fluorinase has also been used for these radiofluorinations which, by allowing more facile removal of biocatalyst, may allow more efficient syntheses of these compounds with higher RCY.[371,372] The ^{18}F-FDA nucleoside produced using fluorinase can also act as the substrate for a Kuhn−Roth oxidation, allowing the chemo-enzymatic synthesis of ^{18}F fluoroacetate (**97**), a common PET-imaging agent in neurology and oncology.[365,366,373]

Combination of the above cascade with a phytase enzyme allows the nucleobase portion of ^{18}F-FDA to be replaced with a hydroxyl group, affording 5′-^{18}F ribose (**101**).[370,374] **101** is of particular interest because of the potential to exploit this primed substrate in bioconjugation due to its propensity for ring-opening.[375,376] Bioconjugation has also been achieved by using nucleotide mimics containing chemical handles for conjugation onto a peptide target.[377,378] It was noted by crystallography that H2 of the adenine ring of SAM was positioned toward the surface of the enzyme, suggesting that analogues with substitution here may be accepted by the fluorinase.[378] Introduction of an acetylene group allows facile attachment of a tagged peptide to 5′-chloro adenosine using an azide−alkyne cycloaddition "click" reaction. Because of the orientation of this substitution, the reverse chlorinase reaction of fluorinase could then be used to afford fluorination of the adenosine portion, affording ^{18}F-labeled peptide **102**, a good PET imaging agent in rats.[378] A simplified protocol, starting from "untagged" nucleoside **103**, was subsequently reported, which allowed the preparation of the same agent in 90% RCY.[377]

5.2.2. Production of Fluorinated Natural Products. Natural products are a common source of medicinal compounds. Given that fluorination can often have a profound effect upon the bioactivity of a compound,[13,14] methods for the facile preparation of fluorinated natural product analogues through fermentation are therefore desirable.[336] In an attempt to achieve this goal, fluorinase enzymes have been incorporated into non-natural biosynthetic pathways (Figure 32).

A seminal example employing this rationale involved replacing the gene encoding salL with the fluorinase encoding gene flA in the salinosporamide (**104**) producer *Salinospora tropica*. The resulting mutant produced fluorosalinosporamide (**105**) directly by fermentation.[379,380] Fluorinated polyketide natural product analogues have also been obtained from substrate-directed biosynthesis. Through incubation of the reconstituted minimum components of the artinochodin PKS machinery with fluoroacetyl-CoA (**106**), a fluorinated octaketide could be produced, which is then subsequently cyclized to **107**.[381] R_2CHF groups have also been introduced into polyketide natural products by feeding of fluoro-malonate.[382] It was demonstrated that a malonyl-CoA synthetase could also accept fluoromalonate **109** which, when combined with part of the assembly module of the erythromycin PKS, allowed generation of the fluorinated polyketide analogue **110**.[382] The ability to produce fluorinated substrates for PKS machinery, from fluoroacetate or similar starting materials generated by integration of fluorinases in vivo,[383] has the potential to allow the facile synthesis of fluorinated natural product analogues through fermentation.

6. FUTURE PERSPECTIVES

The enzymes discussed herein show great versatility for regio- and stereocontrolled halogenations of unactivated C−H bonds. Such transformations, affording highly bioactive compounds and ubiquitous synthetic intermediates, could have a great impact on the efficiency of synthesis in all sectors of the chemical industry. Given the importance of selectivity in these transformations, those classes of halogenases that utilize freely diffusing hypohalous acid intermediates (vide infra) are unlikely to afford transformations of great value, with the exception of the limited number of examples that appear to show some selectivity.[76] The α-KG and flavin-dependent halogenases, which show exquisite control of their active halogenation species and hence afford regio- and stereoselective transformations, are the most promising. Together with fluorinases, the only class capable of introducing fluorine to organic

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

compounds, these classes of halogenase will likely attract the most attention for industrial biocatalytic applications.

Despite this potential, however, there is still some way to go before these enzymes can be reliably used on an industrial scale. The fluorinases have been applied on a laboratory scale for the ^{18}F-labeling of small molecules and for the production of fluorinated natural product analogues,[367−369,373,374,377−380] although much research is still focused on the discovery of additional flourinases, and little attempt has been made to determine the scalability of these transformations. Immobilization of the fluorinases, which would allow more facile removal of biocatalyst and therefore potentially higher RCY, in addition to higher purity products for medical applications and more productive enzyme preparations, is a promising direction.[371] Unlike the other classes of halogenase discussed herein, this class ultimately requires a methionine leaving group to be installed, which fluoride can then displace. Although this can be transiently installed by using the reverse chlorinase reaction of the fluorinases,[377] the need for prior functionalization could limit the efficiency of synthesis using these enzymes. The unique ability of this class to produce organofluorine compounds, unlike the other classes which require oxidation of halide to an activated species prior to chlorination, means that the transformations involving the class will still be of great value.

The Fe(II)/α-KG-dependent halogenases are capable of halogenating aliphatic C−H bonds without the need for a leaving group. This class was initially believed to be of little value for biocatalysis due their selectivity for carrier-protein tethered amino acid substrates, although the recent discovery of a α-Fe(II)/α-KG-dependent halogenase capable of catalyzing the regio- and stereoselective halogenation of freely diffusing substrates offers fresh promise for this class of halogenase.[108−111] The substrate scope of this class is yet to be explored extensively, as is their amenability to engineering for modulation of substrate scope and selectivity, but the limited number of examples thus far is promising and is likely to broaden should further enzymes of this class be discovered. The demonstration that an Fe(II)/α-KG dependent hydroxylase can be engineered to provide a halogenase biocatalyst is also promising because this could allow the engineering of other α-KG/Fe(II)-dependent hydroxylases, which accept a broad range of freely diffusing substrates, into additional biocatalysts for the halogenation of aliphatic C−H bonds.[111,113]

The flavin-dependent halogenases are the most promising biocatalysts for the halogenation of aromatic C−H bonds and as such have received much attention. The tryptophan halogenases are the most extensively studied class to date, with a number of enzymes known which can halogenate complementary positions of a range of substrates.[145,150,155,218,220,221,225] Given the number of phenolic flavin-dependent halogenases demonstrated to be involved in the biosynthesis of natural products, a great deal of which can accept freely diffusing substrates (vide infra), the application of these enzymes will likely significantly broaden the substrate scope of these biotransformations. The Fl-Hals are still of limited scalability, however, due to low productivity and stability. The efforts described above on engineering other flavin-dependent monooxygenases demonstrate the strides that can be made in process development, and in using rational protein engineering or directed evolution, toward developing flavin-dependent hydroxylases, epoxidases, and BVMOs for practical biocatalysis and illustrates the need to apply such rationale to the improvement of the less well-developed flavin-dependent halogenases. As the monooxygenase screening methods reported so far are dependent upon detection of a specific reaction product or sequence of reactions occurring, they are not appropriate for application to the flavin-dependent halogenases, although methods that could be applied to their directed evolution are slowly emerging and have been demonstrated feasible.[223,240,241] In combination with what is known about the modification of their substrate scope, regioselectivity, and stability through mutagenesis,[150,155,222−225] in addition to the discovery of naturally occurring thermostable Fl-Hals,[145] a number of reliable starting points for further engineering efforts are known. Other monooxygenases may also provide useful information about loci for targeting further engineering. Enzymes with lower rates of flavin decoupling, for example, could be used to focus engineering of Fl-Hals with improved stabilization of hydroperoxy-flavin, thereby reducing uncoupling and enzyme deactivation and potentially improving reaction efficiency and biocatalyst stability.[275] Mutation of such cofactor binding sites has been demonstrated to have a positive effect upon catalytic parameters with Fl-Hals and other enzymes.[242,384] Knowledge of the mechanisms of oxygen activation by the monooxygenases could also be used to promote formation and stabilization of this key intermediate.[385] It is possible however that a lack of productivity with the Fl-Hals could be due largely to the fact that the rate-determining step (electrophilic aromatic substitution) is very energetically demanding and therefore engineering could be focused upon stabilization of the Wheland intermediate or increasing electrophilicity of the halogenating species. The outcomes from the rational methods to improve the oxidative and thermostability of BVMO could also be useful to Fl-Hal engineering. As both classes suffer from deactivation from ROS, the mutation of active site and surface sulfur-containing residues to less oxidatively labile residues, found to improve the oxidative stability of BVMOs,[334] could develop halogenase biocatalysts more appropriate for industrial applications.

The methods used to improve monooxygenase practicability and productivity at the reaction and process engineering level may be more directly applicable to the Fl-Hals in the short term. For example, there has been little published work on the effect of oxygenation upon Fl-Hal-mediated biotransformations beyond mechanistic studies,[128,132] and therefore optimization in this regard may be a valuable avenue to peruse in their future development. It appears that the use of additional enzymes to destroy ROS is not commonplace in protocols for using Fl-Hals,[227] and therefore it may be useful to determine the extent of uncoupling in these reactions as well as its impact on enzyme productivity. Fairly straightforward and inexpensive methods of SFPR such as biphasic reactions and adsorbent resins do not seem to have been explored with the Fl-Hals either, and this may also provide some improvement of their scalability. Immobilization of the Fl-Hals does seem to have had an impact on their practicability,[150,244,267] and further work on immobilization methods may be promising, especially because heterogeneous materials may allow a flow-style biocatalytic halogenation, and hence SFPR. Such a system could allow the concentration of substrate, product, and ROS to be limited while allowing significant material to be halogenated with the same biocatalyst. Recovery of the biocatalyst in this manner may also allow more efficient halogenase-transition metal cascades. Additionally, work on adapting the Fl-Hals to accept

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

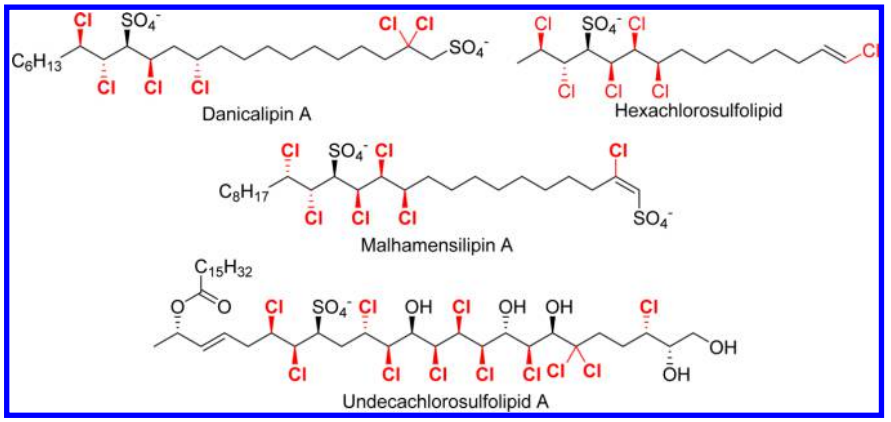


Figure 33. Chlorosulfolipid natural products identified to date identified from algae and toxic mussels.

larger, bioactive, substrates[223,225] may mean that ultimately they find application as a means of late-stage C−H activation.

Although this Review has focused largely upon halogenase enzymes and classes that are well understood and therefore closer to application in synthesis, halogenases that could be of significant utility may yet be discovered because there are halogenated natural products known with biosynthetic pathways that are yet to be fully elucidated. Study of the halogenation reactions involved in the biosynthesis of these compounds may lead to the discovery of synthetically relevant halogenation biocatalysts. Chlorosulfolipids, for example, were first isolated from algae in the 1970s,[386−388] but to date the enzymes responsible for installation of the halogen atoms have not been identified.[389] The halogenases from chlorosulfolipid biosynthesis may be of particular interest as biocatalysts because natural products from this class contain numerous chlorine atoms with defined stereochemistry,[390−393] and the halogenases involved appear to be capable of utilizing chloride and bromide interchangeably.[394] With these natural products thought to be derived from fatty acids,[395,396] the enzymes involved are likely capable of the stereo- and regioselective halogenation of aliphatic CH bonds, similar to the Fe(II)/α-KG halogenases. Additionally, the range of halogenation patterns observed in this class of natural products (Figure 33) suggests exquisite control by a set of halogenase enzymes. Identification of the enzymes involved may therefore reveal biocatalysts of great synthetic utility. Study of these halogenases has likely been hampered by the difficulties associated with mining marine genomes, although the recent identification of a Fl-Hal from the metagenome of a marine sponge adds further promise to the successful identification and characterization of enzymes from such environments.[397]

AUTHOR INFORMATION

Corresponding Author

*E-mail: jason.micklefield@manchester.ac.uk.

ORCID

Jason Micklefield: 0000-0001-8951-4873

Notes

The authors declare no competing financial interest.

Biographies

Jonathan Latham obtained his M.Chem. in Chemistry with Medicinal Chemistry from the University of Manchester in 2013, where he spent his final year project in the Micklefield laboratory studying the application of flavin-dependent halogenases for biocatalysis. After graduation, he remained at the University of Manchester to study for a Ph.D. under the tutelage of Profs. Jason Micklefield and Michael Greaney on the integration of flavin-dependent halogenases with chemo-catalytic transformations.

Eileen Brandenburger studied Chemistry at the Justus-Liebig-University in Germany where she obtained her Master's degree in 2013. In November 2013, she started her Ph.D. in Chemical Biology at the University of Manchester within the Manchester Institute of Biotechnology under the supervision of Prof. Jason Micklefield. Her research is focused on the development of novel halogenase enzymes for application in synthetic pathways.

Sarah Shepherd graduated from the University of Warwick in 2010 with a M.Chem. in Chemistry with Medicinal Chemistry. She then joined the Micklefield lab and received her Ph.D. in Chemical Biology in 2014. Her research interests include biocatalysis and biosynthesis.

Binuraj Menon has two Master's degrees in Pharmaceutical Chemistry (2003) and Biocatalysis (2006). He obtained his Ph.D. in 2009 on Molecular Enzymology with Prof. Nigel S. Scrutton at the University of Manchester. Before joining the Micklefield group in 2013, he was a postdoctoral researcher in structural biology with Dr. David Leys at the same university. Binuraj's research interests are in exploring novel proteins, and their structural, functional, and biophysical characterization for direct application in the pharmaceutical industry.

Jason Micklefield graduated from the University of Cambridge in 1993 with a Ph.D. in Chemistry, with Prof. Sir Alan R. Battersby FRS. He then moved to the University of Washington as a NATO postdoctoral fellow working with Prof. Heinz G. Floss. In 1995 he became a lecturer in organic chemistry at the University of London, Birkbeck College, before moving to the University of Manchester in 1998 where he is Professor of Chemical Biology within the School of Chemistry and the Manchester Institute of Biotechnology. His research interests include biosynthesis, biocatalysis, and riboswitches.

REFERENCES

(1) Ruiz-Castillo, P.; Buchwald, S. L. Applications of Palladium-Catalyzed C−N Cross-Coupling Reactions. *Chem. Rev.* **2016**, *116* (19), 12564−12649.
(2) Han, F.-S. Transition-Metal-Catalyzed Suzuki-Miyaura Cross-Coupling Reactions: A Remarkable Advance from Palladium to Nickel Catalysts. *Chem. Soc. Rev.* **2013**, *42* (12), 5270−5298.
(3) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Palladium-Catalyzed Cross-Coupling Reactions in Total Synthesis. *Angew. Chem., Int. Ed.* **2005**, *44* (29), 4442−4489.
(4) Hartwig, J. F. Carbon-Heteroatom Bond Formation Catalysed by Organometallic Complexes. *Nature* **2008**, *455* (7211), 314−322.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(5) Tasker, S. Z.; Standley, E. A.; Jamison, T. F. Recent Advances in Homogeneous Nickel Catalysis. *Nature* **2014**, *509* (7500), 299−309.
(6) Kuranaga, T.; Sesoko, Y.; Inoue, M. Cu-Mediated Enamide Formation in the Total Synthesis of Complex Peptide Natural Products. *Nat. Prod. Rep.* **2014**, *31* (4), 514−532.
(7) Monnier, F.; Taillefer, M. Catalytic C-C, C-N, and C-O Ullmann-Type Coupling Reactions. *Angew. Chem., Int. Ed.* **2009**, *48* (38), 6954−6971.
(8) Beletskaya, I. P.; Cheprakov, A. V. The Complementary Competitors: Palladium and Copper in C−N Cross-Coupling Reactions. *Organometallics* **2012**, *31* (22), 7753−7808.
(9) Slagt, V. F.; de Vries, A. H. M.; de Vries, J. G.; Kellogg, R. M. Practical Aspects of Carbon−Carbon Cross-Coupling Reactions Using Heteroarenes. *Org. Process Res. Dev.* **2010**, *14* (1), 30−47.
(10) Xu, Z.; Yang, Z.; Liu, Y.; Lu, Y.; Chen, K.; Zhu, W. Halogen Bond: Its Role Beyond Drug−Target Binding Affinity for Drug Discovery and Development. *J. Chem. Inf. Model.* **2014**, *54* (1), 69−78.
(11) Lu, Y.; Liu, Y.; Xu, Z.; Li, H.; Liu, H.; Zhu, W. Halogen Bonding for Rational Drug Design and New Drug Discovery. *Expert Opin. Drug Discovery* **2012**, *7* (5), 375−383.
(12) Zhou, Y.; Wang, J.; Gu, Z.; Wang, S.; Zhu, W.; Aceña, J. L.; Soloshonok, V. A.; Izawa, K.; Liu, H. Next Generation of Fluorine-Containing Pharmaceuticals, Compounds Currently in Phase Ii−Iii Clinical Trials of Major Pharmaceutical Companies: New Structural Trends and Therapeutic Areas. *Chem. Rev.* **2016**, *116* (2), 422−518.
(13) Wang, J.; Sánchez-Roselló, M.; Aceña, J. L.; del Pozo, C.; Sorochinsky, A. E.; Fustero, S.; Soloshonok, V. A.; Liu, H. Fluorine in Pharmaceutical Industry: Fluorine-Containing Drugs Introduced to the Market in the Last Decade (2001−2011). *Chem. Rev.* **2014**, *114* (4), 2432−2506.
(14) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in Medicinal Chemistry. *Chem. Soc. Rev.* **2008**, *37* (2), 320−330.
(15) Harris, C. M.; Kannan, R.; Kopecka, H.; Harris, T. M. The Role of the Chlorine Substituents in the Antibiotic Vancomycin: Preparation and Characterization of Mono- and Didechlorovancomycin. *J. Am. Chem. Soc.* **1985**, *107* (23), 6652−6658.
(16) Jeschke, P. The Unique Role of Halogen Substituents in the Design of Modern Agrochemicals. *Pest Manage. Sci.* **2010**, *66* (1), 10−27.
(17) Jeschke, P. The Unique Role of Fluorine in the Design of Active Ingredients for Modern Crop Protection. *ChemBioChem* **2004**, *5* (5), 570−589.
(18) Berger, G.; Soubhye, J.; Meyer, F. Halogen Bonding in Polymer Science: From Crystal Engineering to Functional Supramolecular Polymers and Materials. *Polym. Chem.* **2015**, *6* (19), 3559−3580.
(19) Tang, M. L.; Bao, Z. Halogenated Materials as Organic Semiconductors. *Chem. Mater.* **2011**, *23* (3), 446−455.
(20) Amanchukwu, C. V.; Harding, J. R.; Shao-Horn, Y.; Hammond, P. T. Understanding the Chemical Stability of Polymers for Lithium−Air Batteries. *Chem. Mater.* **2015**, *27* (2), 550−561.
(21) Jiang, S.; Zhang, L.; Cui, D.; Yao, Z.; Gao, B.; Lin, J.; Wei, D. The Important Role of Halogen Bond in Substrate Selectivity of Enzymatic Catalysis. *Sci. Rep.* **2016**, *6*, 34750.
(22) Sirimulla, S.; Bailey, J. B.; Vegesna, R.; Narayan, M. Halogen Interactions in Protein−Ligand Complexes: Implications of Halogen Bonding for Rational Drug Design. *J. Chem. Inf. Model.* **2013**, *53* (11), 2781−2791.
(23) Cavallo, G.; Metrangolo, P.; Milani, R.; Pilati, T.; Priimagi, A.; Resnati, G.; Terraneo, G. The Halogen Bond. *Chem. Rev.* **2016**, *116* (4), 2478−2601.
(24) Lu, Y.; Wang, Y.; Zhu, W. Nonbonding Interactions of Organic Halogens in Biological Systems: Implications for Drug Discovery and Biomolecular Design. *Phys. Chem. Chem. Phys.* **2010**, *12* (18), 4543−4551.
(25) Parisini, E.; Metrangolo, P.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen Bonding in Halocarbon-Protein Complexes: A Structural Survey. *Chem. Soc. Rev.* **2011**, *40* (5), 2267−2278.
(26) Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. Halogen Bonding: The σ-Hole. *J. Mol. Model.* **2007**, *13* (2), 291−296.
(27) Wolters, L. P.; Schyman, P.; Pavan, M. J.; Jorgensen, W. L.; Bickelhaupt, F. M.; Kozuch, S. The Many Faces of Halogen Bonding: A Review of Theoretical Models and Methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (6), 523−540.
(28) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2013**, *56* (4), 1363−1388.
(29) Gribble, G. W. *Naturally Occurring Organohalogen Compounds - a Comprehensive Update*; Springer Vienna: Vienna, 2010.
(30) Gribble, G. W. Natural Organohalogens: A New Frontier for Medicinal Agents? *J. Chem. Educ.* **2004**, *81* (10), 1441.
(31) Gribble, G. W. *Modern Alkaloids*; Wiley-VCH: Weinheim, 2008.
(32) Weichold, V.; Milbredt, D.; van Pée, K.-H. Specific Enzymatic Halogenation—from the Discovery of Halogenated Enzymes to Their Applications in Vitro and in Vivo. *Angew. Chem., Int. Ed.* **2016**, *55* (22), 6374−6389.
(33) Brown, S.; O'Connor, S. E. Halogenase Engineering for the Generation of New Natural Product Analogues. *ChemBioChem* **2015**, *16* (15), 2129−2135.
(34) Smith, D. R. M.; Grüschow, S.; Goss, R. J. M. Scope and Potential of Halogenases in Biosynthetic Applications. *Curr. Opin. Chem. Biol.* **2013**, *17* (2), 276−283.
(35) Prakash, G. K. S.; Mathew, T.; Hoole, D.; Esteves, P. M.; Wang, Q.; Rasul, G.; Olah, G. A. N-Halosuccinimide/BF_3−H_2O, Efficient Electrophilic Halogenating Systems for Aromatics. *J. Am. Chem. Soc.* **2004**, *126* (48), 15770−15776.
(36) Alonso, F.; Beletskaya, I. P.; Yus, M. Metal-Mediated Reductive Hydrodehalogenation of Organic Halides. *Chem. Rev.* **2002**, *102* (11), 4009−4092.
(37) Petrone, D. A.; Ye, J.; Lautens, M. Modern Transition-Metal-Catalyzed Carbon−Halogen Bond Formation. *Chem. Rev.* **2016**, *116* (14), 8003−8104.
(38) Muralirajan, K.; Haridharan, R.; Prakash, S.; Cheng, C.-H. Rhodium(III)-Catalyzed *in situ* Oxidizing Directing Group-Assisted C-H Bond Activation and Olefination: A Route to 2-Vinylanilines. *Adv. Synth. Catal.* **2015**, *357* (4), 761−766.
(39) Teskey, C. J.; Lui, A. Y. W.; Greaney, M. F. Ruthenium-Catalyzed Meta-Selective C-H Bromination. *Angew. Chem., Int. Ed.* **2015**, *54* (40), 11677−11680.
(40) Sandtorv, A. H. Transition Metal-Catalyzed C-H Activation of Indoles. *Adv. Synth. Catal.* **2015**, *357* (11), 2403−2435.
(41) Xu, H.; Shang, M.; Dai, H.-X.; Yu, J.-Q. Ligand-Controlled Para-Selective C−H Arylation of Monosubstituted Arenes. *Org. Lett.* **2015**, *17* (15), 3830−3833.
(42) Yu, J.-Q.; Shi, Z. *C-H Activation*; Springer: Berlin, Heidelberg, 2010.
(43) Leblanc, C.; Vilter, H.; Fournier, J. B.; Delage, L.; Potin, P.; Rebuffet, E.; Michel, G.; Solari, P. L.; Feiters, M. C.; Czjzek, M. Vanadium Haloperoxidases: From the Discovery 30 Years Ago to X-Ray Crystallographic and V K-Edge Absorption Spectroscopic Studies. *Coord. Chem. Rev.* **2015**, *301−302*, 134−146.
(44) Deboer, E.; Vankooyk, Y.; Tromp, M. G. M.; Plat, H.; Wever, R. Bromoperoxidase from Ascophyllum-Nodosum - a Novel Class of Enzymes Containing Vanadium as a Prosthetic Group. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1986**, *869* (1), 48−53.
(45) Soedjak, H. S.; Butler, A. Characterization of Vanadium Bromoperoxidase from Macrocystis and Fucus - Reactivity of Vanadium Bromoperoxidase toward Acyl and Alkyl Peroxides and Bromination of Amines. *Biochemistry* **1990**, *29* (34), 7974−7981.
(46) Morris, D. R.; Hager, L. P. Chloroperoxidase I: Isolation and Properties of Crystalline Glycoprotein. *J. Biol. Chem.* **1966**, *241*, 1763−1768.
(47) Wagner, C.; Molitor, I. M.; Konig, G. M. Critical View on the Monochlorodimedone Assay Utilized to Detect Haloperoxidase Activity. *Phytochemistry* **2008**, *69* (2), 323−332.
(48) Bernhardt, P.; Okino, T.; Winter, J. M.; Miyanaga, A.; Moore, B. S. A Stereoselective Vanadium-Dependent Chloroperoxidase in

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

Bacterial Antibiotic Biosynthesis. *J. Am. Chem. Soc.* **2011**, *133* (12), 4268−4270.
(49) Diethelm, S.; Teufel, R.; Kaysser, L.; Moore, B. S. A Multitasking Vanadium-Dependent Chloroperoxidase as an Inspiration for the Chemical Synthesis of the Merochlorins. *Angew. Chem., Int. Ed.* **2014**, *53* (41), 11023−11026.
(50) Hasan, Z.; Renirie, R.; Kerkman, R.; Ruijssenaars, H. J.; Hartog, A. F.; Wever, R. Laboratory-Evolved Vanadium Chloroperoxidase Exhibits 100-Fold Higher Halogenating Activity at Alkaline pH: Catalytic Effects from First and Second Coordination Sphere Mutations. *J. Biol. Chem.* **2006**, *281* (14), 9738−9744.
(51) Gwon, H. J.; Teruhiko, I.; Shigeaki, H.; Baik, S. H. Identification of Novel Non-Metal Haloperoxidases from the Marine Metagenome. *J. Microbiol. Biotechnol.* **2014**, *24* (6), 835−842.
(52) de Boer, E.; Plat, H.; Tromp, M. G. M.; Wever, R.; Franssen, M. C. R.; van der Plas, H. C.; Meijer, E. M.; Schoemaker, H. E. Vanadium Containing Bromoperoxidase: An Example of an Oxidoreductase with High Operational Stability in Aqueous and Organic Media. *Biotechnol. Bioeng.* **1987**, *30* (5), 607−610.
(53) Hunter-Cevera, J. C.; Sotos, L. Screening for a "New" Enzyme in Nature: Haloperoxidase Production by Death Valley Dematiaceous Hyphomycetes. *Microb. Ecol.* **1986**, *12* (1), 121−127.
(54) Hartung, J.; Brücher, O.; Hach, D.; Schulz, H.; Vilter, H.; Ruick, G. Bromoperoxidase Activity and Vanadium Level of the Brown Alga Ascophyllum Nodosum. *Phytochemistry* **2008**, *69* (16), 2826−2830.
(55) Gao, Q.; Jiang, Y.; Gao, X.; Hu, M.; Li, S.; Zhai, Q. Activation Function of Chloroperoxidase in the Presence of Metal Ions at Elevated Temperature from 25 to 55°C. *Chin. J. Chem.* **2009**, *27* (7), 1291−1294.
(56) Hager, L. P.; Morris, D. R.; Brown, F. S.; Eberwein, H. Chloroperoxidase Ii: Utilisation of Halogen Anions. *J. Biol. Chem.* **1966**, *241*, 1769−1777.
(57) Taurog, A.; Howells, E. M. Enzymatic Iodination of Tyrosine and Thyroglobulin with Chloroperoxidase. *J. Biol. Chem.* **1966**, *241*, 1329−1339.
(58) Fujimori, D. G.; Walsh, C. T. What's New in Enzymatic Halogenations. *Curr. Opin. Chem. Biol.* **2007**, *11* (5), 553−560.
(59) Sundaramoorthy, M.; Terner, J.; Poulos, T. L. The Crystal Structure of Chloroperoxidase: A Heme Peroxidase−Cytochrome P450 Functional Hybrid. *Structure* **1995**, *3* (12), 1367−1378.
(60) Sundaramoorthy, M.; Terner, J.; Poulos, T. L. Stereochemistry of the Chloroperoxidase Active Site: Crystallographic and Molecular-Modeling Studies. *Chem. Biol.* **1998**, *5* (9), 461−473.
(61) Kuhnel, K.; Blankenfeldt, W.; Terner, J.; Schlichting, I. Crystal Structures of Chloroperoxidase with Its Bound Substrates and Complexed with Formate, Acetate, and Nitrate. *J. Biol. Chem.* **2006**, *281* (33), 23990−23998.
(62) Littlechild, J. Haloperoxidases and Their Role in Biotransformation Reactions. *Curr. Opin. Chem. Biol.* **1999**, *3* (1), 28−34.
(63) Yi, X.; Conesa, A.; Punt, P. J.; Hager, L. P. Examining the Role of Glutamic Acid 183 in Chloroperoxidase Catalysis. *J. Biol. Chem.* **2003**, *278* (16), 13855−13859.
(64) Poulos, T. L. Heme Enzyme Structure and Function. *Chem. Rev.* **2014**, *114* (7), 3919−3962.
(65) Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. Halogenation Strategies in Natural Product Biosynthesis. *Chem. Biol.* **2008**, *15* (2), 99−109.
(66) Vazquez-Duhalt, R.; Ayala, M.; Marquez-Rocha, F. J. Biocatalytic Chlorination of Aromatic Hydrocarbons by Chloroperoxidase of Caldariomyces Fumago. *Phytochemistry* **2001**, *58* (6), 929−933.
(67) Franssen, M. C. R.; Vanderplas, H. C. The Chlorination of Barbituric-Acid and Some of Its Derivatives by Chloroperoxidase. *Bioorg. Chem.* **1987**, *15* (1), 59−70.
(68) Geigert, J.; Neidleman, S. L.; Dalietos, D. J. Novel Haloperoxidase Substrates - Alkynes and Cyclopropanes. *J. Biol. Chem.* **1983**, *258*, 2273−2277.
(69) Geigert, J.; Neidleman, S. L.; Dalietos, D. J.; Dewitt, S. K. Haloperoxidases - Enzymatic-Synthesis of Alpha,Beta-Halohydrins from Gaseous Alkenes. *Appl. Environ. Microbiol.* **1983**, *45*, 366−374.
(70) Geigert, J.; Neidleman, S. L.; Dalietos, D. J.; Dewitt, S. K. Novel Haloperoxidase Reaction - Synthesis of Dihalogenated Products. *Appl. Environ. Microbiol.* **1983**, *45*, 1575−1581.
(71) Wannstedt, C.; Rotella, D.; Siuda, J. F. Chloroperoxidase Mediated Halogenation of Phenols. *Bull. Environ. Contam. Toxicol.* **1990**, *44* (2), 282−287.
(72) Yaipakdee, P.; Robertson, L. W. Enzymatic Halogenation of Flavanones and Flavones. *Phytochemistry* **2001**, *57* (3), 341−347.
(73) Vilter, H. Peroxidases from Phaeophyceae - a Vanadium(V)-Dependent Peroxidase from Ascophyllum-Nodosum 0.5. *Phytochemistry* **1984**, *23* (7), 1387−1390.
(74) Wever, R.; de Boer, E.; Plat, H.; E. Krenn, B. Vanadium — an Element Involved in the Biosynthesis of Halogenated Compounds and Nitrogen Fixation. *FEBS Lett.* **1987**, *216* (1), 1−3.
(75) Butler, A.; Walker, J. V. Marine Haloperoxidases. *Chem. Rev.* **1993**, *93* (5), 1937−1944.
(76) Carter-Franklin, J. N.; Butler, A. Vanadium Bromoperoxidase-Catalyzed Biosynthesis of Halogenated Marine Natural Products. *J. Am. Chem. Soc.* **2004**, *126* (46), 15060−15066.
(77) Coupe, E. E.; Smyth, M. G.; Fosberry, A. P.; Hall, R. M.; Littlechild, J. A. The Dodecameric Vanadium-Dependent Haloperoxidase from the Marine Algae Corallina Officinalis: Cloning, Expression, and Refolding of the Recombinant Enzyme. *Protein Expression Purif.* **2007**, *52* (2), 265−272.
(78) Kamenarska, Z.; Taniguchi, T.; Ohsawa, N.; Hiraoka, M.; Itoh, N. A Vanadium-Dependent Bromoperoxidase in the Marine Red Alga Kappaphycus Alvarezii (Doty) Doty Displays Clear Substrate Specificity. *Phytochemistry* **2007**, *68* (10), 1358−1366.
(79) Kaneko, K.; Washio, K.; Umezawa, T.; Matsuda, F.; Morikawa, M.; Okino, T. Cdna Cloning and Characterization of Vanadium-Dependent Bromoperoxidases from the Red Alga Laurencia Nipponica. *Biosci., Biotechnol., Biochem.* **2014**, *78* (8), 1310−1319.
(80) Itoh, N.; Izumi, Y.; Yamada, H. Characterization of Nonheme Type Bromoperoxidase in Corallina Pilulifera. *J. Biol. Chem.* **1986**, *261*, 5194−5200.
(81) Plat, H.; Krenn, B. E.; Wever, R. The Bromoperoxidase from the Lichen *Xanthoria Parietina* Is a Novel Vanadium Enzyme. *Biochem. J.* **1987**, *248* (1), 277−279.
(82) van Schijndel, J. W. P. M.; Simons, L. H.; Vollenbroek, E. G. M.; Wever, R. The Vanadium Chloroperoxidase from the Fungus, Curvularia Inaequalis. *FEBS Lett.* **1993**, *336* (2), 239−242.
(83) Kaysser, L.; Bernhardt, P.; Nam, S. J.; Loesgen, S.; Ruby, J. G.; Skewes-Cox, P.; Jensen, P. R.; Fenical, W.; Moore, B. S. Merochlorins a-D, Cyclic Meroterpenoid Antibiotics Biosynthesized in Divergent Pathways with Vanadium-Dependent Chloroperoxidases. *J. Am. Chem. Soc.* **2012**, *134* (29), 11988−11991.
(84) Carter, J. N.; Beatty, K. E.; Simpson, M. T.; Butler, A. Reactivity of Recombinant and Mutant Vanadium Bromoperoxidase from the Red Alga Corallina Officinalis. *J. Inorg. Biochem.* **2002**, *91* (1), 59−69.
(85) Hemrika, W.; Renirie, R.; Macedo-Ribeiro, S.; Messerschmidt, A.; Wever, R. Heterologous Expression of the Vanadium-Containing Chloroperoxidase from Curvularia Inaequalis Insaccharomyces Cerevisiae and Site-Directed Mutagenesis of the Active Site Residues His496, Lys353, Arg360, and Arg490. *J. Biol. Chem.* **1999**, *274* (34), 23820−23827.
(86) Ohshiro, T.; Hemrika, W.; Aibara, T.; Wever, R.; Izumi, Y. Expression of the Vanadium-Dependent Bromoperoxidase Gene from a Marine Macro-Alga Corallina Pilulifera in Saccharomyces Cerevisiae and Characterization of the Recombinant Enzyme. *Phytochemistry* **2002**, *60* (6), 595−601.
(87) Yamada, R.; Higo, T.; Yoshikawa, C.; China, H.; Ogino, H. Improvement of the Stability and Activity of the Bpo-A1 Haloperoxidase from Streptomyces Aureofaciens by Directed Evolution. *J. Biotechnol.* **2014**, *192*, 248−254.
(88) Yamada, R.; Higo, T.; Yoshikawa, C.; China, H.; Yasuda, M.; Ogino, H. Random Mutagenesis and Selection of Organic Solvent-Stable Haloperoxidase from Streptomyces Aureofaciens. *Biotechnol. Prog.* **2015**, *31* (4), 917−924.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(89) Littlechild, J.; Garcia-Rodriguez, E.; Dalby, A.; Isupov, M. Structural and Functional Comparisons between Vanadium Haloperoxidase and Acid Phosphatase Enzymes. *J. Mol. Recognit.* **2002**, *15* (5), 291−296.
(90) Soedjak, H. S.; Everett, R. R.; Butler, A. The Novel Nonheme Vanadium Bromoperoxidase from Marine-Algae - Phosphate Inactivation. *J. Ind. Microbiol.* **1991**, *8* (1), 37−44.
(91) Tanaka, N.; Wever, R. Inhibition of Vanadium Chloroperoxidase from the Fungus Curvularia Inaequalis by Hydroxylamine, Hydrazine and Azide and Inactivation by Phosphate. *J. Inorg. Biochem.* **2004**, *98* (4), 625−631.
(92) Messerschmidt, A.; Prade, L.; Wever, R. Implications for the Catalytic Mechanism of the Vanadium-Containing Enzyme Chloroperoxidase from the Fungus Curvularia Inaequalis by X-Ray Structures of the Native and Peroxide Form. *Biol. Chem.* **1997**, *378* (3−4), 309−315.
(93) Messerschmidt, A.; Wever, R. X-Ray Structure of a Vanadium-Containing Enzyme: Chloroperoxidase from the Fungus Curvularia Inaequalis. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (1), 392−396.
(94) Weyand, M.; Hecht, H. J.; Kiess, M.; Liaud, M. F.; Vilter, H.; Schomburg, D. X-Ray Structure Determination of a Vanadium-Dependent Haloperoxidase from Ascophyllum Nodosum at 2.0 Angstrom Resolution. *J. Mol. Biol.* **1999**, *293* (3), 595−611.
(95) Littlechild, J. Structural Studies on the Dodecameric Vanadium Bromoperoxidase from Corallina Species. *Coord. Chem. Rev.* **2003**, *237* (1−2), 65−76.
(96) Itoh, N.; Hasan, A. K. M. Q.; Izumi, Y.; Yamada, H. Substrate Specificity, Regiospecificity and Stereospecificity of Halogenation Reactions Catalyzed by Non-Heme-Type Bromoperoxidase of Corallina Pilulifera. *Eur. J. Biochem.* **1988**, *172* (2), 477−484.
(97) Wischang, D.; Brücher, O.; Hartung, J. Bromoperoxidases and Functional Enzyme Mimics as Catalysts for Oxidative Bromination—a Sustainable Synthetic Approach. *Coord. Chem. Rev.* **2011**, *255* (19−20), 2204−2217.
(98) Wischang, D.; Hartung, J.; Hahn, T.; Ulber, R.; Stumpf, T.; Fecher-Trost, C. Vanadate(V)-Dependent Bromoperoxidase Immobilized on Magnetic Beads as Reusable Catalyst for Oxidative Bromination. *Green Chem.* **2011**, *13* (1), 102−108.
(99) Macedo-Ribeiro, S.; Hemrika, W.; Renirie, R.; Wever, R.; Messerschmidt, A. X-Ray Crystal Structures of Active Site Mutants of the Vanadium-Containing Chloroperoxidase from the Fungus Curvularia Inaequalis. *JBIC, J. Biol. Inorg. Chem.* **1999**, *4* (2), 209−219.
(100) Winter, J. M.; Moore, B. S. Exploring the Chemistry and Biology of Vanadium-Dependent Haloperoxidases. *J. Biol. Chem.* **2009**, *284* (28), 18577−18581.
(101) Teufel, R.; Kaysser, L.; Villaume, M. T.; Diethelm, S.; Carbullido, M. K.; Baran, P. S.; Moore, B. S. One-Pot Enzymatic Synthesis of Merochlorin a and B. *Angew. Chem., Int. Ed.* **2014**, *53* (41), 11019−11022.
(102) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. Syrb2 in Syringomycin E Biosynthesis Is a Nonheme Feii A-Ketoglutarate- and O2-Dependent Halogenase. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (29), 10111−10116.
(103) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. Cryptic Chlorination by a Non-Haem Iron Enzyme During Cyclopropyl Amino Acid Biosynthesis. *Nature* **2005**, *436* (7054), 1191−1194.
(104) Sitachitta, N.; Márquez, B. L.; Thomas Williamson, R.; Rossi, J.; Ann Roberts, M.; Gerwick, W. H.; Nguyen, V.-A.; Willis, C. L. Biosynthetic Pathway and Origin of the Chlorinated Methyl Group in Barbamide and Dechlorobarbamide, Metabolites from the Marine Cyanobacterium Lyngbya Majuscula. *Tetrahedron* **2000**, *56* (46), 9103−9113.
(105) Sitachitta, N.; Rossi, J.; Roberts, M. A.; Gerwick, W. H.; Fletcher, M. D.; Willis, C. L. Biosynthesis of the Marine Cyanobacterial Metabolite Barbamide. 1. Origin of the Trichloromethyl Group. *J. Am. Chem. Soc.* **1998**, *120* (28), 7131−7132.
(106) Galonić, D. P.; Vaillancourt, F. H.; Walsh, C. T. Halogenation of Unactivated Carbon Centers in Natural Product Biosynthesis: Trichlorination of Leucine During Barbamide Biosynthesis. *J. Am. Chem. Soc.* **2006**, *128* (12), 3900−3901.
(107) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. Crystal Structure of the Non-Haem Iron Halogenase SyrB2 in Syringomycin Biosynthesis. *Nature* **2006**, *440* (7082), 368−371.
(108) Hillwig, M. L.; Liu, X. A New Family of Iron-Dependent Halogenases Acts on Freestanding Substrates. *Nat. Chem. Biol.* **2014**, *10* (11), 921−923.
(109) Hillwig, M. L.; Zhu, Q.; Liu, X. Biosynthesis of Ambiguine Indole Alkaloids in Cyanobacterium Fischerella Ambigua. *ACS Chem. Biol.* **2014**, *9* (2), 372−377.
(110) Hillwig, M. L.; Fuhrman, H. A.; Ittiamornkul, K.; Sevco, T. J.; Kwak, D. H.; Liu, X. Identification and Characterization of a Welwitindolinone Alkaloid Biosynthetic Gene Cluster in the Stigonematalean Cyanobacterium Hapalosiphon Welwitschii. *ChemBioChem* **2014**, *15* (5), 665−669.
(111) Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron- and 2-Oxo-Glutarate-Dependent Enzyme Welo5. *Nat. Chem. Biol.* **2016**, *12* (8), 636−640.
(112) Matthews, M. L.; Neumann, C. S.; Miles, L. A.; Grove, T. L.; Booker, S. J.; Krebs, C.; Walsh, C. T.; Bollinger, J. M. Substrate Positioning Controls the Partition between Halogenation and Hydroxylation in the Aliphatic Halogenase, SyrB2. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (42), 17723−17728.
(113) Mitchell, A. J.; Dunham, N. P.; Bergman, J. A.; Wang, B.; Zhu, Q.; Chang, W.-C.; Liu, X.; Boal, A. K. Structure-Guided Reprogramming of a Hydroxylase to Halogenate Its Small Molecule Substrate. *Biochemistry* **2017**, *56* (3), 441−444.
(114) Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K.-H.; Naismith, J. H. Tryptophan 7-Halogenase (PrnA) Structure Suggests a Mechanism for Regioselective Chlorination. *Science* **2005**, *309* (5744), 2216−2219.
(115) Mascotti, M. L.; Juri Ayub, M.; Furnham, N.; Thornton, J. M.; Laskowski, R. A. Chopping and Changing: The Evolution of the Flavin-Dependent Monooxygenases. *J. Mol. Biol.* **2016**, *428* (15), 3131−3146.
(116) Huijbers, M. M.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J. Flavin Dependent Monooxygenases. *Arch. Biochem. Biophys.* **2014**, *544*, 2−17.
(117) van Berkel, W. J. H.; Kamerbeek, N. M.; Fraaije, M. W. Flavoprotein Monooxygenases, a Diverse Class of Oxidative Biocatalysts. *J. Biotechnol.* **2006**, *124* (4), 670−689.
(118) Visitsatthawong, S.; Chenprakhon, P.; Chaiyen, P.; Surawatanawong, P. Mechanism of Oxygen Activation in a Flavin-Dependent Monooxygenase: A Nearly Barrierless Formation of C4a-Hydroperoxyflavin Via Proton-Coupled Electron Transfer. *J. Am. Chem. Soc.* **2015**, *137* (29), 9363−9374.
(119) Badieyan, S.; Bach, R. D.; Sobrado, P. Mechanism of N-Hydroxylation Catalyzed by Flavin-Dependent Monooxygenases. *J. Org. Chem.* **2015**, *80* (4), 2139−2147.
(120) Morrison, E.; Kantz, A.; Gassner, G. T.; Sazinsky, M. H. Structure and Mechanism of Styrene Monooxygenase Reductase: New Insight into the Fad-Transfer Reaction. *Biochemistry* **2013**, *52* (35), 6063−6075.
(121) Kamerbeek, N. M.; Janssen, D. B.; van Berkel, W. J. H.; Fraaije, M. W. Baeyer−Villiger Monooxygenases, an Emerging Family of Flavin-Dependent Biocatalysts. *Adv. Synth. Catal.* **2003**, *345* (6−7), 667−678.
(122) Pang, A. H.; Garneau-Tsodikova, S.; Tsodikov, O. V. Crystal Structure of Halogenase Plta from the Pyoluteorin Biosynthetic Pathway. *J. Struct. Biol.* **2015**, *192* (3), 349−357.
(123) Yeh, E.; Garneau, S.; Walsh, C. T. Robust in Vitro Activity of Rebf and Rebh, a Two-Component Reductase/Halogenase, Generating 7-Chlorotryptophan During Rebeccamycin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (11), 3960−3965.
(124) Onaka, H.; Taniguchi, S.; Igarashi, Y.; Furumai, T. Characterization of the Biosynthetic Gene Cluster of Rebeccamycin from

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

Lechevalieria Aerocolonigenes ATCC 39243. *Biosci., Biotechnol., Biochem.* **2003**, *67* (1), 127−138.
(125) Zeng, J.; Zhan, J. A Novel Fungal Flavin-Dependent Halogenase for Natural Product Biosynthesis. *ChemBioChem* **2010**, *11* (15), 2119−2123.
(126) Ferrara, M.; Perrone, G.; Gambacorta, L.; Epifani, F.; Solfrizzo, M.; Gallo, A. Identification of a Halogenase Involved in the Biosynthesis of Ochratoxin a in Aspergillus Carbonarius. *Appl. Environ. Microbiol.* **2016**, *82* (18), 5631−5641.
(127) Keller, S.; Wage, T.; Hohaus, K.; Hölzer, M.; Eichhorn, E.; van Pée, K.-H. Purification and Partial Characterization of Tryptophan 7-Halogenase (PrnA) from Pseudomonas Fluorescens. *Angew. Chem., Int. Ed.* **2000**, *39* (13), 2300−2302.
(128) Unversucht, S.; Hollmann, F.; Schmid, A.; van Pée, K.-H. $FADH_2$-Dependence of Tryptophan 7-Halogenase. *Adv. Synth. Catal.* **2005**, *347* (7−8), 1163−1167.
(129) van Pee, K. H.; Patallo, E. P. Flavin-Dependent Halogenases Involved in Secondary Metabolism in Bacteria. *Appl. Microbiol. Biotechnol.* **2006**, *70* (6), 631−641.
(130) Zhu, X.; De Laurentis, W.; Leang, K.; Herrmann, J.; Ihlefeld, K.; van Pée, K.-H.; Naismith, J. H. Structural Insights into Regioselectivity in the Enzymatic Chlorination of Tryptophan. *J. Mol. Biol.* **2009**, *391* (1), 74−85.
(131) Dong, C.; Kotzsch, A.; Dorward, M.; van Pee, K.-H.; Naismith, J. H. Crystallization and X-Ray Diffraction of a Halogenating Enzyme, Tryptophan 7-Halogenase, from Pseudomonas Fluorescens. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (8), 1438−1440.
(132) Yeh, E.; Cole, L. J.; Barr, E. W.; Bollinger, J. M.; Ballou, D. P.; Walsh, C. T. Flavin Redox Chemistry Precedes Substrate Chlorination During the Reaction of the Flavin-Dependent Halogenase Rebh. *Biochemistry* **2006**, *45* (25), 7904−7912.
(133) Yeh, E.; Blasiak, L. C.; Koglin, A.; Drennan, C. L.; Walsh, C. T. Chlorination by a Long-Lived Intermediate in the Mechanism of Flavin-Dependent Halogenases. *Biochemistry* **2007**, *46* (5), 1284−1292.
(134) Bitto, E.; Huang, Y.; Bingman, C. A.; Singh, S.; Thorson, J. S.; Phillips, G. N. The Structure of Flavin-Dependent Tryptophan 7-Halogenase Rebh. *Proteins: Struct., Funct., Genet.* **2008**, *70* (1), 289−293.
(135) Flecks, S.; Patallo, E. P.; Zhu, X.; Ernyei, A. J.; Seifert, G.; Schneider, A.; Dong, C.; Naismith, J. H.; van Pée, K.-H. New Insights into the Mechanism of Enzymatic Chlorination of Tryptophan. *Angew. Chem., Int. Ed.* **2008**, *47* (49), 9533−9536.
(136) Podzelinska, K.; Latimer, R.; Bhattacharya, A.; Vining, L. C.; Zechel, D. L.; Jia, Z. Chloramphenicol Biosynthesis: The Structure of CmlS, a Flavin-Dependent Halogenase Showing a Covalent Flavin−Aspartate Bond. *J. Mol. Biol.* **2010**, *397* (1), 316−331.
(137) Kirner, S.; Hammer, P. E.; Hill, D. S.; Altmann, A.; Fischer, I.; Weislo, L. J.; Lanahan, M.; van Pee, K. H.; Ligon, J. M. Functions Encoded by Pyrrolnitrin Biosynthetic Genes from Pseudomonas Fluorescens. *J. Bacteriol.* **1998**, *180*, 1939−1943.
(138) Kirner, S.; Krauss, S.; Sury, G.; Lam, S. T.; Ligon, J. M.; vanPee, K. H. The Non-Haem Chloroperoxidase from Pseudomonas Fluorescens and Its Relationship to Pyrrolnitrin Biosynthesis. *Microbiology* **1996**, *142*, 2129−2135.
(139) Sanchez, C.; Butovich, I. A.; Brana, A. F.; Rohr, J.; Mendez, C.; Salas, J. A. The Biosynthetic Gene Cluster for the Antitumor Rebeccamycin: Characterization and Generation of Indolocarbazole Derivatives. *Chem. Biol.* **2002**, *9* (4), 519−531.
(140) Jiang, W.; Heemstra, J. R., Jr.; Forseth, R. R.; Neumann, C. S.; Manaviazar, S.; Schroeder, F. C.; Hale, K. J.; Walsh, C. T. Biosynthetic Chlorination of the Piperazate Residue in Kutzneride Biosynthesis by Kthp. *Biochemistry* **2011**, *50* (27), 6063−6072.
(141) Heemstra, J. R.; Walsh, C. T. Tandem Action of the O2- and Fadh2-Dependent Halogenases KtzQ and KtzR Produce 6,7-Dichlorotryptophan for Kutzneride Assembly. *J. Am. Chem. Soc.* **2008**, *130* (43), 14024−14025.
(142) Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A.; Walsh, C. T. Cloning and Characterization of the Biosynthetic Gene Cluster for Kutznerides. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (42), 16498−16503.
(143) Zeng, J.; Zhan, J. Characterization of a Tryptophan 6-Halogenase from Streptomyces Toxytricini. *Biotechnol. Lett.* **2011**, *33* (8), 1607−1613.
(144) Milbredt, D.; Patallo, E. P.; van Pée, K.-H. A Tryptophan 6-Halogenase and an Amidotransferase Are Involved in Thienodolin Biosynthesis. *ChemBioChem* **2014**, *15* (7), 1011−1020.
(145) Menon, B. R. K.; Latham, J.; Dunstan, M. S.; Brandenburger, E.; Klemstein, U.; Leys, D.; Karthikeyan, C.; Greaney, M. F.; Shepherd, S. A.; Micklefield, J. Structure and Biocatalytic Scope of Thermophilic Flavin-Dependent Halogenase and Flavin Reductase Enzymes. *Org. Biomol. Chem.* **2016**, *14* (39), 9354−9361.
(146) Wang, Y.; Wang, J.; Yu, S.; Wang, F.; Ma, H.; Yue, C.; Liu, M.; Deng, Z.; Huang, Y.; Qu, X. Identifying the Minimal Enzymes for Unusual Carbon-Sulfur Bond Formation in Thienodolin Biosynthesis. *ChemBioChem* **2016**, *17* (9), 799−803.
(147) Xu, L.; Han, T.; Ge, M.; Zhu, L.; Qian, X. Discovery of the New Plant Growth-Regulating Compound LyxlF2 Based on Manipulating the Halogenase in Amycolatopsis Orientalis. *Curr. Microbiol.* **2016**, *73* (3), 335−340.
(148) Wang, Y.; Wang, J.; Yu, S.; Wang, F.; Ma, H.; Yue, C.; Liu, M.; Deng, Z.; Huang, Y.; Qu, X. Corrigendum: Identifying the Minimal Enzymes for Unusual Carbon-Sulfur Bond Formation in Thienodolin Biosynthesis. *ChemBioChem* **2016**, *17* (9), 876.
(149) Milbredt, D.; Patallo, E. P.; van Pee, K. H. Characterization of the Aminotransferase Thdn from Thienodolin Biosynthesis in Streptomyces Albogriseolus. *ChemBioChem* **2016**, *17* (19), 1859−1864.
(150) Shepherd, S. A.; Menon, B. R. K.; Fisk, H.; Struck, A.-W.; Levy, C.; Leys, D.; Micklefield, J. A Structure-Guided Switch in the Regioselectivity of a Tryptophan Halogenase. *ChemBioChem* **2016**, *17* (9), 821−824.
(151) Zehner, S.; Kotzsch, A.; Bister, B.; Süssmuth, R. D.; Méndez, C.; Salas, J. A.; van Pée, K.-H. A Regioselective Tryptophan 5-Halogenase Is Involved in Pyrroindomycin Biosynthesis in Streptomyces Rugosporus LI-42D005. *Chem. Biol.* **2005**, *12* (4), 445−452.
(152) Du, Y. L.; Ryan, K. S. Expansion of Bisindole Biosynthetic Pathways by Combinatorial Construction. *ACS Synth. Biol.* **2015**, *4* (6), 682−688.
(153) Ryan, K. S. Biosynthetic Gene Cluster for the Cladoniamides, Bis-Indoles with a Rearranged Scaffold. *PLoS One* **2011**, *6* (8), e23694.
(154) Ortega, M. A.; Cogan, D. P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G. N.; Maffioli, S.; Donadio, S.; Sosio, M.; Escano, J.; et al. Two Flavoenzymes Catalyze the Post-Translational Generation of 5-Chlorotryptophan and 2-Aminovinyl-Cysteine During NAI-107 Biosynthesis. *ACS Chem. Biol.* **2017**, *12* (2), 548−557.
(155) Shepherd, S. A.; Karthikeyan, C.; Latham, J.; Struck, A.-W.; Thompson, M. L.; Menon, B. R. K.; Styles, M. Q.; Levy, C.; Leys, D.; Micklefield, J. Extending the Biocatalytic Scope of Regiocomplementary Flavin-Dependent Halogenase Enzymes. *Chem. Sci.* **2015**, *6* (6), 3454−3460.
(156) El Gamal, A.; Agarwal, V.; Diethelm, S.; Rahman, I.; Schorn, M. A.; Sneed, J. M.; Louie, G. V.; Whalen, K. E.; Mincer, T. J.; Noel, J. P.; et al. Biosynthesis of Coral Settlement Cue Tetrabromopyrrole in Marine Bacteria by a Uniquely Adapted Brominase−Thioesterase Enzyme Pair. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (14), 3797−3802.
(157) Hammer, P. E.; Hill, D. S.; Lam, S. T.; VanPee, K. H.; Ligon, J. M. Four Genes from Pseudomonas Fluorescens That Encode the Biosynthesis of Pyrrolnitrin. *Appl. Environ. Microb.* **1997**, *63*, 2147−2154.
(158) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. Dichlorination of a Pyrrolyl-S-Carrier Protein by $FADH_2$-Dependent Halogenase Plta During Pyoluteorin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (39), 13843−13848.
(159) Wynands, I.; van Pee, K. H. A Novel Halogenase Gene from the Pentachloropseudilin Producer Actinoplanes Sp ATCC 33002 and

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

Detection of in Vitro Halogenase Activity. *FEMS Microbiol. Lett.* **2004**, *237* (2), 363−367.
(160) Zhao, Y.; Yan, B.; Yang, T.; Jiang, J.; Wei, H.; Zhu, X. Purification and Crystallographic Analysis of a Fad-Dependent Halogenase from Streptomyces Sp. JCM9888. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2015**, *71* (8), 972−976.
(161) Mantovani, S. M.; Moore, B. S. Flavin-Linked Oxidase Catalyzes Pyrrolizine Formation of Dichloropyrrole-Containing Polyketide Extender Unit in Chlorizidine A. *J. Am. Chem. Soc.* **2013**, *135* (48), 18032−18035.
(162) Alvarez-Mico, X.; Jensen, P. R.; Fenical, W.; Hughes, C. C. Chlorizidine, a Cytotoxic 5H-pyrrolo[2,1-a]isoindol-5-one-Containing Alkaloid from a Marine Streptomyces Sp. *Org. Lett.* **2013**, *15* (5), 988−991.
(163) Speck, K.; Magauer, T. The Chemistry of Isoindole Natural Products. *Beilstein J. Org. Chem.* **2013**, *9*, 2048−2078.
(164) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. Biosynthesis of Polybrominated Aromatic Organic Compounds by Marine Bacteria. *Nat. Chem. Biol.* **2014**, *10* (8), 640−647.
(165) Maansson, M.; Vynne, N. G.; Klitgaard, A.; Nybo, J. L.; Melchiorsen, J.; Nguyen, D. D.; Sanchez, L. M.; Ziemert, N.; Dorrestein, P. C.; Andersen, M. R. An Integrated Metabolomic and Genomic Mining Workflow to Uncover the Biosynthetic Potential of Bacteria. *mSystems* **2016**, *1* (3), e00028-15.
(166) Flecks, S.; Patallo, E. R.; Zhu, X. F.; Ernyei, A. J.; Seifert, G.; Schneider, A.; Dong, C. J.; Naismith, J. H.; van Pee, K. H. New Insights into the Mechanism of Enzymatic Chlorination of Tryptophan. *Angew. Chem., Int. Ed.* **2008**, *47* (49), 9533−9536.
(167) Zhou, H.; Qiao, K.; Gao, Z.; Vederas, J. C.; Tang, Y. Insights into Radicicol Biosynthesis Via Heterologous Synthesis of Intermediates and Analogs. *J. Biol. Chem.* **2010**, *285* (53), 41412−41421.
(168) Zeng, J.; Lytle, A. K.; Gage, D.; Johnson, S. J.; Zhan, J. Specific Chlorination of Isoquinolines by a Fungal Flavin-Dependent Halogenase. *Bioorg. Med. Chem. Lett.* **2013**, *23* (4), 1001−1003.
(169) Zeng, J.; Valiente, J.; Zhan, J. X. Specific Inhibition of the Halogenase for Radicicol Biosynthesis by Bromide at the Transcriptional Level in Pochonia Chlamydosporia. *Biotechnol. Lett.* **2011**, *33* (2), 333−338.
(170) Wang, S.; Xu, Y.; Maine, E. A.; Wijeratne, E. M.; Espinosa-Artiles, P.; Gunatilaka, A. A.; Molnar, I. Functional Characterization of the Biosynthesis of Radicicol, an Hsp90 Inhibitor Resorcylic Acid Lactone from Chaetomium Chiversii. *Chem. Biol.* **2008**, *15* (12), 1328−1338.
(171) Zhou, H.; Vederas, J. C.; Tang, Y. Insights into Radicicol Biosynthesis Via Heterologous Synthesis of Intermediates and Analogs. *Abstr. Pap. Am. Chem. Soc.* **2011**, *241*.
(172) Xu, F.; Merkley, A.; Yu, D.; Zhan, J. Selective Biochlorination of Hydroxyquinolines by a Flavin-Dependent Halogenase. *Tetrahedron Lett.* **2016**, *57* (47), 5262−5265.
(173) Banani, H.; Marcet-Houben, M.; Ballester, A. R.; Abbruscato, P.; Gonzalez-Candelas, L.; Gabaldon, T.; Spadaro, D. Genome Sequencing and Secondary Metabolism of the Postharvest Pathogen Penicillium Griseofulvum. *BMC Genomics* **2016**, *17* (19), 1−14.
(174) Cacho, R. A.; Chooi, Y. H.; Zhou, H.; Tang, Y. Complexity Generation in Fungal Polyketide Biosynthesis: A Spirocycle-Forming P450 in the Concise Pathway to the Antifungal Drug Griseofulvin. *ACS Chem. Biol.* **2013**, *8* (10), 2322−2330.
(175) Chooi, Y.-H.; Cacho, R.; Tang, Y. Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from Penicillium Aethiopicum. *Chem. Biol.* **2010**, *17* (5), 483−494.
(176) Chankhamjon, P.; Boettger-Schmidt, D.; Scherlach, K.; Urbansky, B.; Lackner, G.; Kalb, D.; Dahse, H. M.; Hoffmeister, D.; Hertweck, C. Biosynthesis of the Halogenated Mycotoxin Aspirochlorine in Koji Mold Involves a Cryptic Amino Acid Conversion. *Angew. Chem., Int. Ed.* **2014**, *53* (49), 13409−13413.
(177) Klausmeyer, P.; McCloud, T. G.; Tucker, K. D.; Cardellina, J. H., 2nd; Shoemaker, R. H. Aspirochlorine Class Compounds from Aspergillus Flavus Inhibit Azole-Resistant Candida Albicans. *J. Nat. Prod.* **2005**, *68* (8), 1300−1302.
(178) Monti, F.; Ripamonti, F.; Hawser, S. P.; Islam, K. Aspirochlorine: A Highly Selective and Potent Inhibitor of Fungal Protein Synthesis. *J. Antibiot.* **1999**, *52* (3), 311−318.
(179) Sato, M.; Winter, J. M.; Kishimoto, S.; Noguchi, H.; Tang, Y.; Watanabe, K. Combinatorial Generation of Chemical Diversity by Redox Enzymes in Chaetoviridin Biosynthesis. *Org. Lett.* **2016**, *18* (6), 1446−1449.
(180) Winter, J. M.; Sato, M.; Sugimoto, S.; Chiou, G.; Garg, N. K.; Tang, Y.; Watanabe, K. Identification and Characterization of the Chaetoviridin and Chaetomugilin Gene Cluster in Chaetomium Globosum Reveal Dual Functions of an Iterative Highly-Reducing Polyketide Synthase. *J. Am. Chem. Soc.* **2012**, *134* (43), 17900−17903.
(181) Kingsland, S. R.; Barrow, R. A. Identification of Chaetoviridin E from a Cultured Microfungus, Chaetomium Sp and Structural Reassignment of Chaetoviridins B and D. *Aust. J. Chem.* **2009**, *62* (3), 269−274.
(182) Massi, F. P.; Sartori, D.; Ferranti Lde, S.; Iamanaka, B. T.; Taniwaki, M. H.; Vieira, M. L.; Fungaro, M. H. Data on the Presence or Absence of Genes Encoding Essential Proteins for Ochratoxin and Fumonisin Biosynthesis in Aspergillus Niger and Aspergillus Welwitschiae. *Data Brief* **2016**, *7*, 704−708.
(183) Susca, A.; Proctor, R. H.; Morelli, M.; Haidukowski, M.; Gallo, A.; Logrieco, A. F.; Moretti, A. Variation in Fumonisin and Ochratoxin Production Associated with Differences in Biosynthetic Gene Content in Aspergillus Niger and A. Welwitschiae Isolates from Multiple Crop and Geographic Origins. *Front. Microbiol.* **2016**, *7*, 1 DOI: 10.3389/fmicb.2016.01412.
(184) Farber, P.; Geisen, R. Analysis of Differentially-Expressed Ochratoxin a Biosynthesis Genes of Penicillium Nordicum. *Eur. J. Plant Pathol.* **2004**, *110* (5−6), 661−669.
(185) Nielsen, M. T.; Nielsen, J. B.; Anyaogu, D. C.; Holm, D. K.; Nielsen, K. F.; Larsen, T. O.; Mortensen, U. H. Heterologous Reconstitution of the Intact Geodin Gene Cluster in Aspergillus Nidulans through a Simple and Versatile PCR Based Approach. *PLoS One* **2013**, *8* (8), e72871.
(186) Yin, Y.; Cai, M. H.; Zhou, X. S.; Li, Z. Y.; Zhang, Y. X. Polyketides in Aspergillus Terreus: Biosynthesis Pathway Discovery and Application. *Appl. Microbiol. Biotechnol.* **2016**, *100* (18), 7787−7798.
(187) Neumann, C. S.; Walsh, C. T.; Kay, R. R. A Flavin-Dependent Halogenase Catalyzes the Chlorination Step in the Biosynthesis of Dictyostelium Differentiation-Inducing Factor 1. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (13), 5798−5803.
(188) Thompson, C. R.; Kay, R. R. The Role of Dif-1 Signaling in Dictyostelium Development. *Mol. Cell* **2000**, *6* (6), 1509−1514.
(189) Mohri, K.; Hata, T.; Kikuchi, H.; Oshima, Y.; Urushihara, H. Defects in the Synthetic Pathway Prevent Dif-1 Mediated Stalk Lineage Specification Cascade in the Non-Differentiating Social Amoeba, Acytostelium Subglobosum. *Biol. Open* **2014**, *3* (6), 553−560.
(190) Motohashi, K. A.; Morita, N.; Kato, A.; Saito, T. Identification of Des-Methyl-Dif-1 Methyltransferase in Dictyostelium Purpureum. *Biosci., Biotechnol., Biochem.* **2012**, *76* (9), 1672−1676.
(191) Knobloch, T.; Harmrolfs, K.; Taft, F.; Thomaszewski, B.; Sasse, F.; Kirschning, A. Mutational Biosynthesis of Ansamitocin Antibiotics: A Diversity-Oriented Approach to Exploit Biosynthetic Flexibility. *ChemBioChem* **2011**, *12* (4), 540−547.
(192) Yu, T. W.; Bai, L.; Clade, D.; Hoffmann, D.; Toelzer, S.; Trinh, K. Q.; Xu, J.; Moss, S. J.; Leistner, E.; Floss, H. G. The Biosynthetic Gene Cluster of the Maytansinoid Antitumor Agent Ansamitocin from Actinosynnema Pretiosum. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (12), 7968−7973.
(193) Spiteller, P.; Bai, L.; Shang, G.; Carroll, B. J.; Yu, T. W.; Floss, H. G. The Post-Polyketide Synthase Modification Steps in the Biosynthesis of the Antitumor Agent Ansamitocin by Actinosynnema Pretiosum. *J. Am. Chem. Soc.* **2003**, *125* (47), 14236−14237.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(194) Wu, Y. Y.; Kang, Q. J.; Shen, Y. M.; Su, W. J.; Bai, L. Q. Cloning and Functional Analysis of the Naphthomycin Biosynthetic Gene Cluster in Streptomyces Sp Cs. *Mol. BioSyst.* **2011**, *7* (8), 2459−2469.
(195) Winter, J. M.; Moffitt, M. C.; Zazopoulos, E.; McAlpine, J. B.; Dorrestein, P. C.; Moore, B. S. Molecular Basis for Chloronium-Mediated Meroterpene Cyclization: Cloning, Sequencing, and Heterologous Expression of the Napyradiomycin Biosynthetic Gene Cluster. *J. Biol. Chem.* **2007**, *282* (22), 16362−16368.
(196) Lu, Y. H.; Yue, C. W.; Shao, M. Y.; Qian, S. Y.; Liu, N.; Bao, Y. X.; Wang, M.; Liu, M. H.; Li, X. Q.; Wang, Y. Y. Molecular Genetic Characterization of an Anthrabenzoxocinones Gene Cluster in Streptomyces Sp FJS31−2 for the Biosynthesis of BE-24566B and Zunyimycin Ale. *Molecules* **2016**, *21* (6), 711.
(197) Eustaquio, A. S.; Gust, B.; Luft, T.; Li, S. M.; Chater, K. F.; Heide, L. Clorobiocin Biosynthesis in Streptomyces: Identification of the Halogenase and Generation of Structural Analogs. *Chem. Biol.* **2003**, *10* (3), 279−288.
(198) Pojer, F.; Li, S. M.; Heide, L. Molecular Cloning and Sequence Analysis of the Clorobiocin Biosynthetic Gene Cluster: New Insights into the Biosynthesis of Aminocoumarin Antibiotics. *Microbiology* **2002**, *148*, 3901−3911.
(199) Puk, O.; Huber, P.; Bischoff, D.; Recktenwald, J.; Jung, G.; Sussmuth, R. D.; van Pee, K. H.; Wohlleben, W.; Pelzer, S. Glycopeptide Biosynthesis in Amycolatopsis Mediterranei DSM5908: Function of a Halogenase and a Haloperoxidase/Perhydrolase. *Chem. Biol.* **2002**, *9* (2), 225−235.
(200) Schmartz, P. C.; Zerbe, K.; Abou-Hadeed, K.; Robinson, J. A. Bis-Chlorination of a Hexapeptide-PCP Conjugate by the Halogenase Involved in Vancomycin Biosynthesis. *Org. Biomol. Chem.* **2014**, *12* (30), 5574−5577.
(201) Pelzer, S.; Sussmuth, R.; Heckmann, D.; Recktenwald, J.; Huber, P.; Jung, G.; Wohlleben, W. Identification and Analysis of the Balhimycin Biosynthetic Gene Cluster and Its Use for Manipulating Glycopeptide Biosynthesis in Amycolatopsis Mediterranei Dsm5908. *Antimicrob. Agents Chemother.* **1999**, *43*, 1565−1573.
(202) Chiu, H. T.; Hubbard, B. K.; Shah, A. N.; Eide, J.; Fredenburg, R. A.; Walsh, C. T.; Khosla, C. Molecular Cloning and Sequence Analysis of the Complestatin Biosynthetic Gene Cluster. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (15), 8548−8553.
(203) Park, O. K.; Choi, H. Y.; Kim, G. W.; Kim, W. G. Generation of New Complestatin Analogues by Heterologous Expression of the Complestatin Biosynthetic Gene Cluster from Streptomyces Chartreusis AN1542. *ChemBioChem* **2016**, *17* (18), 1725−1731.
(204) Sosio, M.; Stinchi, S.; Beltrametti, F.; Lazzarini, A.; Donadio, S. The Gene Cluster for the Biosynthesis of the Glycopeptide Antibiotic A40926 by Nonomuraea Species. *Chem. Biol.* **2003**, *10* (6), 541−549.
(205) Li, T. L.; Huang, F.; Haydock, S. F.; Mironenko, T.; Leadlay, P. F.; Spencer, J. B. Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase. *Chem. Biol.* **2004**, *11* (1), 107−119.
(206) Horbal, L.; Ostash, B.; Luzhetskyy, A.; Walker, S.; Kalinowski, J.; Fedorenko, V. A Gene Cluster for the Biosynthesis of Moenomycin Family Antibiotics in the Genome of Teicoplanin Producer Actinoplanes Teichomyceticus. *Appl. Microbiol. Biotechnol.* **2016**, *100* (17), 7629−7638.
(207) Sosio, M.; Bianchi, A.; Bossi, E.; Donadio, S. Teicoplanin Biosynthesis Genes in Actinoplanes Teichomyceticus. *Antonie van Leeuwenhoek* **2000**, *78* (3−4), 379−384.
(208) Buedenbender, S.; Rachid, S.; Müller, R.; Schulz, G. E. Structure and Action of the Myxobacterial Chondrochloren Halogenase CndH: A New Variant of FAD-Dependent Halogenases. *J. Mol. Biol.* **2009**, *385* (2), 520−530.
(209) Van Lanen, S. G.; Lin, S.; Horsman, G. P.; Shen, B. Characterization of Sgce6, the Flavin Reductase Component Supporting Fad-Dependent Halogenation and Hydroxylation in the Biosynthesis of the Enediyne Antitumor Antibiotic C-1027. *FEMS Microbiol. Lett.* **2009**, *300* (2), 237−241.
(210) Lin, S.; Huang, T.; Shen, B. Tailoring Enzymes Acting on Carrier Protein-Tethered Substrates in Natural Product Biosynthesis. *Methods Enzymol.* **2012**, *516*, 321−343.
(211) Lin, S. J.; Van Lanen, S. G.; Shen, B. Characterization of the Two-Component, Fad-Dependent Monooxygenase Sgcc That Requires Carrier Protein-Tethered Substrates for the Biosynthesis of the Enediyne Antitumor Antibiotic C-1027. *J. Am. Chem. Soc.* **2008**, *130* (20), 6616−6623.
(212) Chen, J.-S.; Su, M.; Shao, L.; Wang, Y.-X.; Lin, H.-M.; Chen, D.-J. Investigation of Halogenation During the Biosynthesis of Ramoplanin in Actinoplanes Sp. ATCC33076. *Appl. Microbiol. Biotechnol.* **2016**, *100* (1), 289−298.
(213) Yin, X. H.; Zabriskie, T. M. The Enduracidin Biosynthetic Gene Cluster from Streptomyces Fungicidicus. *Microbiology* **2006**, *152*, 2969−2983.
(214) Yin, X. H.; Chen, Y.; Zhang, L.; Wang, Y.; Zabriskie, T. M. Enduracidin Analogues with Altered Halogenation Patterns Produced by Genetically Engineered Strains of Streptomyces Fungicidicus. *J. Nat. Prod.* **2010**, *73* (4), 583−589.
(215) Han, J. G.; Chen, J. S.; Shao, L.; Zhang, J. L.; Dong, X. J.; Liu, P. Y.; Chen, D. J. Production of the Ramoplanin Activity Analogue by Double Gene Inactivation. *PLoS One* **2016**, *11* (5), e0154121.
(216) Rank, C.; Klejnstrup, M. L.; Petersen, L. M.; Kildgaard, S.; Frisvad, J. C.; Held Gotfredsen, C.; Ostenfeld Larsen, T. Comparative Chemistry of Aspergillus Oryzae (Rib40) and A. Flavus (NRRL 3357). *Metabolites* **2012**, *2* (1), 39−56.
(217) Chankhamjon, P.; Tsunematsu, Y.; Ishida-Ito, M.; Sasa, Y.; Meyer, F.; Boettger-Schmidt, D.; Urbansky, B.; Menzel, K.-D.; Scherlach, K.; Watanabe, K.; et al. Regioselective Dichlorination of a Non-Activated Aliphatic Carbon Atom and Phenolic Bismethylation by a Multifunctional Fungal Flavoenzyme. *Angew. Chem., Int. Ed.* **2016**, *55* (39), 11955−11959.
(218) Frese, M.; Guzowska, P. H.; Voß, H.; Sewald, N. Regioselective Enzymatic Halogenation of Substituted Tryptophan Derivatives Using the FAD-Dependent Halogenase RebH. *ChemCatChem* **2014**, *6* (5), 1270−1276.
(219) Hölzer, M.; Burd, W.; Reißig, H.-U.; Pée, K.-H. v. Substrate Specificity and Regioselectivity of Tryptophan 7-Halogenase from Pseudomonas Fluorescens BL915. *Adv. Synth. Catal.* **2001**, *343* (6−7), 591−595.
(220) Payne, J. T.; Andorfer, M. C.; Lewis, J. C. Regioselective Arene Halogenation Using the FAD-Dependent Halogenase RebH. *Angew. Chem., Int. Ed.* **2013**, *52* (20), 5271−5274.
(221) Payne, J. T.; Lewis, J. C. Upgrading Nature's Tools: Expression Enhancement and Preparative Utility of the Halogenase RebH. *Synlett* **2014**, *25*, A−E.
(222) Lang, A.; Polnick, S.; Nicke, T.; William, P.; Patallo, E. P.; Naismith, J. H.; van Pée, K.-H. Changing the Regioselectivity of the Tryptophan 7-Halogenase PrnA by Site-Directed Mutagenesis. *Angew. Chem., Int. Ed.* **2011**, *50* (13), 2951−2953.
(223) Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C. Directed Evolution of RebH for Catalyst-Controlled Halogenation of Indole C-H Bonds. *Chem. Sci.* **2016**, *7* (6), 3720−3729.
(224) Poor, C. B.; Andorfer, M. C.; Lewis, J. C. Improving the Stability and Catalyst Lifetime of the Halogenase RebH by Directed Evolution. *ChemBioChem* **2014**, *15* (9), 1286−1289.
(225) Payne, J. T.; Poor, C. B.; Lewis, J. C. Directed Evolution of RebH for Site-Selective Halogenation of Large Biologically Active Molecules. *Angew. Chem., Int. Ed.* **2015**, *54* (14), 4226−4230.
(226) Bar-Even, A.; Noor, E.; Savir, Y.; Liebermeister, W.; Davidi, D.; Tawfik, D. S.; Milo, R. The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters. *Biochemistry* **2011**, *50* (21), 4402−4410.
(227) Muffler, K.; Retzlaff, M.; van Pée, K.-H.; Ulber, R. Optimisation of Halogenase Enzyme Activity by Application of a Genetic Algorithm. *J. Biotechnol.* **2007**, *127* (3), 425−433.
(228) Liao, H. H. Thermostable Mutants of Kanamycin Nucleotidyltransferase Are Also More Stable to Proteinase K, Urea, Detergents,

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

and Water-Miscible Organic Solvents. *Enzyme Microb. Technol.* **1993**, *15* (4), 286−292.
(229) Zhao, H.; Arnold, F. H. Directed Evolution Converts Subtilisin E into a Functional Equivalent of Thermitase. *Protein Eng., Des. Sel.* **1999**, *12* (1), 47−53.
(230) Haki, G. D.; Rakshit, S. K. Developments in Industrially Important Thermostable Enzymes: A Review. *Bioresour. Technol.* **2003**, *89* (1), 17−34.
(231) Wu, I.; Arnold, F. H. Engineered Thermostable Fungal Cel6a and Cel7a Cellobiohydrolases Hydrolyze Cellulose Efficiently at Elevated Temperatures. *Biotechnol. Bioeng.* **2013**, *110* (7), 1874−1883.
(232) Chen, X.; Zhang, B.; Zhang, W.; Wu, X.; Zhang, M.; Chen, T.; Liu, G.; Dyson, P. Genome Sequence of Streptomyces Violaceusniger Strain SPC6, a Halotolerant Streptomycete That Exhibits Rapid Growth and Development. *Genome Announc* **2013**, *1* (4), e00494-13.
(233) Harrison, J.; Studholme, D. J. Recently Published Streptomyces Genome Sequences. *Microb. Biotechnol.* **2014**, *7* (5), 373−380.
(234) Takahashi, S.; Furuya, T.; Ishii, Y.; Kino, K.; Kirimura, K. Characterization of a Flavin Reductase from a Thermophilic Dibenzothiophene-Desulfurizing Bacterium, Bacillus Subtilis WU-S2B. *J. Biosci. Bioeng.* **2009**, *107* (1), 38−41.
(235) Strickler, S. S.; Gribenko, A. V.; Gribenko, A. V.; Keiffer, T. R.; Tomlinson, J.; Reihle, T.; Loladze, V. V.; Makhatadze, G. I. Protein Stability and Surface Electrostatics: A Charged Relationship. *Biochemistry* **2006**, *45* (9), 2761−2766.
(236) Smith, M. D.; Rosenow, M. A.; Wang, M.; Allen, J. P.; Szostak, J. W.; Chaput, J. C. Structural Insights into the Evolution of a Non-Biological Protein: Importance of Surface Residues in Protein Fold Optimization. *PLoS One* **2007**, *2* (5), e467.
(237) Chan, C.-H.; Wilbanks, C. C.; Makhatadze, G. I.; Wong, K.-B. Electrostatic Contribution of Surface Charge Residues to the Stability of a Thermophilic Protein: Benchmarking Experimental and Predicted Pka Values. *PLoS One* **2012**, *7* (1), e30296.
(238) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *J. Am. Chem. Soc.* **2007**, *129* (33), 10110−10112.
(239) Rees, D. C.; Robertson, A. D. Some Thermodynamic Implications for the Thermostability of Proteins. *Protein Sci.* **2001**, *10* (6), 1187−1194.
(240) Schnepel, C.; Minges, H.; Frese, M.; Sewald, N. A High-Throughput Fluorescence Assay to Determine the Activity of Tryptophan Halogenases. *Angew. Chem., Int. Ed.* **2016**, *55* (45), 14159−14163.
(241) Hosford, J.; Shepherd, S. A.; Micklefield, J.; Wong, L. S. A High-Throughput Assay for Arylamine Halogenation Based on a Peroxidase-Mediated Quinone−Amine Coupling with Applications in the Screening of Enzymatic Halogenations. *Chem. - Eur. J.* **2014**, *20* (50), 16759−16763.
(242) Belsare, K. D.; Andorfer, M. C.; Cardenas, F. S.; Chael, J. R.; Park, H. J.; Lewis, J. C. A Simple Combinatorial Codon Mutagenesis Method for Targeted Protein Engineering. *ACS Synth. Biol.* **2017**, *6* (3), 416−420.
(243) Latham, J.; Henry, J.-M.; Sharif, H. H.; Menon, B. R. K.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J. Integrated Catalysis Opens New Arylation Pathways Via Regiodivergent Enzymatic C−H Activation. *Nat. Commun.* **2016**, *7*, 11873.
(244) Frese, M.; Sewald, N. Enzymatic Halogenation of Tryptophan on a Gram Scale. *Angew. Chem., Int. Ed.* **2015**, *54* (1), 298−301.
(245) Sheldon, R. A.; Arends, I. W. C. E.; Hanefeld, U. *Green Chemistry and Catalysis*; Wiley-VCH Verlag GmbH: Weinheim, 2007.
(246) Cao, L.; van Rantwijk, F.; Sheldon, R. A. Cross-Linked Enzyme Aggregates: A Simple and Effective Method for the Immobilization of Penicillin Acylase. *Org. Lett.* **2000**, *2* (10), 1361−1364.
(247) Sheldon, R. A. Enzyme Immobilization: The Quest for Optimum Performance. *Adv. Synth. Catal.* **2007**, *349* (8−9), 1289−1307.
(248) van Pelt, S.; Quignard, S.; Kubac, D.; Sorokin, D. Y.; van Rantwijk, F.; Sheldon, R. A. Nitrile Hydratase Cleas: The Immobilization and Stabilization of an Industrially Important Enzyme. *Green Chem.* **2008**, *10* (4), 395−400.
(249) Runguphan, W.; Qu, X.; O'Connor, S. E. Integrating Carbon-Halogen Bond Formation into Medicinal Plant Metabolism. *Nature* **2010**, *468* (7322), 461−464.
(250) Glenn, W. S.; Nims, E.; O'Connor, S. E. Reengineering a Tryptophan Halogenase to Preferentially Chlorinate a Direct Alkaloid Precursor. *J. Am. Chem. Soc.* **2011**, *133* (48), 19346−19349.
(251) Roy, A. D.; Grüschow, S.; Cairns, N.; Goss, R. J. M. Gene Expression Enabling Synthetic Diversification of Natural Products: Chemogenetic Generation of Pacidamycin Analogs. *J. Am. Chem. Soc.* **2010**, *132* (35), 12243−12245.
(252) Sánchez, C.; Zhu, L.; Braña, A. F.; Salas, A. P.; Rohr, J.; Méndez, C.; Salas, J. A. Combinatorial Biosynthesis of Antitumor Indolocarbazole Compounds. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (2), 461−466.
(253) Kim, E.; Moore, B. S.; Yoon, Y. J. Reinvigorating Natural Product Combinatorial Biosynthesis with Synthetic Biology. *Nat. Chem. Biol.* **2015**, *11* (9), 649−659.
(254) Du, Y.-L.; Ryan, K. S. Expansion of Bisindole Biosynthetic Pathways by Combinatorial Construction. *ACS Synth. Biol.* **2015**, *4* (6), 682−688.
(255) Wenda, S.; Illner, S.; Mell, A.; Kragl, U. Industrial Biotechnology-the Future of Green Chemistry? *Green Chem.* **2011**, *13* (11), 3007−3047.
(256) Grüschow, S.; Rackham, E. J.; Elkins, B.; Newill, P. L. A.; Hill, L. M.; Goss, R. J. M. New Pacidamycin Antibiotics through Precursor-Directed Biosynthesis. *ChemBioChem* **2009**, *10* (2), 355−360.
(257) Fräbel, S.; Krischke, M.; Staniek, A.; Warzecha, H. Recombinant Flavin-Dependent Halogenases Are Functional in Tobacco Chloroplasts without Co-Expression of Flavin Reductase Genes. *Biotechnol. J.* **2016**, *11* (12), 1586−1594.
(258) Wang, S.; Zhang, S.; Xiao, A.; Rasmussen, M.; Skidmore, C.; Zhan, J. Metabolic Engineering of Escherichia Coli for the Biosynthesis of Various Phenylpropanoid Derivatives. *Metab. Eng.* **2015**, *29*, 153−159.
(259) Li, X.-Z.; Wei, X.; Zhang, C.-J.; Jin, X.-L.; Tang, J.-J.; Fan, G.-J.; Zhou, B. Hypohalous Acid-Mediated Halogenation of Resveratrol and Its Role in Antioxidant and Antimicrobial Activities. *Food Chem.* **2012**, *135* (3), 1239−1244.
(260) Weissman, K. J. Mutasynthesis − Uniting Chemistry and Genetics for Drug Discovery. *Trends Biotechnol.* **2007**, *25* (4), 139−142.
(261) Durak, L. J.; Payne, J. T.; Lewis, J. C. Late-Stage Diversification of Biologically Active Molecules Via Chemoenzymatic C−H Functionalization. *ACS Catal.* **2016**, *6* (3), 1451−1454.
(262) Gröger, H.; Hummel, W. Combining the 'Two Worlds' of Chemocatalysis and Biocatalysis Towards Multi-Step One-Pot Processes in Aqueous Media. *Curr. Opin. Chem. Biol.* **2014**, *19* (0), 171−179.
(263) Reetz, M. T. Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **2013**, *135* (34), 12480−12496.
(264) Sato, H.; Hummel, W.; Gröger, H. Cooperative Catalysis of Noncompatible Catalysts through Compartmentalization: Wacker Oxidation and Enzymatic Reduction in a One-Pot Process in Aqueous Media. *Angew. Chem., Int. Ed.* **2015**, *54* (15), 4488−4492.
(265) Bruggink, A.; Schoevaart, R.; Kieboom, T. Concepts of Nature in Organic Synthesis: Cascade Catalysis and Multistep Conversions in Concert. *Org. Process Res. Dev.* **2003**, *7* (5), 622−640.
(266) Kilpin, K. J.; Dyson, P. J. Enzyme Inhibition by Metal Complexes: Concepts, Strategies and Applications. *Chem. Sci.* **2013**, *4* (4), 1410−1419.
(267) Frese, M.; Schnepel, C.; Minges, H.; Voß, H.; Feiner, R.; Sewald, N. Modular Combination of Enzymatic Halogenation of Tryptophan with Suzuki−Miyaura Cross-Coupling Reactions. *ChemCatChem* **2016**, *8* (10), 1799−1803.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(268) Long, T. R.; Bowden, N. B. Polydimethylsiloxane Thimbles. *Encyclopedia of Reagents for Organic Synthesis* **2001**, DOI: 10.1002/047084289X.re130.
(269) MillerIi, A. L.; Bowden, N. B. Site-Isolation and Recycling of Pdcl2 Using Pdms Thimbles. *J. Org. Chem.* **2009**, *74* (13), 4834−4840.
(270) Yang, Y.; Li, R.; Zhao, Y.; Zhao, D.; Shi, Z. Cu-Catalyzed Direct C6-Arylation of Indoles. *J. Am. Chem. Soc.* **2016**, *138* (28), 8734−8737.
(271) Yang, Y.; Qiu, X.; Zhao, Y.; Mu, Y.; Shi, Z. Palladium-Catalyzed C−H Arylation of Indoles at the C7 Position. *J. Am. Chem. Soc.* **2016**, *138* (2), 495−498.
(272) Holtmann, D.; Fraaije, M. W.; Arends, I. W. C. E.; Opperman, D. J.; Hollmann, F. The Taming of Oxygen: Biocatalytic Oxyfunctionalisations. *Chem. Commun.* **2014**, *50* (87), 13180−13200.
(273) Huijbers, M. M. E.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J. H. Flavin Dependent Monooxygenases. *Arch. Biochem. Biophys.* **2014**, *544*, 2−17.
(274) Cochrane, R. V.; Vederas, J. C. Highly Selective but Multifunctional Oxygenases in Secondary Metabolism. *Acc. Chem. Res.* **2014**, *47* (10), 3148−3161.
(275) Holtmann, D.; Hollmann, F. The Oxygen Dilemma: A Severe Challenge for the Application of Monooxygenases? *ChemBioChem* **2016**, *17* (15), 1391−1398.
(276) Bučko, M.; Gemeiner, P.; Schenkmayerová, A.; Krajčovič, T.; Rudroff, F.; Mihovilovič, M. D. Baeyer-Villiger Oxidations: Biotechnological Approach. *Appl. Microbiol. Biotechnol.* **2016**, *100* (15), 6585−6599.
(277) Leisch, H.; Morley, K.; Lau, P. C. K. Baeyer−Villiger Monooxygenases: More Than Just Green Chemistry. *Chem. Rev.* **2011**, *111* (7), 4165−4222.
(278) Ellis, H. R. The Fmn-Dependent Two-Component Monooxygenase Systems. *Arch. Biochem. Biophys.* **2010**, *497* (2), 1−12.
(279) Suske, W. A.; Held, M.; Schmid, A.; Fleischmann, T.; Wubbolts, M. G.; Kohler, H.-P. E. Purification and Characterization of 2-Hydroxybiphenyl 3-Monooxygenase, a Novel NADH-Dependent, FAD-Containing Aromatic Hydroxylase from Pseudomonas Azelaica Hbp1. *J. Biol. Chem.* **1997**, *272* (39), 24257−24265.
(280) Schlaich, N. L. Flavin-Containing Monooxygenases in Plants: Looking Beyond Detox. *Trends Plant Sci.* **2007**, *12* (9), 412−418.
(281) Crosa, J. H.; Walsh, C. T. Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. *Microbiol. Mol. Biol. Rev.* **2002**, *66* (2), 223−249.
(282) Laden, B. P.; Tang, Y.; Porter, T. D. Cloning, Heterologous Expression, and Enzymological Characterization of Human Squalene Monooxygenase. *Arch. Biochem. Biophys.* **2000**, *374* (2), 381−388.
(283) Breton, J.; Avanzi, N.; Magagnin, S.; Covini, N.; Magistrelli, G.; Cozzi, L.; Isacchi, A. Functional Characterization and Mechanism of Action of Recombinant Human Kynurenine 3-Hydroxylase. *Eur. J. Biochem.* **2000**, *267* (4), 1092−1099.
(284) Schwarcz, R.; Bruno, J. P.; Muchowski, P. J.; Wu, H.-Q. Kynurenines in the Mammalian Brain: When Physiology Meets Pathology. *Nat. Rev. Neurosci.* **2012**, *13* (7), 465−477.
(285) Amaral, M.; Levy, C.; Heyes, D. J.; Lafite, P.; Outeiro, T. F.; Giorgini, F.; Leys, D.; Scrutton, N. S. Structural Basis of Kynurenine 3-Monooxygenase Inhibition. *Nature* **2013**, *496* (7445), 382−385.
(286) Mooney, A.; Ward, P. G.; O'Connor, K. E. Microbial Degradation of Styrene: Biochemistry, Molecular Genetics, and Perspectives for Biotechnological Applications. *Appl. Microbiol. Biotechnol.* **2006**, *72* (1), 1.
(287) ten Brink, G. J.; Arends, I. W. C. E.; Sheldon, R. A. The Baeyer−Villiger Reaction: New Developments toward Greener Procedures. *Chem. Rev.* **2004**, *104* (9), 4105−4124.
(288) Berry, D. F.; Francis, A. J.; Bollag, J. M. Microbial Metabolism of Homocyclic and Heterocyclic Aromatic Compounds under Anaerobic Conditions. *Microbiol. Rev.* **1987**, *51*, 43−59.
(289) Alfieri, A.; Malito, E.; Orru, R.; Fraaije, M. W.; Mattevi, A. Revealing the Moonlighting Role of NADP in the Structure of a Flavin-Containing Monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (18), 6572−6577.
(290) Malito, E.; Alfieri, A.; Fraaije, M. W.; Mattevi, A. Crystal Structure of a Baeyer−Villiger Monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (36), 13157−13162.
(291) Tischler, D.; Schlömann, M.; van Berkel, W. J. H.; Gassner, G. T. Fad C(4a)-Hydroxide Stabilized in a Naturally Fused Styrene Monooxygenase. *FEBS Lett.* **2013**, *587* (23), 3848−3852.
(292) Tischler, D.; Gröning, J. A. D.; Kaschabek, S. R.; Schlömann, M. One-Component Styrene Monooxygenases: An Evolutionary View on a Rare Class of Flavoproteins. *Appl. Biochem. Biotechnol.* **2012**, *167* (5), 931−944.
(293) Massey, V. Activation of Molecular Oxygen by Flavins and Flavoproteins. *J. Biol. Chem.* **1994**, *269*, 22459−22462.
(294) Orru, R.; Dudek, H. M.; Martinoli, C.; Torres Pazmiño, D. E.; Royant, A.; Weik, M.; Fraaije, M. W.; Mattevi, A. Snapshots of Enzymatic Baeyer-Villiger Catalysis: Oxygen Activation and Intermediate Stabilization. *J. Biol. Chem.* **2011**, *286* (33), 29284−29291.
(295) Meyer, A.; Schmid, A.; Held, M.; Westphal, A. H.; Röthlisberger, M.; Kohler, H.-P. E.; van Berkel, W. J. H.; Witholt, B. Changing the Substrate Reactivity of 2-Hydroxybiphenyl 3-Monooxygenase from Pseudomonas Azelaica Hbp1 by Directed Evolution. *J. Biol. Chem.* **2002**, *277* (7), 5575−5582.
(296) Baldwin, C. V. F.; Woodley, J. M. On Oxygen Limitation in a Whole Cell Biocatalytic Baeyer−Villiger Oxidation Process. *Biotechnol. Bioeng.* **2006**, *95* (3), 362−369.
(297) O'Sullivan, L. M.; Patel, S.; Ward, J. M.; Woodley, J. M.; Doig, S. D. Large Scale Production of Cyclohexanone Monooxygenase from Escherichia Coli Top10 PQR239. *Enzyme Microb. Technol.* **2001**, *28* (2), 265−274.
(298) Schmid, A.; Vereyken, I.; Held, M.; Witholt, B. Preparative Regio- and Chemoselective Functionalization of Hydrocarbons Catalyzed by Cell Free Preparations of 2-Hydroxybiphenyl 3-Monooxygenase. *J. Mol. Catal. B: Enzym.* **2001**, *11* (4−6), 455−462.
(299) Churakova, E.; Kluge, M.; Ullrich, R.; Arends, I.; Hofrichter, M.; Hollmann, F. Specific Photobiocatalytic Oxyfunctionalization Reactions. *Angew. Chem., Int. Ed.* **2011**, *50* (45), 10716−10719.
(300) Hollmann, F.; Taglieber, A.; Schulz, F.; Reetz, M. T. A Light-Driven Stereoselective Biocatalytic Oxidation. *Angew. Chem., Int. Ed.* **2007**, *46* (16), 2903−2906.
(301) Hollmann, F.; Hofstetter, K.; Habicher, T.; Hauer, B.; Schmid, A. Direct Electrochemical Regeneration of Monooxygenase Subunits for Biocatalytic Asymmetric Epoxidation. *J. Am. Chem. Soc.* **2005**, *127* (18), 6540−6541.
(302) Hollmann, F.; Lin, P.-C.; Witholt, B.; Schmid, A. Stereospecific Biocatalytic Epoxidation: The First Example of Direct Regeneration of a Fad-Dependent Monooxygenase for Catalysis. *J. Am. Chem. Soc.* **2003**, *125* (27), 8209−8217.
(303) Paul, C. E.; Tischler, D.; Riedel, A.; Heine, T.; Itoh, N.; Hollmann, F. Nonenzymatic Regeneration of Styrene Monooxygenase for Catalysis. *ACS Catal.* **2015**, *5* (5), 2961−2965.
(304) Mifsud, M.; Gargiulo, S.; Iborra, S.; Arends, I. W. C. E.; Hollmann, F.; Corma, A. Photobiocatalytic Chemistry of Oxidoreductases Using Water as the Electron Donor. *Nat. Commun.* **2014**, *5*, 1 DOI: 10.1038/ncomms4145.
(305) Schrewe, M.; Julsing, M. K.; Buhler, B.; Schmid, A. Whole-Cell Biocatalysis for Selective and Productive C-O Functional Group Introduction and Modification. *Chem. Soc. Rev.* **2013**, *42* (15), 6346−6377.
(306) Held, M.; Schmid, A.; Kohler, H.-P. E.; Suske, W.; Witholt, B.; Wubbolts, M. G. An Integrated Process for the Production of Toxic Catechols from Toxic Phenols Based on a Designer Biocatalyst. *Biotechnol. Bioeng.* **1999**, *62* (6), 641−648.
(307) Held, M.; Suske, W.; Schmid, A.; Engesser, K.-H.; Kohler, H.-P. E.; Witholt, B.; Wubbolts, M. G. Preparative Scale Production of 3-Substituted Catechols Using a Novel Monooxygenase from Pseudomonas Azelaica HBP1. *J. Mol. Catal. B: Enzym.* **1998**, *5* (1−4), 87−93.
(308) Hilker, I.; Gutierrez, M. C.; Furstoss, R.; Ward, J.; Wohlgemuth, R.; Alphand, V. Preparative Scale Baeyer-Villiger Biooxidation at High Concentration Using Recombinant Escherichia

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

Coli and in Situ Substrate Feeding and Product Removal Process. *Nat. Protoc.* **2008**, *3* (3), 546−554.
(309) Hilker, I.; Wohlgemuth, R.; Alphand, V.; Furstoss, R. Microbial Transformations 59: First Kilogram Scale Asymmetric Microbial Baeyer-Villiger Oxidation with Optimized Productivity Using a Resin-Based in situ SFPR Strategy. *Biotechnol. Bioeng.* **2005**, *92* (6), 702−710.
(310) Hilker, I.; Alphand, V.; Wohlgemuth, R.; Furstoss, R. Microbial Transformations, 56. Preparative Scale Asymmetric Baeyer−Villiger Oxidation Using a Highly Productive "Two-in-One" Resin-Based in situ SFPR Concept. *Adv. Synth. Catal.* **2004**, *346* (2−3), 203−214.
(311) Hilker, I.; Gutiérrez, M. C.; Alphand, V.; Wohlgemuth, R.; Furstoss, R. Microbiological Transformations 57. Facile and Efficient Resin-Based in situ SFPR Preparative-Scale Synthesis of an Enantiopure "Unexpected" Lactone Regioisomer Via a Baeyer−Villiger Oxidation Process. *Org. Lett.* **2004**, *6* (12), 1955−1958.
(312) Doig, S. D.; Avenell, P. J.; Bird, P. A.; Gallati, P.; Lander, K. S.; Lye, G. J.; Wohlgemuth, R.; Woodley, J. M. Reactor Operation and Scale-up of Whole Cell Baeyer-Villiger Catalyzed Lactone Synthesis. *Biotechnol. Prog.* **2002**, *18* (5), 1039−1046.
(313) Baldwin, C. V. F.; Wohlgemuth, R.; Woodley, J. M. The First 200-L Scale Asymmetric Baeyer−Villiger Oxidation Using a Whole-Cell Biocatalyst. *Org. Process Res. Dev.* **2008**, *12* (4), 660−665.
(314) Simpson, H. D.; Alphand, V.; Furstoss, R. Microbiological Transformations: 49. Asymmetric Biocatalysed Baeyer−Villiger Oxidation: Improvement Using a Recombinant Escherichia Coli Whole Cell Biocatalyst in the Presence of an Adsorbent Resin. *J. Mol. Catal. B: Enzym.* **2001**, *16* (2), 101−108.
(315) Melgarejo-Torres, R.; Castillo-Araiza, C. O.; Dutta, A.; Bény, G.; Torres-Martinez, D.; Gutiérrez-Rojas, M.; Lye, G. J.; Huerta-Ochoa, S. Mathematical Model of a Three Phase Partitioning Bioreactor for Conversion of Ketones Using Whole Cells. *Chem. Eng. J.* **2015**, *260*, 765−775.
(316) Yang, J.; Wang, S.; Lorrain, M.-J.; Rho, D.; Abokitse, K.; Lau, P. C. K. Bioproduction of Lauryl Lactone and 4-Vinyl Guaiacol as Value-Added Chemicals in Two-Phase Biotransformation Systems. *Appl. Microbiol. Biotechnol.* **2009**, *84* (5), 867−876.
(317) Panke, S.; Held, M.; Wubbolts, M. G.; Witholt, B.; Schmid, A. Pilot-Scale Production of (S)-Styrene Oxide from Styrene by Recombinant Escherichia Coli Synthesizing Styrene Monooxygenase. *Biotechnol. Bioeng.* **2002**, *80* (1), 33−41.
(318) Bong, Y. K.; Clay, M. D.; Collier, Steven, J.; Mijts, B.; Vogel, M.; Zhang, X.; Zhu, J.; Nazor, J.; Smith, D. J.; Song, S. Codexis, Inc.: U.S., 2016.
(319) Reetz, M. T.; Wu, S. Greatly Reduced Amino Acid Alphabets in Directed Evolution: Making the Right Choice for Saturation Mutagenesis at Homologous Enzyme Positions. *Chem. Commun.* **2008**, *2008* (43), 5499−5501.
(320) Schulz, F.; Leca, F.; Hollmann, F.; Reetz, M. T. Towards Practical Biocatalytic Baeyer-Villiger Reactions: Applying a Thermostable Enzyme in the Gram-Scale Synthesis of Optically-Active Lactones in a Two-Liquid-Phase System. *Beilstein J. Org. Chem.* **2005**, *1*, 10.
(321) Reetz, M. T.; Daligault, F.; Brunner, B.; Hinrichs, H.; Deege, A. Directed Evolution of Cyclohexanone Monooxygenases: Enantioselective Biocatalysts for the Oxidation of Prochiral Thioethers. *Angew. Chem., Int. Ed.* **2004**, *43* (31), 4078−4081.
(322) Reetz, M. T.; Brunner, B.; Schneider, T.; Schulz, F.; Clouthier, C. M.; Kayser, M. M. Directed Evolution as a Method to Create Enantioselective Cyclohexanone Monooxygenases for Catalysis in Baeyer−Villiger Reactions. *Angew. Chem., Int. Ed.* **2004**, *43* (31), 4075−4078.
(323) Clouthier, C. M.; Kayser, M. M.; Reetz, M. T. Designing New Baeyer−Villiger Monooxygenases Using Restricted Casting. *J. Org. Chem.* **2006**, *71* (22), 8431−8437.
(324) Watts, A. B.; Beecher, J.; Whitcher, C. S.; Littlechild, J. A. A Method for Screening Baeyer-Villiger Monooxygenase Activity against Monocyclic Ketones. *Biocatal. Biotransform.* **2002**, *20* (3), 209−215.
(325) Gutierrez, M. C.; Sleegers, A.; Simpson, H. D.; Alphand, V.; Furstoss, R. The First Fluorogenic Assay for Detecting a Baeyer-Villigerase Activity in Microbial Cells. *Org. Biomol. Chem.* **2003**, *1* (20), 3500−3506.
(326) Wahler, D.; Reymond, J.-L. The Adrenaline Test for Enzymes. *Angew. Chem., Int. Ed.* **2002**, *41* (7), 1229−1232.
(327) Kirschner, A.; Bornscheuer, U. T. Directed Evolution of a Baeyer−Villiger Monooxygenase to Enhance Enantioselectivity. *Appl. Microbiol. Biotechnol.* **2008**, *81* (3), 465−472.
(328) Linares-Pastén, J. A.; Chávez-Lizárraga, G.; Villagomez, R.; Mamo, G.; Hatti-Kaul, R. A Method for Rapid Screening of Ketone Biotransformations: Detection of Whole Cell Baeyer−Villiger Monooxygenase Activity. *Enzyme Microb. Technol.* **2012**, *50* (2), 101−106.
(329) Sicard, R.; Chen, L. S.; Marsaioli, A. J.; Reymond, J.-L. A Fluorescence-Based Assay for Baeyer−Villiger Monooxygenases, Hydroxylases and Lactonases. *Adv. Synth. Catal.* **2005**, *347* (7−8), 1041−1050.
(330) Balke, K.; Kadow, M.; Mallin, H.; Sa; Bornscheuer, U. T. Discovery, Application and Protein Engineering of Baeyer-Villiger Monooxygenases for Organic Synthesis. *Org. Biomol. Chem.* **2012**, *10* (31), 6249−6265.
(331) Meyer, A.; Würsten, M.; Schmid, A.; Kohler, H.-P. E.; Witholt, B. Hydroxylation of Indole by Laboratory-Evolved 2-Hydroxybiphenyl 3-Monooxygenase. *J. Biol. Chem.* **2002**, *277* (37), 34161−34167.
(332) Balke, K.; Schmidt, S.; Genz, M.; Bornscheuer, U. T. Switching the Regioselectivity of a Cyclohexanone Monooxygenase toward (+)-Trans-Dihydrocarvone by Rational Protein Design. *ACS Chem. Biol.* **2016**, *11* (1), 38−43.
(333) Wu, S.; Acevedo, J. P.; Reetz, M. T. Induced Allostery in the Directed Evolution of an Enantioselective Baeyer−Villiger Monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (7), 2775−2780.
(334) Opperman, D. J.; Reetz, M. T. Towards Practical Baeyer−Villiger-Monooxygenases: Design of Cyclohexanone Monooxygenase Mutants with Enhanced Oxidative Stability. *ChemBioChem* **2010**, *11* (18), 2589−2596.
(335) Deng, H.; O'Hagan, D.; Schaffrath, C. Fluorometabolite Biosynthesis and the Fluorinase from Streptomyces Cattleya. *Nat. Prod. Rep.* **2004**, *21* (6), 773−784.
(336) O'Hagan, D.; Deng, H. Enzymatic Fluorination and Biotechnological Developments of the Fluorinase. *Chem. Rev.* **2015**, *115* (2), 634−649.
(337) Harper, D. B.; O'Hagan, D. The Fluorinated Natural Products. *Nat. Prod. Rep.* **1994**, *11* (0), 123−133.
(338) Wagner, C.; El Omari, M.; König, G. M. Biohalogenation: Nature's Way to Synthesize Halogenated Metabolites. *J. Nat. Prod.* **2009**, *72* (3), 540−553.
(339) O'Hagan, D.; Perry, R.; Lock, J. M.; Meyer, J. J. M.; Dasaradhi, L.; Hamilton, J. T. G.; Harper, D. B. High Levels of Monofluoroacetate in Dichapetalum Braunii. *Phytochemistry* **1993**, *33* (5), 1043−1045.
(340) Zechel, D. L.; Reid, S. P.; Stoll, D.; Nashiru, O.; Warren, R. A. J.; Withers, S. G. Mechanism, Mutagenesis, and Chemical Rescue of a B-Mannosidase from Cellulomonas Fimi. *Biochemistry* **2003**, *42* (23), 7195−7204.
(341) Nashiru, O.; Zechel, D. L.; Stoll, D.; Mohammadzadeh, T.; Warren, R. A. J.; Withers, S. G. B-Mannosynthase: Synthesis of B-Mannosides with a Mutant B-Mannosidase. *Angew. Chem., Int. Ed.* **2001**, *40* (2), 417−420.
(342) Zechel, D. L.; Reid, S. P.; Nashiru, O.; Mayer, C.; Stoll, D.; Jakeman, D. L.; Warren, R. A. J.; Withers, S. G. Enzymatic Synthesis of Carbon−Fluorine Bonds. *J. Am. Chem. Soc.* **2001**, *123* (18), 4350−4351.
(343) Maguire, A. R.; Meng, W.-D.; Roberts, S. M.; Willetts, A. J. Synthetic Approaches Towards Nucleocidin and Selected Analogues; Anti-Hiv Activity in 4′-Fluorinated Nucleoside Derivatives. *J. Chem. Soc., Perkin Trans. 1* **1993**, No. 15, 1795−1808.
(344) Sanada, M.; Miyano, T.; Iwadare, S.; Williamson, J. M.; Arison, B. H. J.; Smith, L.; Douglas, A. W.; Liesch, J. M.; Inamine, E. *J. Antibiot.* **1986**, *39*, 259.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(345) O'Hagan, D.; Schaffrath, C.; Cobb, S. L.; Hamilton, J. T. G.; Murphy, C. D. Biochemistry: Biosynthesis of an Organofluorine Molecule. *Nature* **2002**, *416* (6878), 279−279.
(346) Cadicamo, C. D.; Courtieu, J.; Deng, H.; Meddour, A.; O'Hagan, D. Enzymatic Fluorination in Streptomyces Cattleya Takes Place with an Inversion of Configuration Consistent with an Sn2 Reaction Mechanism. *ChemBioChem* **2004**, *5* (5), 685−690.
(347) O'Hagan, D.; Goss, R. J. M.; Meddour, A.; Courtieu, J. Assay for the Enantiomeric Analysis of [^{2}H1]-Fluoroacetic Acid: Insight into the Stereochemical Course of Fluorination During Fluorometabolite Biosynthesis in Streptomyces Cattleya. *J. Am. Chem. Soc.* **2003**, *125* (2), 379−387.
(348) Dong, C.; Huang, F.; Deng, H.; Schaffrath, C.; Spencer, J. B.; O'Hagan, D.; Naismith, J. H. Crystal Structure and Mechanism of a Bacterial Fluorinating Enzyme. *Nature* **2004**, *427* (6974), 561−565.
(349) Zhu, X.; Robinson, D. A.; McEwan, A. R.; O'Hagan, D.; Naismith, J. H. Mechanism of Enzymatic Fluorination in Streptomyces Cattleya. *J. Am. Chem. Soc.* **2007**, *129* (47), 14597−14604.
(350) Senn, H. M.; O'Hagan, D.; Thiel, W. Insight into Enzymatic C−F Bond Formation from Qm and Qm/Mm Calculations. *J. Am. Chem. Soc.* **2005**, *127* (39), 13643−13655.
(351) Vincent, M. A.; Hillier, I. H. The Solvated Fluoride Anion Can Be a Good Nucleophile. *Chem. Commun.* **2005**, *2005* (47), 5902−5903.
(352) Lohman, D. C.; Edwards, D. R.; Wolfenden, R. Catalysis by Desolvation: The Catalytic Prowess of Sam-Dependent Halide-Alkylating Enzymes. *J. Am. Chem. Soc.* **2013**, *135* (39), 14473−14475.
(353) Deng, H.; Ma, L.; Bandaranayaka, N.; Qin, Z.; Mann, G.; Kyeremeh, K.; Yu, Y.; Shepherd, T.; Naismith, J. H.; O'Hagan, D. Identification of Fluorinases from Streptomyces Sp MA37, Norcardia Brasiliensis, and Actinoplanes Sp N902−109 by Genome Mining. *ChemBioChem* **2014**, *15* (3), 364−368.
(354) Wang, Y.; Deng, Z.; Qu, X. Characterization of a Sam-Dependent Fluorinase from a Latent Biosynthetic Pathway for Fluoroacetate and 4-Fluorothreonine Formation in Nocardia Brasiliensis. *F1000Research* **2014**, *61*, 1.
(355) Huang, S.; Ma, L.; Tong, M. H.; Yu, Y.; O'Hagan, D.; Deng, H. Fluoroacetate Biosynthesis from the Marine-Derived Bacterium Streptomyces Xinghaiensis NRRL B-24674. *Org. Biomol. Chem.* **2014**, *12* (27), 4828−4831.
(356) Eustáquio, A. S.; Härle, J.; Noel, J. P.; Moore, B. S. S-Adenosyl-L-Methionine Hydrolase (Adenosine-Forming), a Conserved Bacterial and Archeal Protein Related to SAM-Dependent Halogenases. *ChemBioChem* **2008**, *9* (14), 2215−2219.
(357) Eustáquio, A. S.; McGlinchey, R. P.; Liu, Y.; Hazzard, C.; Beer, L. L.; Florova, G.; Alhamadsheh, M. M.; Lechner, A.; Kale, A. J.; Kobayashi, Y.; et al. Biosynthesis of the Salinosporamide a Polyketide Synthase Substrate Chloroethylmalonyl-Coenzyme a from S-Adenosyl-L-Methionine. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (30), 12295−12300.
(358) Murphy, C. D.; O'Hagan, D.; Schaffrath, C. Identification of a PLP-Dependent Threonine Transaldolase: A Novel Enzyme Involved in 4-Fluorothreonine Biosynthesis in Streptomyces Cattleya. *Angew. Chem., Int. Ed.* **2001**, *40* (23), 4479−4481.
(359) Deng, H.; Cobb, S. L.; McEwan, A. R.; McGlinchey, R. P.; Naismith, J. H.; O'Hagan, D.; Robinson, D. A.; Spencer, J. B. The Fluorinase from Streptomyces Cattleya Is Also a Chlorinase. *Angew. Chem., Int. Ed.* **2006**, *45* (5), 759−762.
(360) Eustaquio, A. S.; Pojer, F.; Noel, J. P.; Moore, B. S. Discovery and Characterization of a Marine Bacterial SAM-Dependent Chlorinase. *Nat. Chem. Biol.* **2008**, *4* (1), 69−74.
(361) Deng, H.; O'Hagan, D. The Fluorinase, the Chlorinase and the DUF-62 Enzymes. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 582−592.
(362) Deng, H.; Botting, C. H.; Hamilton, J. T. G.; Russell, R. J. M.; O'Hagan, D. S-Adenosyl-L-Methionine:Hydroxide Adenosyltransferase: A SAM Enzyme. *Angew. Chem., Int. Ed.* **2008**, *47* (29), 5357−5361.
(363) Deng, H.; McMahon, S. A.; Eustaquio, A. S.; Moore, B. S.; Naismith, J. H.; O'Hagan, D. Mechanistic Insights into Water Activation in SAM Hydroxide Adenosyltransferase (DUF-62). *ChemBioChem* **2009**, *10* (15), 2455−2459.
(364) Wängler, C.; Niedermoser, S.; Chin, J.; Orchowski, K.; Schirrmacher, E.; Jurkschat, K.; Iovkova-Berends, L.; Kostikov, A. P.; Schirrmacher, R.; Wängler, B. One-Step ^{18}F-Labeling of Peptides for Positron Emission Tomography Imaging Using the Sifa Methodology. *Nat. Protoc.* **2012**, *7* (11), 1946−1955.
(365) Ponde, D. E.; Dence, C. S.; Oyama, N.; Kim, J.; Tai, Y.-C.; Laforest, R.; Siegel, B. A.; Welch, M. J. ^{18}F-Fluoroacetate: A Potential Acetate Analog for Prostate Tumor Imaging−in vivo Evaluation of ^{18}F-Fluoroacetate Versus ^{11}C-Acetate. *J. Nucl. Med.* **2007**, *48*, 420−428.
(366) Marik, J.; Ogasawara, A.; Martin-McNulty, B.; Ross, J.; Flores, J. E.; Gill, H. S.; Tinianow, J. N.; Vanderbilt, A. N.; Nishimura, M.; Peale, F.; et al. Pet of Glial Metabolism Using 2-^{18}F-Fluoroacetate. *J. Nucl. Med.* **2009**, *50* (6), 982−990.
(367) Deng, H.; Cobb, S. L.; Gee, A. D.; Lockhart, A.; Martarello, L.; McGlinchey, R. P.; O'Hagan, D.; Onega, M. Fluorinase Mediated C-^{18}F Bond Formation, an Enzymatic Tool for PET Labelling. *Chem. Commun.* **2006**, *6* (6), 652−654.
(368) Martarello, L.; Schaffrath, C.; Deng, H.; Gee, A. D.; Lockhart, A.; O'Hagan, D. The First Enzymatic Method for C−18f Bond Formation: The Synthesis of 5′-[^{18}F]-Fluoro-5′-Deoxyadenosine for Imaging with PET. *J. Labelled Compd. Radiopharm.* **2003**, *46* (13), 1181−1189.
(369) Winkler, M.; Domarkas, J.; Schweiger, L. F.; O'Hagan, D. Fluorinase-Coupled Base Swaps: Synthesis of [^{18}F]-5′-Deoxy-5′-Fluorouridines. *Angew. Chem., Int. Ed.* **2008**, *47* (52), 10141−10143.
(370) Deng, H.; Cobb, S. L.; Gee, A. D.; Lockhart, A.; Martarello, L.; McGlinchey, R. P.; O'Hagan, D.; Onega, M. Fluorinase Mediated C-^{18}F Bond Formation, an Enzymatic Tool for PET Labelling. *Chem. Commun.* **2006**, *2006* (6), 652−654.
(371) Sergeev, M. E.; Morgia, F.; Javed, M. R.; Doi, M.; Keng, P. Y. Polymer-Immobilized Fluorinase: Recyclable Catalyst for Fluorination Reactions. *J. Mol. Catal. B: Enzym.* **2013**, *92*, 51−56.
(372) Sergeev, M. E.; Morgia, F.; Javed, M. R.; Doi, M.; Keng, P. Y. Enzymatic Radiofluorination: Fluorinase Accepts Methylaza-Analog of SAM as Substrate for FDA Synthesis. *J. Mol. Catal. B: Enzym.* **2013**, *97*, 74−79.
(373) Li, X.-G.; Domarkas, J.; O'Hagan, D. Fluorinase Mediated Chemoenzymatic Synthesis of [^{18}F]-Fluoroacetate. *Chem. Commun.* **2010**, *46* (41), 7819−7821.
(374) Onega, M.; Domarkas, J.; Deng, H.; Schweiger, L. F.; Smith, T. A. D.; Welch, A. E.; Plisson, C.; Gee, A. D.; O'Hagan, D. An Enzymatic Route to 5-Deoxy-5-[^{18}F]Fluoro-D-Ribose, a [^{18}F]-Fluorinated Sugar for PET Imaging. *Chem. Commun.* **2010**, *46* (1), 139−141.
(375) Li, X.-G.; Dall'Angelo, S.; Schweiger, L. F.; Zanda, M.; O'Hagan, D. [^{18}F]-5-Fluoro-5-Deoxyribose, an Efficient Peptide Bioconjugation Ligand for Positron Emission Tomography (PET) Imaging. *Chem. Commun.* **2012**, *48* (43), 5247−5249.
(376) Dall'Angelo, S.; Zhang, Q.; Fleming, I. N.; Piras, M.; Schweiger, L. F.; O'Hagan, D.; Zanda, M. Efficient Bioconjugation of 5-Fluoro-5-Deoxy-Ribose (FDR) to RGD Peptides for Positron Emission Tomography (PET) Imaging of [α]V[β]3 Integrin Receptor. *Org. Biomol. Chem.* **2013**, *11* (27), 4551−4558.
(377) Thompson, S.; Onega, M.; Ashworth, S.; Fleming, I. N.; Passchier, J.; O'Hagan, D. A Two-Step Fluorinase Enzyme Mediated ^{18}F Labelling of an RSG Peptide for Positron Emission Tomography. *Chem. Commun.* **2015**, *51* (70), 13542−13545.
(378) Thompson, S.; Zhang, Q.; Onega, M.; McMahon, S.; Fleming, I.; Ashworth, S.; Naismith, J. H.; Passchier, J.; O'Hagan, D. A Localized Tolerance in the Substrate Specificity of the Fluorinase Enzyme Enables "Last-Step" ^{18}F Fluorination of a Rgd Peptide under Ambient Aqueous Conditions. *Angew. Chem., Int. Ed.* **2014**, *53* (34), 8913−8918.
(379) Eustáquio, A. S.; O'Hagan, D.; Moore, B. S. Engineering Fluorometabolite Production: Fluorinase Expression in Salinispora Tropica Yields Fluorosalinosporamide. *J. Nat. Prod.* **2010**, *73* (3), 378−382.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX

(380) Eustáquio, A. S.; Moore, B. S. Mutasynthesis of Fluorosalinosporamide, a Potent and Reversible Inhibitor of the Proteasome. *Angew. Chem., Int. Ed.* **2008**, *47* (21), 3936−3938.
(381) Hong, H.; Spiteller, D.; Spencer, J. B. Incorporation of Fluoroacetate into an Aromatic Polyketide and Its Influence on the Mode of Cyclization. *Angew. Chem., Int. Ed.* **2008**, *47* (32), 6028−6032.
(382) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C. Y. Expanding the Fluorine Chemistry of Living Systems Using Engineered Polyketide Synthase Pathways. *Science* **2013**, *341* (6150), 1089.
(383) B. Harper, D.; T.G. Hamilton, J.; O' Hagan, D. Identification of Threo-ig-fluoro.9,10-Dihydroxystearic Acid:A Novel Cofluorinated Fatty Acid from Dichapetalum Toxicarium Seeds. *Tetrahedron Lett.* **1990**, *31* (52), 7661−7662.
(384) Cahn, J. K. B.; Baumschlager, A.; Brinkmann-Chen, S.; Arnold, F. H. Mutations in Adenine-Binding Pockets Enhance Catalytic Properties of NAD(P)H-Dependent Enzymes. *Protein Eng. Des. Sel.* **2016**, *29*, 31−38.
(385) McDonald, C. A.; Fagan, R. L.; Collard, F.; Monnier, V. M.; Palfey, B. A. Oxygen Reactivity in Flavoenzymes: Context Matters. *J. Am. Chem. Soc.* **2011**, *133* (42), 16809−16811.
(386) Mercer, E. J.; Davies, C. L. Distribution of Chlorosulpholipids in Algae. *Phytochemistry* **1979**, *18* (3), 457−462.
(387) Mercer, E. I.; Davies, C. L. Chlorosulpholipids in Algae. *Phytochemistry* **1975**, *14* (7), 1545−1548.
(388) Mercer, E. I.; Davies, C. L. Chlorosulpholipids of Tribonema Aequale. *Phytochemistry* **1974**, *13* (8), 1607−1610.
(389) Bedke, D. K.; Vanderwal, C. D. Chlorosulfolipids: Structure, Synthesis, and Biological Relevance. *Nat. Prod. Rep.* **2011**, *28* (1), 15−25.
(390) Nilewski, C.; Geisser, R. W.; Carreira, E. M. Total Synthesis of a Chlorosulpholipid Cytotoxin Associated with Seafood Poisoning. *Nature* **2009**, *457* (7229), 573−576.
(391) Landry, M. L.; Hu, D. X.; McKenna, G. M.; Burns, N. Z. Catalytic Enantioselective Dihalogenation and the Selective Synthesis of (−)-Deschloromytilipin a and (−)-Danicalipin A. *J. Am. Chem. Soc.* **2016**, *138* (15), 5150−5158.
(392) Kawahara, T.; Kumaki, Y.; Kamada, T.; Ishii, T.; Okino, T. Absolute Configuration of Chlorosulfolipids from the Chrysophyta Ochromonas Danica. *J. Org. Chem.* **2009**, *74* (16), 6016−6024.
(393) Fischer, S.; Huwyler, N.; Wolfrum, S.; Carreira, E. M. Synthesis and Biological Evaluation of Bromo- and Fluorodanicalipin A. *Angew. Chem., Int. Ed.* **2016**, *55* (7), 2555−2558.
(394) White, A. R.; Duggan, B. M.; Tsai, S.-C.; Vanderwal, C. D. The Alga Ochromonas Danica Produces Bromosulfolipids. *Org. Lett.* **2016**, *18* (5), 1124−1127.
(395) Mooney, C. L.; Haines, T. H. Chlorination and Sulfation Reactions in the Biosynthesis of Chlorosulfolipids in Ochromonas Danica, in Vivo. *Biochemistry* **1973**, *12* (22), 4469−4472.
(396) Mooney, C. L.; Mahoney, E. M.; Pousada, M.; Haines, T. H. Direct Incorporation of Fatty Acids into the Halosulfatides of Ochromonas Danica. *Biochemistry* **1972**, *11* (25), 4839−4844.
(397) Smith, D. R. M.; Uria, A. R. R.; Helfrich, E. J. N.; Milbredt, D.; van-Pee, K.-H.; Piel, J.; Goss, R. J. M. An Unusual Flavin-Dependent Halogenase from the Metagenome of the Marine Sponge Theonella Swinhoei Wa. *ACS Chem. Biol.* **2017**, DOI: 10.1021/acschembio.6b01115.

DOI: 10.1021/acs.chemrev.7b00032
Chem. Rev. XXXX, XXX, XXX−XXX