Engineering new enzymes for synthesis and biosynthesis

A dissertation submitted to The University of Manchester for the degree of Master of Science by Research in the Faculty of Science and Engineering

2018

Varvara Androulaki

School of Chemistry

Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations

CAN	Acetonitrile
ADH	Alcohol dehydrogenase
Chromo-Hal	Tryptophan 5/6 halogenase
CLEAs	Cross linked enzyme aggregates
CV	Column volumes
E. coli	*Escherichia coli*
FAD	Flavin adenine dinucleotide (oxidised)
$FADH_2$	Reduced flavin adenine dinucleotide
Fl-Hal	Flavin dependent halogenase
Hal	Halogenase
Fre	Flavin reductase
HOCl	Hypochlorous acid
HPLC	High-pressure liquid chromatography
Hrs / h	Hours
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	KiloDalton
KMES	2-(N-Morpholino)ethanesulfonic acid potassium salt
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NAD	Nicotinamide adenine dinucleotide (oxidised form)

NADH	Reduced nicotinamide adenine dinucleotide
LB	Lysogeny broth
LBA	Lysogeny broth agar
OD_{600}	Optical density
PrnA	Tryptophan-7-halogenase
PyrH	Tryptophan-5-halogenase
RebH	Tryptophan-7-halogenase
Rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMC	Suzuki-Miyaura cross-coupling
SttH	Tryptophan-5-halogenase
TFA	Trifluoroacetic acid
TLC	Thin-layer chromatography
v/v	Volume/volume concentration
WT	Wild-type

Abstract

Halogenated moieties are massively used as pharmaceuticals, agrochemicals, polymers and many other valuable materials. Biocatalysis has been evolving over the years with interest growing due to the use of naturally occurring enzymes to carry out reactions with chemo- / regio- or stereo- selectivities, which would be otherwise impossible using traditional organic synthesis. That could offer a more efficient ecological and economical advantage during the production of molecular complexity. More selective halogenations can be achieved via biocatalysis. Contrary to traditional chemical reactions which bring about a plethora of side reactions, biocatalysis could bring a more alternative option.

In the current project, Vmax™ cells were used instead of *E. coli,* considering that they would be a more suitable host for halogenases whole cells assays. Despite the fact that detectable activity was achieved in these Vmax™ expressed halogenases, the whole cells assays were unsuccessful. The promising detectable activity though is a novel finding as this halophile has not been reported previously with expression of halogenases.

In addition, the identification of a novel tryptophan halogenase, *S. chromofuscus* found by bioinformatic analysis was screened against a number of substrates and under different conditions, showing halogenation regioselective to the 5 position, showing detectable conversion in most of the substrates. Future work on this enzyme could expand its substrate scope and improve the yield and/or conversion.

Finally, with the aim of understanding the biosynthesis pathway involved in the production of one of the major polyketide mycotoxins faced in agriculture, the expression and characterisation of the Ochratoxin halogenase, identified in the Ochratoxin A mycotoxin biosynthesis pathway was attempted but not succeeded. Although many optimisations and alterations within the gene were experimented, it was considered that the expression of the halogenase in *E. coli* is not suitable as the original system is a fungal host.

Declaration

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

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance Presentation of Theses Policy You are required to submit your thesis electronically Page 11 of 25 with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=2442 0), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Acknowledgments

First of all, I would like to thank my supervisor Prof. Jason Micklefield for the opportunity of being part of his group and also for his guidance and support during this year.

I would like also to thank Dr. Sarah Shepherd for the training she provided me and her constant support. To all members of the Micklefield group, I would like to say a big thank you for their help and guidance throughout this project.

Also, I would like to thank individually Helen Taylor, Elliott Craven, Abigail Herbert and Saadia Irfan for their support, guidance and patience during this year and for proofreading this document. This accomplishment would not been possible without them.

Finally, I would like to thank my family and friends for their encouragement.

1 Introduction

The field of biocatalysis has been evolving in recent years with interest growing due to the use of naturally occurring enzymes to carry out reactions with chemo- / regio- or stereo-selectivities, which would be otherwise impossible using traditional organic synthesis.[1, 2]

Halogenated moieties are extensively used as pharmaceuticals, agrochemicals, polymers and many other valuable materials, such as synthetic intermediates.[1, 2] Many compounds containing halogen atoms have been introduced in clinical trials or are on the market. (Figure 1)

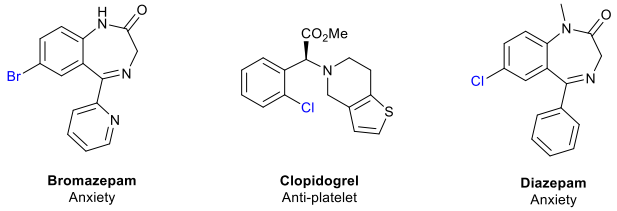


Figure 1: Examples of pharmaceuticals containing organohalogens

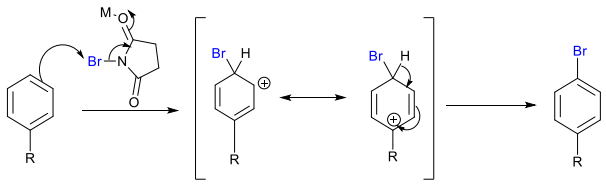
It has been estimated that there are over 4000 organohalogens produced by living organisms, however the quantitative levels in our biosphere are expected to be much larger, with this inventory likely to be incomplete. Most of these compounds are produced by marine life, which can be mainly explained by the rich chloride and bromide content found in aquatic environments, allowing them to incorporate such ions with ease.[3, 4]

Halogenation is a common reaction in organic synthesis, but despite this, the selective introduction of halogens at mechanistically less favoured positions remains challenging.[5]

Current methods for producing organohalide compounds often involves the use of toxic or deleterious reagents including catalysts and organic solvents.[2] Traditional methods of halogenation also tend to lack regiocontrol in the absence of directing groups, rendering some regioisomers inaccessible and resulting in unwanted by-products, including polyhalogenated compounds that require separation and careful disposal due to their toxicity and/or persistence in the environment.[1]

The most common method of halogenation in synthesis involves the generation of a highly electrophilic halogen source, which can be formed via activation of elemental halogen

through the use of a Lewis acid. An alternative method is the use of less toxic halogenating reagents such as N-bromosuccinimide (NBS),[6] the resultant highly conjugated system that forms on breaking of the halide bond acts to increase electrophilicity of the halogen by lowering the relative activation energy. Halogenation of aromatic scaffolds, via electrophilic aromatic substitution, is a highly valuable transformation, as the production of halo-aromatics can allow for further diversification from means of transition metal cross-coupling etc. The breaking of aromaticity to form the carbon-halogen bond is the first and most energetically demanding step of the reaction, resulting in the formation of the carbocation Wheland intermediate, subsequent deprotonation then restores aromaticity producing the halogenated product (Scheme 1).[7]



Scheme 1: Typical mechanism of electrophilic aromatic substitution showing stabilisation of the Wheland intermediate.[7]

As a solution to mitigate both toxicity and lack of selectivity, an interesting method in the production of halogenated compounds is the use of halogenase enzymes for the introduction of halogen atoms, combining an environmentally friendly and more selective process. Many reactions take advantage of organohalogens due to the ability of the C-X bond to undergo oxidative addition to metal centres. The presence of the halogen atom can also provide significant effects on the biological activity towards target molecules, making these compounds highly prevalent in medicinal chemistry.

It has been observed that the loss of the halogen atom reduces the compounds binding affinity for the target and hence it's pharmaceutical activity. This is illustrated in the case of vancomycin (Figure 2), which has been shown to lose its antibiotic activity if non-chlorinated.[2]

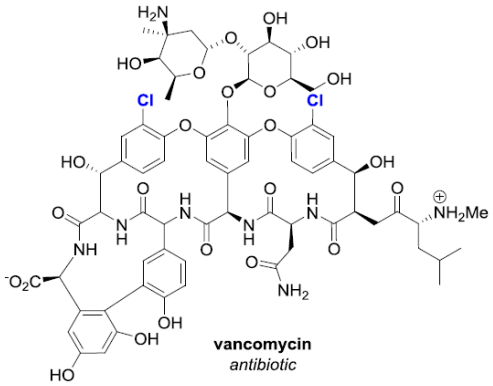


Figure 2: Representation of the antibiotic vancomycin.

The presence of halogen atoms also has a great effect on the bioactivity of compounds, as it can alter the lipophilicity and hydrophobic interactions with targets. The C-X bond is also responsible for additional intermolecular interactions with protein target through the means of halogen bonding, which can vastly improve the specificity and strength of interaction with the protein target. In fact, fluorine is the most extensive halogen used in drugs, due to its similar size to hydrogen and its high electronegativity. These characteristics favour the replacement of hydrogen and the formation of a strong C-F bond without affecting pre-existing interactions with the protein target. A significant number of halogenated natural products isolated from microorganisms have been used as they have strong antimicrobial and antitumor activities.[2]

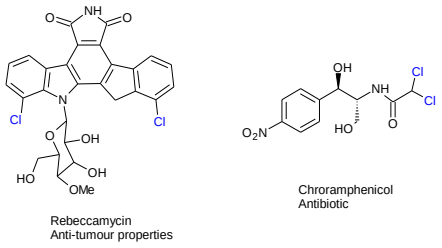


Figure 3: Examples of halogenated natural products.[2]

Most enzymes act through an *in situ* generation of hypohalous acid, derived from halide anions in solution. The halide sources usually come from innocuous halogen salts, which constitute an improvement on synthetic halogenating reagents such as chlorine, bromine or

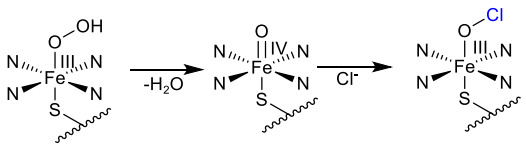
N-halosuccinimides. Halogenating biocatalysts can be divided into four categories according to their respective mechanisms and structural features; these include Heme Haloperoxidase, Vanadium Haloperoxidase, α-ketoglutarate-dependent Halogenases, Flavin dependant halogenases.

2 Halogenases

2.1 Haloperoxidases

2.1.1 Heme dependent halogenases

Many halogenated natural products have been isolated from living organisms leading to the first hypothesis for the existence of halogenating enzymes. The first confirmation came in 1966 with the discovery of chloroperoxidase. Chloroperoxidase is a halogenating enzyme which is thought to be responsible for the chlorination step in the biosynthesis of caldariomycin, isolated from the *Caldariomyces fumago* fungus.[8-10] The enzyme acts through the formation of hypochlorous acid in a heme-containing active site, and is therefore classed as a heme-iron-dependent haloperoxidase.[11] Further studies revealed that this enzyme is responsible not only for chlorinating its natural substrate, 1,3-cyclopentanedione, but also shows chlorinating and even iodinating activity on a wide range of electron rich substrates, such as tyrosine and insulin.[12] The first step of the mechanism is the oxidation of Fe (III) to Fe (IV) which driven by the loss of water, this causes the oxidation to the high valent iron (IV) intermediate, subsequent halide addition reduces the metal centre back down to Fe(III) and substitution with water releases the catalytically active hypohalous acid.. However, lack of specificity is observed due to the freely diffusible nature of the hypohalous acid, thus it is free to halogenate non selectively to any nucleophilic centres on a substrate.[13]

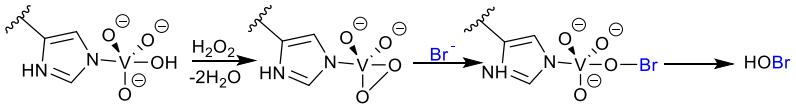


Scheme 2: Catalytic cycle of vanadium-dependent haloperoxidase in hypobromous acid production.

A class of haloperoxidases that were later discovered are the vanadium-dependent haloperoxidases, which act in an analogous way to chloroperoxidase but with an active site containing vanadium instead of heme. From a biocatalytic point of view, the ease production of these enzymes using recombinant expression in *E. coli* and yeast coupled with their higher tolerance of organic solvents make them a useful tool for industrial application.[14]

2.1.2 Vanadium Haloperoxidase

Recently, it has been found that vanadium has a key role in terms of medicinal and enzymatic activity, despite the fact that it was first reported in 1831 when it was isolated from brown and red algae. Vanadium compounds present great enzymatic activity and are highly correlated to medical treatment of diabetes type II as they improve the effect of insulin.[15] The active site contains the catalytic vanadium instead of iron which is found at the base of a 15 to 20 Å channel, with the vanadate ion tethered to the protein by coordination with the imidazole ring of an axial histidine residue and further stabilised by 4 oxygen species in the active site. The mechanism initially begins when hydrogen peroxide binds the vanadate ion forming a peroxovanadium species in the active site. As this species acts as a stronger oxidising agent compared to hydrogen peroxide, it causes the oxidation of the bromide ion to produce the hypobromous acid whilst maintaining a conserved +5 vanadium centre. It has already been shown than the hypobromous acid halogenates several substrates, which confirm substrate specificity.[16]



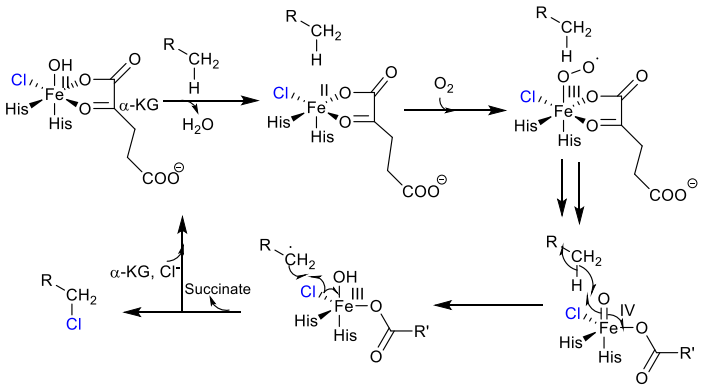
Scheme 3: Catalytic cycle of vanadium haloperoxidase induced halogenation.[15]

In spite of their halogenation activity, the generation of free hypohalous acid still presents issues with regards to substrate specificity and leading to a lack of observed regioselectivity, a major requirement for application as potential biocatalysts.[17]

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

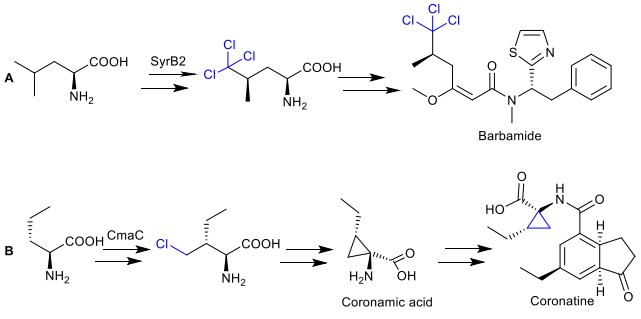
α -ketoglutarate-dependent halogenases are another class of enzyme which are derived from the mononuclear iron enzyme class and have been found to halogenate inactivated aliphatic C-H bonds using α-ketoglutarate (α-KG) as a redox cofactor. The coordination of the iron

centre with active site carboxylate residues compared to that of a porphyrin heme group, coupled with the requirement of substrate tethering via a peptidyl carrier protein (PCP) makes these a distinct class of iron containing halogenase.[2, 18]



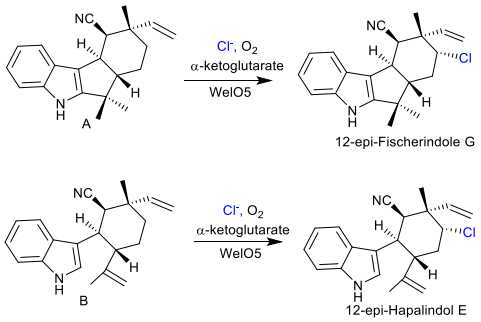
Scheme 4: Proposed mechanism for Fe(II)/α-KG-dependent halogenases.[2]

The first characterised enzyme of this class was SyrB2, which is responsible for the biosynthesis of Syringomycin E by *Pseudomonas syringae*. This enzyme was found to halogenate the PCP-tethered L-threonine specifically at C4 position.[19] Two other Fe (II)/α-KG halogenases have also been found to be responsible for the biosynthesis of other natural products. For instance, the biosynthesis of barbamide was found to occur by two Fe (II)/α-KG halogenases (BarB1 and BarB2). The trichloromethyl group of barbamide is derived from trichlorination of a carrier protein-tethered L-leucine by these two enzymes (Scheme 5A).[20] Furthermore, biosynthesis of the cyclopropyl ring of coronamic acid 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 en route to coronatine (Scheme 5 B).[21]



Scheme 5: Known involvement of Fe (II)/α-ketoglutarate dependent halogenases in the biosynthesis of natural product.

Because of the requirement of substrate tethering for activity and limited substrate scope, largely confined to natural amino acids related to their biosynthetic pathway, these enzymes have greatly limited biocatalytic applications.[22, 23] However, the recent discovery of WelO5, a halogenase responsible for the chlorination of A and B to 12-*epi*-fischerindole U and 12-*epi*-hapalindole C (Scheme 6),[25] demonstrates a promising example of an α-KG dependent halogenase that avoids these issues and provides regio- and stereo-selectivity. WelO5 lacks the requirement for a carrier protein for activity and has a substrate turnover much higher than the other α-KG dependent halogenases already discovered,[24] and is therefore of interest for biocatalytic purposes.



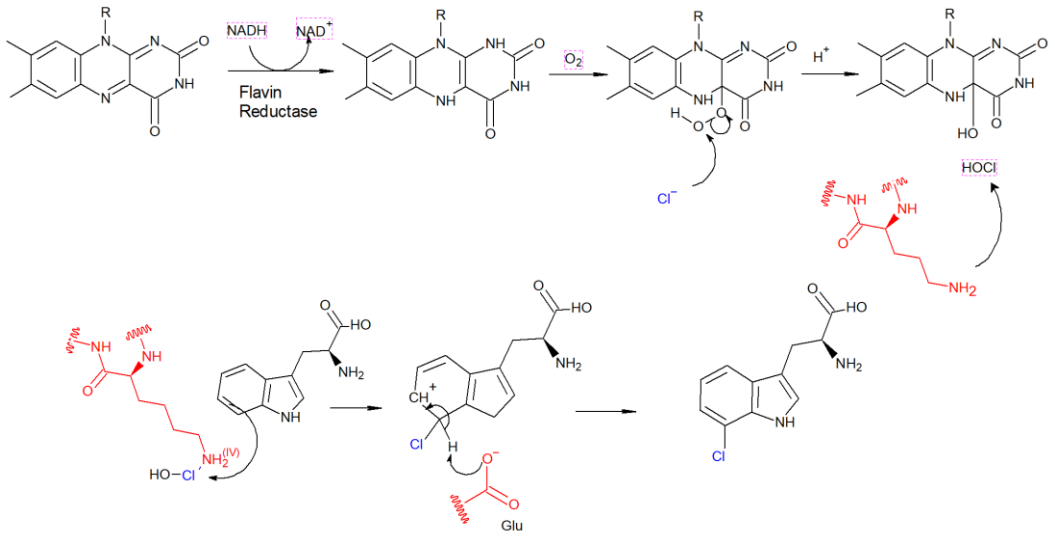
Scheme 6: WelO5 catalysed halogenation of indolinone natural products.

2.3 Flavin-Dependent Halogenases

The flavin-dependent halogenase enzymes are considered one of the most important and interesting classes of halogenases in terms of their biocatalytic applications. These enzymes are classified into the superfamily of flavin-dependent monooxygenases, which use molecular oxygen to form C4α-hydroperoxy flavin from reduced-flavin, the active molecule responsible for the enzyme activity.[27, 28] Contrary to the previous class of halogenases, Flavin dependent halogenases (FlHals) do not require a metal centre for their catalytic activity.[29] The classification of the FlHals as "two distinct component" monooxygenases is due to the fact that $FADH_2$ is usually derived from a second flavin reductase enzyme which facilitates the reduction of FAD to produce $FADH_2$ which is required for activation of molecular oxygen.[28] The absence of the flavin reductase enzyme can be related to the loss of enzyme efficiency due to the persistence of the more stable oxidised FAD and the auto-oxidation of $FADH_2$ in

solution.[30] As a consequence, it is highly recommended these two enzymes are used in conjunction, the way they appear in living organisms.

Despite the fact that it was initially considered that FlHals operate in a way identical to the other Flavin-dependent monooxygenases, where the substrate directly attacks the C4α-hydroperoxy Flavin or a chlorinated Flavin-moiety,[31, 32] crystallography revealed a different mechanistic action. The structures of tryptophan-7-halogenases (PrnA and RebH), and tryptophan-5-halogenase (PyrH) gave insight into the regiocontrol and mechanism of these enzymes, and consequently demonstrated distinct differences from that of other monooxygenase enzymes. More specifically, the crystallography of PrnA illustrated that $FADH_2$ and substrate were bound in spatially unconnected sites, separated by a tunnel approximately 10 Å in length. Due to its size, the direct interaction between the cofactor intermediate C4α-hydroperoxy flavin and the substrate is unlikely.[33, 34] Further studies have shown that the reaction of the $FADH_2$ with the molecular oxygen, generating of the C4α-hydroxy flavin, took place prior to halogenation, proposing that the chloride attacks the C4α-hydroperoxy flavin to form hypochlorous acid.[35] (Scheme 7)



Scheme 7: Proposed mechanism for regioselective FAD dependent halogenation

Experiments have revealed in the absence of the substrate, an enzyme-chloride adduct is formed, with an active site lysine abstracting the halide from the hypohalous acid to form a chloroamine intermediate. The role of this lysine residue is key in the determination of C-H selectivity of the substrate and has been demonstrated in recent studies that this residue is pivotal in maintaining halogenation activity.[36] It is hypothesised that the enzyme-chloride

adduct is bound covalently to the amino group or that the lysine interacts with the hypochlorous acid via hydrogen bonding. The electrophilicity of the chlorine atom is enhanced by surrounding glutamate residues and these also act to facilitate the base mediated deprotonation of the halogenated Wheland intermediate in the electrophilic aromatic substitution step.[37]

2.3.1 Tryptophan FAD dependent halogenases

Flavin-dependent halogenases can be divided into further subcategories based on the substrate structure they halogenate such as pyrrole halogenases like PltA ans PrnC, phenolic halogenases and tryptophan halogenases. Tryptophan halogenases are the most studied among these halogenases because they appear to accept multiple substrates other than tryptophan.[2, 38, 39]

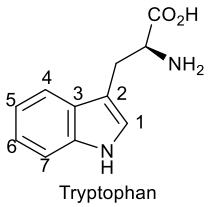


Figure 4: Structure of tryptophan

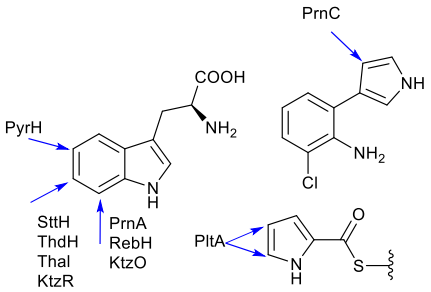
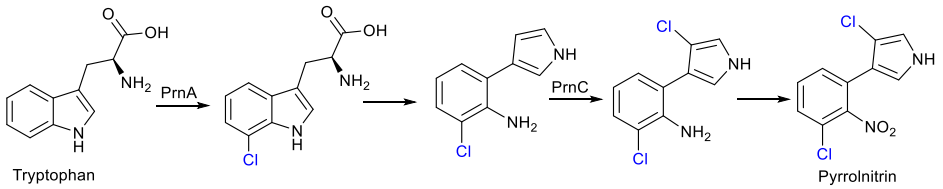


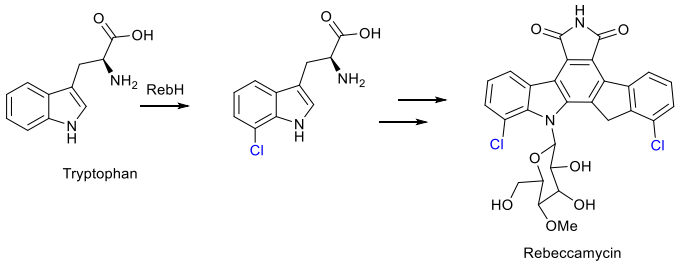
Figure 5: Natural substrates and products of flavin-dependent halogenases. Arrows illustrate the halogenation position.

PrnA was the first tryptophan halogenase, identified and is responsible for the biosynthesis of the antifungal pyrrolnitrin produced by *Pseudomonas fluorescens.*[40] From the compound structure it has been indicated that PrnA chlorinates at the C7 position of the tryptophan molecule which occurs before the formation of the pyrrolitic intermediate and then a second halogenase (PrnC) is responsible for chlorination of this intermediate. (Scheme 8)



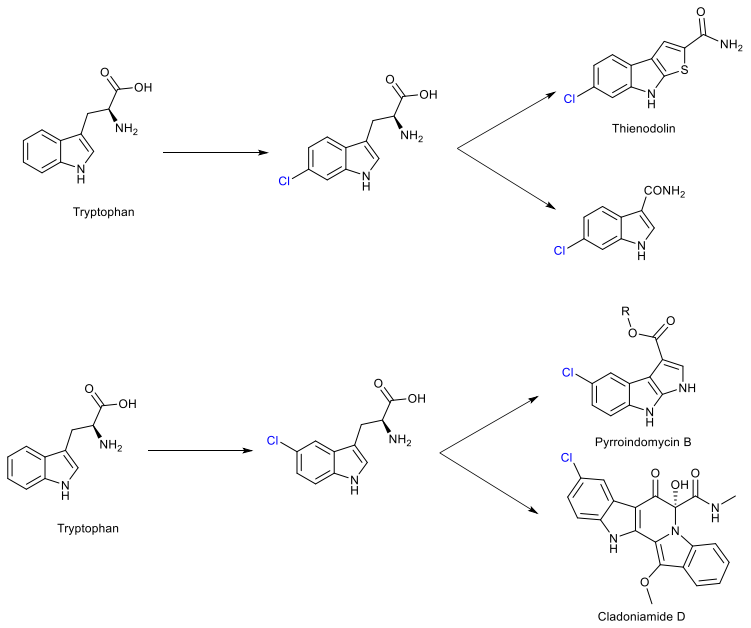
Scheme 8: Biosynthesis of Pyrrolnitrin by tryptophan halogenases (PrnA)

RebH, which is also responsible for the chlorination at the C7 position of tryptophan in the biosynthesis of the antitumor compound rebeccamycin, was later identified. RebH, which was isolated from the terrestrial bacteria *Lechevalieria aerocolonigenes* shares over 50% sequence identity to PrnA.[31]



Scheme 9: Biosynthesis of Rebeccamycin by tryptophan halogenases (RebH)

PyrH is the only fully characterised 5-tryptophan halogenases and is responsible for the biosynthesis of pyrroindomycin in *Streptomyces rugosporus* and Cladoniamide D in *Streptomyces uncialis*.[36, 41] Further examples include 6-tryptophan halogenases, such as ThdH and Thal, with more recent studies showing full identification and characterisation of KtzQ and KtzR which are also responsible for the chlorination steps of tryptophan at C7 and C6 positions, respectively. Moreover, SttH, a tryptophan-6-halogenase, has been shown to be involved in the biosynthesis of lipstatin in the bacterium *Streptomyces toxytricini.*[42]



Scheme 10: Biosynthesis of natural products using tryptophan 5- and 6- halogenases

Flavin-dependent phenolic halogenases have recently been found to act on free standing substrates, enzymes such as the halogenase RadH isolated from the terrestrial fungi *Chaetomium chiversi* is amongst this class. Among other fungal enzymes, RadH, which catalyses the biohalogenation in the biosynthesis of radicicol, halogenates ortho to the hydroxyl group. The ochratoxin halogenase is another phenolic halogenase and is a key enzyme within the biosynthetic pathway of the production of ochratoxin A (OTA) and has been shown to halogenate para to the hydroxyl group.[43]

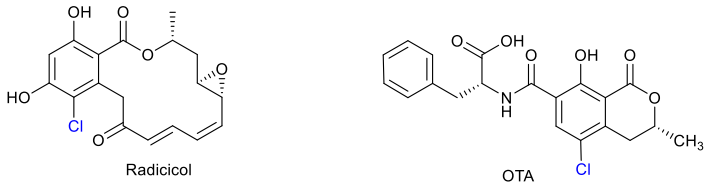


Figure 6: Natural substrates and products of a selection of flavin-dependent halogenases.

2.4 Flavin-Dependent Halogenases as Biocatalysts

Among all halogenases discovered, flavin-dependent halogenases are considered highly suitable for synthetic applications as they halogenate with regioselectivity on free standing moieties, as discussed above. More specifically, tryptophan halogenases present high regioselectivity as they halogenate the indole moiety of tryptophan at C5, C6 or C7 positions,

but also accept various other aromatic substrates with distinct regioselectivities.[44] Tryptophan halogenases are a key focus in biocatalytic studies, showing a vast number of compatible substrates, such as tryptophan derivatives, benzamides and benzoic acids.[43, 45] RebH seems to show halogenation activity predominately at the C7 position of several non-native indole derivatives and also accepts aromatic amine and hydroxyl functionalisation, directing the halogenation ortho.[43] RebH, like most of the tryptophan halogenases, is also highly dependent on the chosen substrate, as the regioselectivity can change to the more activated position, depending on the functionalities in the molecule, and dihalogenation can occur as a result of halogenation being driven by the electron donating groups. Dihalogenation can be minimised by decreasing the reaction time so that only a single mono-halogenated regioisomer is allowed to be formed.[46]

PrnA, as discussed previously, seems to act in a different way to RebH as it halogenates certain non-natural indolic substrates at the C2 position, which is highly electronically favoured.[47] Because of the limited number of potential substrates, multiple efforts have been made for further studies focused on PyrH and SttH, which are a tryptophan-5-halogenase and tryptophan-6-halogenase, respectively. Many studies have shown that both enzymes primarily halogenate anthranilamide para to the position of the amine group, with a regioisomer mixture of both ortho and para products formed by PrnA.[43] For the enzymes to be used as biocatalysts, the regioselectivity is essential. Tryptophan halogenases, and especially PrnA, have been modified to improve and alter their regioselectivity via mutagenesis, as is illustrated by Lang *et al*. Mutagenesis of 103 in PrnA from a lysine residue to alanine enables the halogenation at the C7 and C5 positions with 2:1 ratio, contrary to the wild type of PrnA, where only the halogenated C7 product is formed.[48] Further studies have shown that mutagenesis in the active site of PrnA provide an orientation for the binding of the substrate anthranilic acid to give specific site directed halogenation. Random, rational and directed mutagenesis approaches have already been used to improve and alter the regioselectivity. RebH was the first enzyme to be mutated so as to be more thermostable. More detailed screening showed that two active mutants of RebH were able to halogenate an indolic substrate at C5 and C6 positions with high specificity and reactivity. RebH was shown to halogenate much larger substrates when it was randomly mutated.[49]

In spite of the positive results and the potential applications of the Fl-Hals as biocatalysts, there are still some improvements to be made with regards to poor stability and low turnover numbers.[50] The identification of more stable enzymes from thermophilic organisms can be a way to improve biocatalytic properties, among other advantages, increased stability and solvent tolerance are some major characteristics of these thermostable enzymes.[51] From previous experiments, it has been observed that the thermophilic tryptophan-6-halogenase from *Streptomyces violaceus niger* has a 10 °C higher melting point than the wild type.[52] This halogenase was also tested with thermophilic flavin reductase that permitted an increased temperature (45 °C) of the biocatalytic reactions with various non-natural substrates. The regioselectivity of the halogenase is comparable to that of SttH, and from crystallographic experiments the thermophilic halogenase presents 40% differences in polar residues lying at their surface, with the most similarities to SttH being found in the polar residues at the surface, which is related to increased surface charge. Surface charge is considered to improve stability, due to hydrogen bonding interactions with water and reduction of protein aggregation.[53]

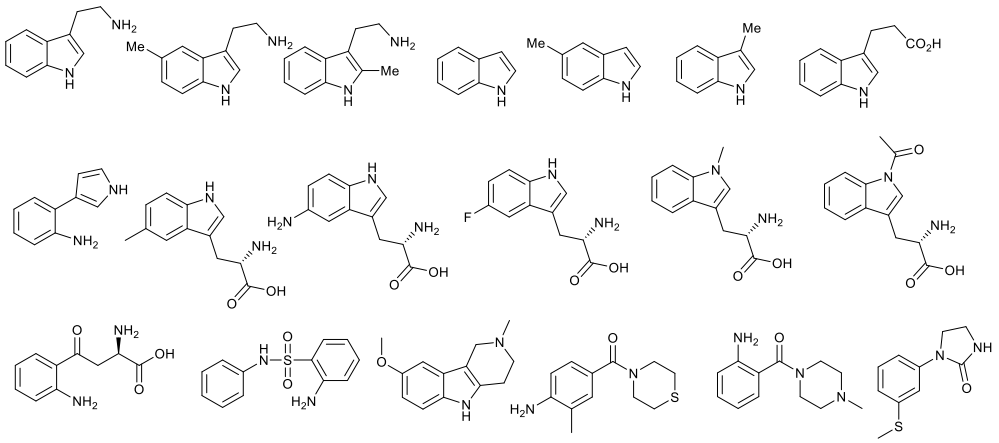
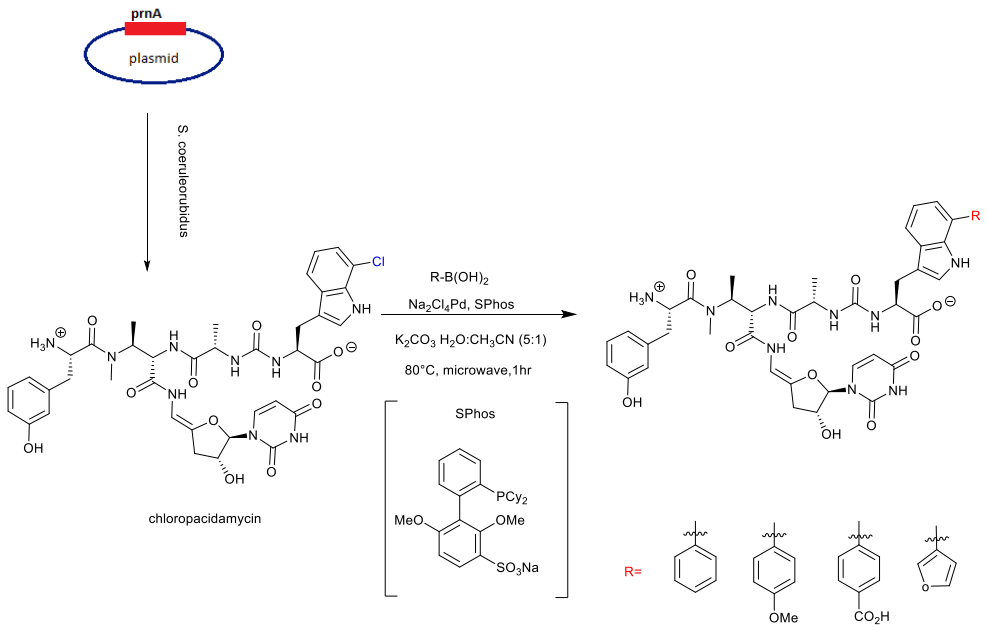


Figure 7: Non-natural substrates of Tryptophan Fl-Hals.[2, 54]

2.5 Suzuki Cross-coupling using halogenase enzymes

Suzuki-Miyaura coupling (SMC) is the most studied cross-coupling, which was first introduced in 1979 and is related to the C-C bond formation through the use of an organohalide and a boronic acid or ester. The versatile formation of complex molecular structures from two independent adducts, utilising a large range of boronic acids, are important advantages of

Suzuki coupling.[55, 56] The widespread use of palladium cross-coupling within organic synthesis has led to the combination of cross-coupling reactions with regioselective halogenation, as to achieve new bond formation at specific positions. The C-H activation controls the functionalisation of the substrate to the position with the highest C-H acidity or coordination, leaving other positions inaccessible.[51] A notable example is the production of chloropacidamycin from *S. coeruleorubidus* with the use of PrnA to introduce a Cl at the C7 position, leading to the production of 7-aryl pacidamycin moieties using Suzuki cross-coupling.[57]



Scheme 11: Production of chloropacidamycin from *S. coeruleorubidus* using Suzuki cross-coupling.[57]

In this way, the highly biologically active compounds and synthetic intemediates required in industry for halogenation reactions can be afforded by the enzymes as explained above. The Fl-Hals among all the enzymes already discussed are considered the most promising in terms of regioselectivity. They have shown precise regio-controlled halogenation reactions and are essential in medicinal chemistry. Current work within the Micklefield group is been carried out to optimise the conditions of halogenation reactions, as well as extending knowledge by testing a various number of tryptophan derivatives and non-native substrates.

2.6 Ochratoxin Halogenase

Ochratoxin A (OTA) is a major polyketide mycotoxin and is mainly derived by several *Aspergillus* and *Penicillium* species. It was first isolated in 1965 from the culture broth of *Aspergillus ochraceus*.[58, 59] Nephrotoxic, teratogenic, immunotoxic, neurotoxic and hepatotoxic effects have been reported with this compound.[60] But also OTA has been reported as a human carcinogen.[61] Nephrotoxicity studies have shown that with doses of 1 and 4 ppm, the kidney seems to lose its colour and necrosis is observed in a period of 3–4 months.[62]

The biosynthetic genes of fungal mycotoxins, including the OTA, are located within clusters and show high relation in identification among the genes that are present in the same cluster. OTA is a secondary metabolite which has been found in a large number of food products such as grapes, coffee, nuts cereals, spices and many others. OTA found in grapes proves to be not only a major health issue but also an economic issue in all southern European countries, and in those with similar climate.[63] OTA has a high risk for food safety as it accumulates in human and animal organs, which has led to the implementation of legal regulations for some agricultural products. More recent studies have shown that the most significant species within the genus *Aspergillus* section *Circumdat*i for OTA production are *A*. *steynii* and *A*. *westerdijkiae* and in *Aspergillus* section *Nigri A. carbonarius, A. niger and A. welwitschiae* respectively.

In the OTA biosynthetic cluster, the most significant genes are a polyketide synthase (PKS) and a non-ribosomal peptide synthase (NRPS).[64] The production of many polyketide metabolites with great chemical diversity is highly related to multiple PKS genes which are usually present in fungal genomes. A single gene organised into eight types of functional domains is responsible for encoding PKS enzymes which have multifunctional roles.[65] Surprisingly, PKS and NRPS genes in *Penicillium* are not related to those in *Aspergillus* genus, demonstrating that they are specific of the genus. In addition, other genes which participate in OTA biosynthesis encode cytochrome P450 monooxygenases (P450), AcOTA-P450 and a basic leucine zipper (bZIP) transcription factor, AcOTAbZIP. The P450 role in the OTA biosynthetic pathway in *A. ochraceus* and *A. steynii* species has been already reported, where it seems to catalyse the oxidase step responsible to form the precursor metabolite OTβ. The presence of bZIP transcription factor is also essential for the OTA gene expression. It is

common among fungi that for the biosynthetic enzymes, gene expression with one or more transcription factors are usually needed. BZIP are proteins responsible for several metabolic processes.[66-68] In fact, the AcOTApks gene which is responsible for the PKS seems to catalyse the isocoumaric group at the very first steps of the OTA biosynthetic pathway when starting from acetate and malonate, which provide the pentaketide form of the OTA. The AcOTAnrps gene is also responsible for the NRPS which links the phenylalanine to dihydroisocoumarin ring, a step which is assumed to occur before the chlorination. It has been observed that the previous step in the OTA biosynthesis produces the ochratoxin B (OTB) analogue of OTA, but without the chlorine atom at C5 position.[60, 63, 69] The introduction of the halogen in the OTA is not related to the high toxicity of the molecule as *in vitro* assays have shown even similar toxicities in same cases between OTA and OTB where there is hydrogen instead of chlorine.[60]

The addition of the halogen atom (chlorine) to ochratoxin B by a chloroperoxidase or halogenase enzyme does not occur, as studies on the OTA pathway have shown. A new halogenase gene has been described in the OTA biosynthetic cluster. After identification and sequencing of the genomic DNA of *A. carbonarius,* the region with the PKS and NRPS genes allowed the identification of a gene responsible for the expression of a flavin-dependent halogenase. The characterisation of this gene has been achieved, *AcOTAhal*. The P450 and bZIP genes followed after the PKS, NRPS and Hal gene, in the biosynthetic cluster.[60] (Figure 8)

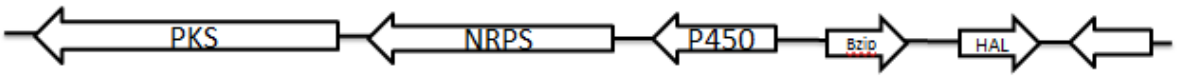


Figure 8: Structure of the OTA cluster in *Aspergillus carbonarius*.[60]

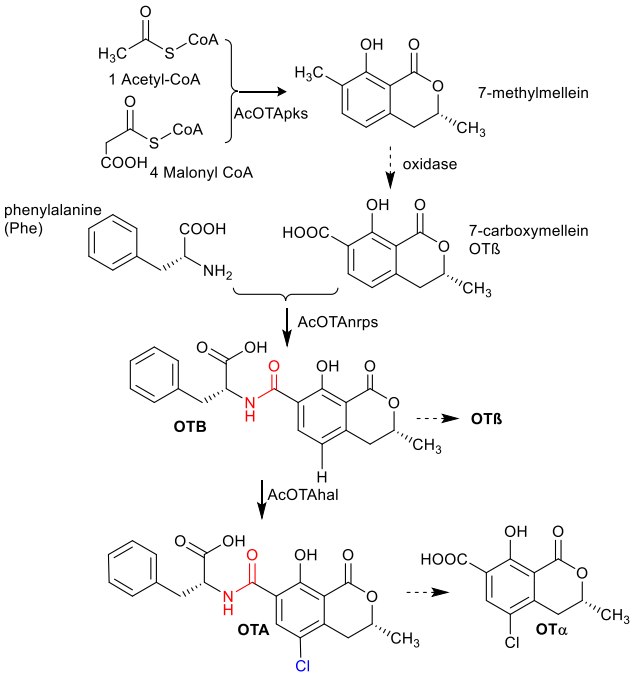
Multiple studies have revealed many genome sequences of species containing the biosynthetic pathway of OTA, which can be found within databases, and through these sequences, the study of OTA could be expanded. Undoubtedly, this information could contribute to minimisation of the risk as a result of the presence of OTA in agricultural products.

From recent studies it has been proved that in both *A. westerdijkiae* and *A. steynii* species the HAL was the most expressed gene, followed by PKS, P450 and NRPS encoding genes.

Furthermore, the expression of bZIP was also present in these species. All the five genes have been expressed by the *A. carbonarius* and *A. niger* species. Additionally, the P450 encoding gene presents the highest expression level in *A. carbonarius,* while the PKS encoding gene was the most expressed in *A. niger*. Finally, in *P. nordicum,* the most expressed gene of the cluster was HAL with much higher levels in comparison to the rest of the genes.

Nevertheless, a full understanding of the fungi mechanisms for production of mycotoxins is essential and could be accomplished by a deep study of the expression of the genes which are take place in the OTA biosynthesis. It is significant that all the different species responsible for the OTA biosynthesis seem to be related with the high expression of a HAL, which suggests that the chlorination in production of the OTA is made by a HAL enzyme[60] and not by chloroperoxidase as was reported in previous studies.[70] Ochratoxin α is the chlorinated isocoumaric derivative of OTA which is not connected with phenylalanine.[63]

The halogenase enzyme appears to be analogous to a RadH flavin-dependent halogenase of *A. niger*. The *A. Niger* halogenase gene has been also identified in the OTA cluster.[71] It has been proved that the genomic regions of the OTA biosynthetic clusters in both *A. carbonarius* and *A. niger* are organised in the same way in terms of order and direction of transcription[72], which is related to the homology between the halogenases in *A. carbonarius* and *A. niger* much more than homology between PKSs and NRPSs, which are also key proteins in the OTA biosynthesis. An additional reason could be the functional domains which seem to be presented as very conserved.[73, 60]



Scheme 12: Representation of the OTA biosynthetic pathways, with all the enzymes which participate. The dashed arrows depict hypothesised steps.[19]

3 Aims of the project

There were two main aims of the project outlined in this literature. Firstly, this project focusses on extending the biocatalytic scope of a number of flavin dependent halogenase enzymes. This will aid the potential production of new halogenated pharmaceuticals in the future.

Initially, the project involved using Vmax™ cells to test a range of halogenases in whole cell reactions without the use of purified enzymes. Whole cell reactions have many benefits, avoiding tedious purification methods and the possibility of the enzymes becoming inactive. Whole cell reactions have been attempted in *E. coli* previously, with little success. Vmax™ cells, which originate from marine bacteria, *Vibrio natriegens,* are very fast growing, provide high yield and aid in the expression of insoluble or inactive proteins. Vmax™ is used in this case for purification of the halogenases as it was thought that halogenases might be expressed in higher yield compared to in *E. coli* strains, due to common issues with insolubility and stability. The implementation of Vmax™ cells instead of *E. coli* is also rationalised due to the fact that the marine organism is a halophile, meaning that its cell membrane should be more permeable to and it should be more tolerable to high salt/halide rich environments. It is speculated that perhaps this organism may have reduced expression of enzymes catalysing the breakdown of tryptophan. This means that halogenase enzymes produced in this organism, compared to *E. coli*, may halogenate their natural substrate at higher efficiency in whole cell assays.

In addition, novel work is performed with a chromo halogenase (chromo-Hal) which was identified via bioinformatic screening as either a putative 5 or 6 tryptophan halogenase derived from *Streptomyces chromofuscus*. This novel halogenase is tested with its natural substrate, tryptophan, and the optimal conditions for its activity will be identified. NMR analysis is performed on the products to attempt to identify the position of halogenation. If this enzyme could be confirmed as a 5 tryptophan halogenase this would be a massive achievement in this field, as there is only one other fully characterised halogenase regioselective to the 5 position, PyrH.[36] Chromo-Hal is also tested against various substrates to expand its catalytic scope.

Following on with the identification and characterization of novel halogenases, the second part of the project is focused on the ochratoxin halogenase (Ochra-Hl). The main aim is the expression of the Ochra-Hl and the Ochra-P450 enzymes, which are both known to participate in the biosynthesis of OTA – however have not been characterized yet in the literature. The biosynthesis pathway of Ochratoxin has been speculated; however, through the expression of these enzymes specific substrates could then be tested to elucidate their biosynthetic mechanisms.

Therefore, through this project, both the repertoire of halogenases available and the understanding of their substrate specificity will be expanded to aid the production of pharmaceuticals in the future. Novel halogenases will be characterised and their regioselectivity confirmed with various potential substrates.

4 Results and discussion

4.1 Halogenase enzymes

4.1.1 Gateway cloning in Vmax™ cells

Firstly, the possibility of expressing halogenases in Vmax™ cells for whole cell reactions was explored to expand the substrate scope of the enzymes. Due to the partial natural resistance the Vmax™ strain has towards kanamycin, it was first necessary to ligate the halogenase encoding constructs into vectors which express resistance to non-kanamycin antibiotics. In doing so this would mitigate issues that could occur during the transformation and selection process, as it will avoid the production of false positives. It was therefore necessary to change the vector containing the halogenase genes (PrnA, PyrH, RebH and SttH) from pET-28a to a new vector with resistance to a different antibiotic. Through gateway cloning, the halogenase genes were introduced into pDEST17.

PCR was performed for amplification of the genes prior to gateway cloning (Figure 9). Due to difficulties during PCR of PrnA and SttH, a gradient PCR was attempted over 50 °C - 70 °C and it showed that a lower annealing temperature (50 °C) worked for PrnA. All temperatures were successful for the SttH gene.

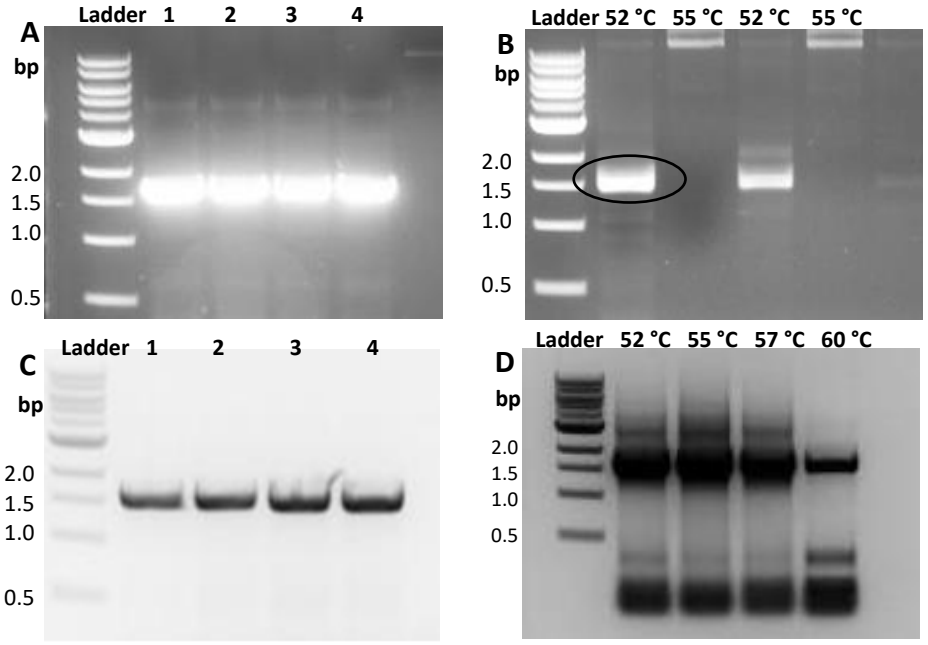


Figure 9: Images of DNA gel showing PCR of various halogneases, where A: RebH (replicates 1-4), B: PrnA (annealing temperatures indicated), C: PyrH (replicates 1-4) and D: SttH (annealing temperatures indicated). NEB 1 kb ladder used.

Electroporation was performed to introduce the plasmids to Vmax™ cells for the whole cell assays. PrnA did not provide any colonies, so only the other three enzymes were carried forward for whole cell assays. Expression and purification of RebH and SttH was carried out successfully, in order to use to confirm enzyme activity (Figure 10).

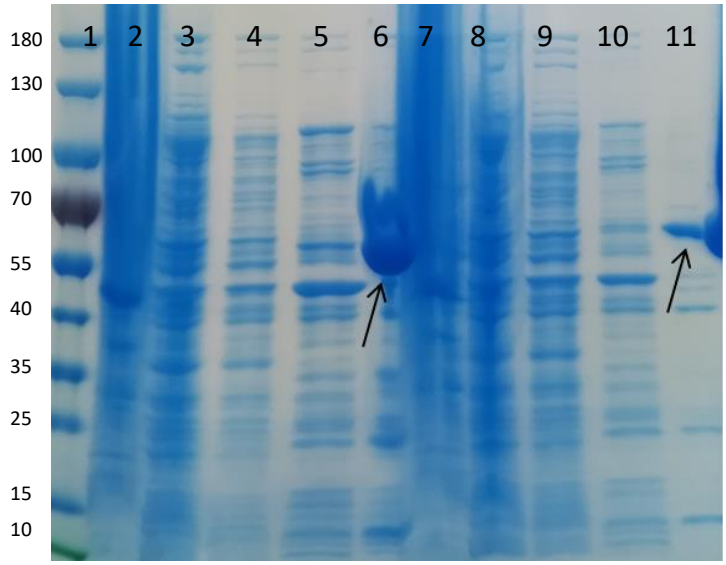


Figure 10: Image of SDS-PAGE showing the expression of SttH (58 kDa) and RebH (64 kDa), where from left to right 1: Protein ladder, 2: insoluble fraction, 3: flow through, 4: Wash 1, 5: Wash 2, 6: elution SttH, 7: insoluble fraction, 8: flow through, 9: Wash 1, 10: Wash 2, 11: elution RebH. The expected bands for both enzymes are indicated with arrows. Thermofisher PageRuler Prestained Protein Ladder is used.

Table 1: Concentration of RebH and SttH produced using Vmax™ cells in 1 L culture.

Enzymes	Concentration (mg/mL)
RebH	10
SttH	26

4.1.2 Vmax™ Assays

It was essential to confirm the activity of halogenases against Trp, their natural substrate, using MgCl as a chloride source. Following confirmation of the activity of purified enzymes from these strains, whole cell assays using Vmax™ cells could be explored.

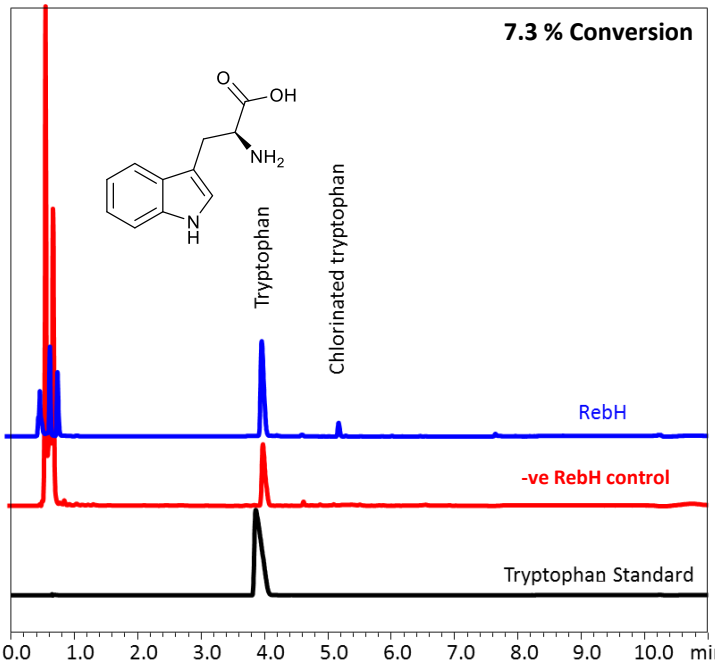


Figure 11: HPLC chromatogram showing the purified RebH halogenase assay with Trp.

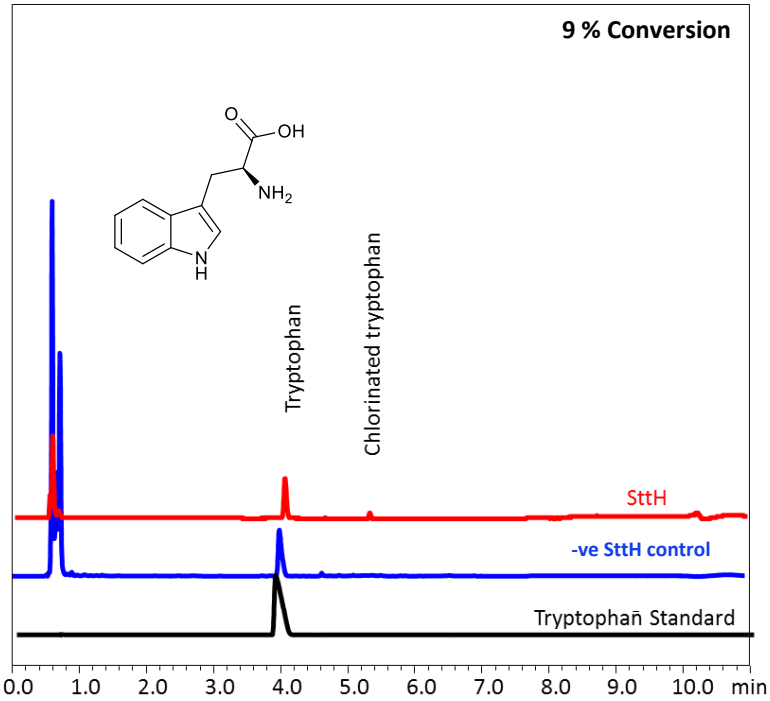


Figure 12: HPLC chromatogram showing the purified SttH halogenase assay with Trp.

Purified halogenases showed some activity against trp, although conversions were low. This low conversion could be due to misfolding of the protein during expression in Vmax cells or other issues with expressing these enzymes in this new environment. The expression of these enzymes in Vmax™ cells could be optimised in the future. However, the main interest in

expressing these enzymes in Vmax focusses on achieving conversion in whole cell assays. Therefore, despite the low concentration of the expressed enzymes and its activity against Trp, it was decided to test these halogenases enzymes in whole cell reactions.

4.1.3 Vmax™ Whole cell and lysate assays

Attention was focused to testing halogenases in Vmax™ cells in a whole cell reaction without the use of purified enzyme. SttH, RebH and PyrH were all tested (Figure 13, Figure 14, Figure 15).

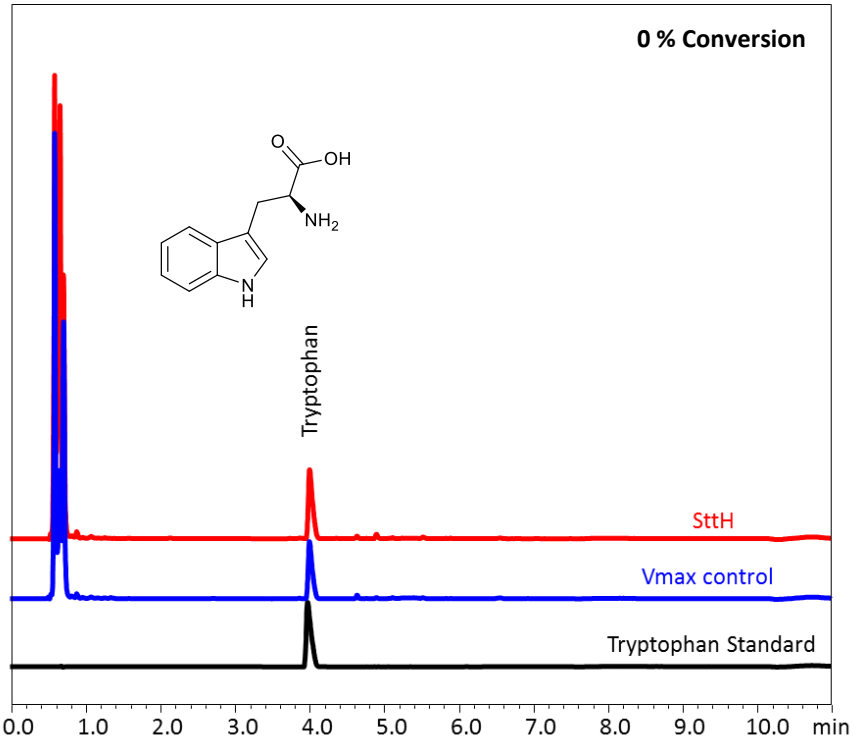


Figure 13: HPLC chromatogram showing the SttH halogenase whole cell reaction assay. Vmax™ control = Vmax™ whole cells not containing SttH.

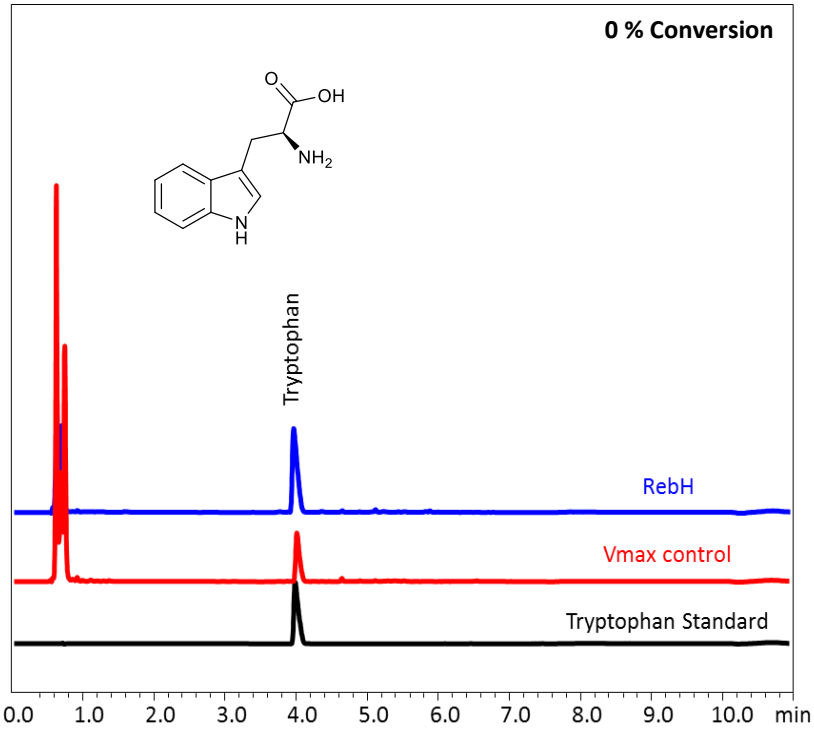


Figure 14: HPLC chromatogram showing the RebH halogenase whole cell reaction assay. Vmax™ control = Vmax™ whole cells not containing RebH.

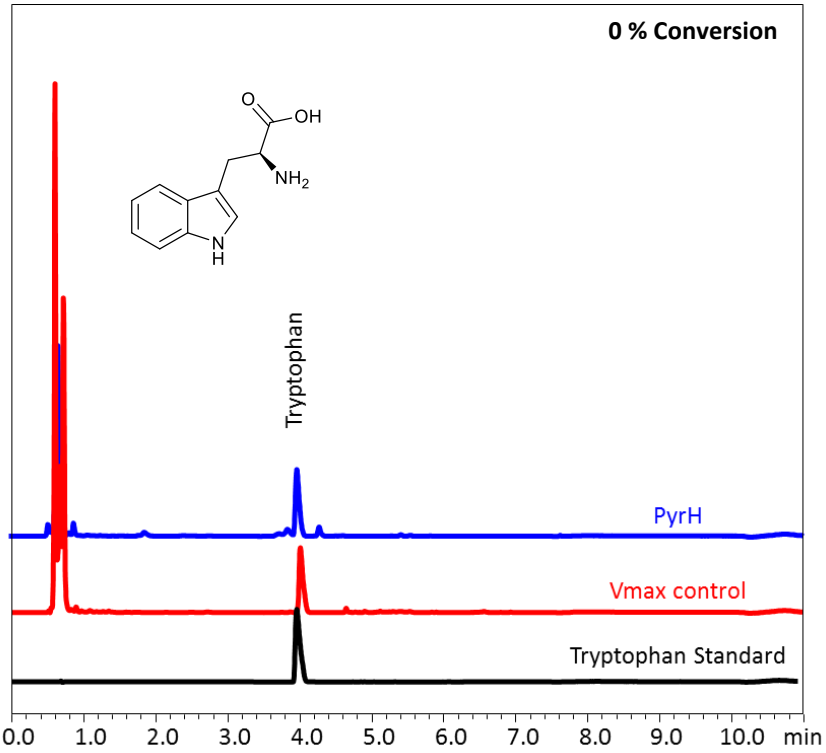


Figure 15: HPLC chromatogram showing the PyrH halogenase whole cells reaction assay. Vmax™ control = Vmax™ whole cells not containing PyrH.

Unfortunately, the whole cell assays in Vmax™ did not show any conversion. Therefore, it was decided to attempt the reaction with the cell lysate with the use of cofactors. Cofactors were added separately, as these are usually expressed by the cell. SttH was chosen to test in these lysate assays, as it was most highly expressed. Lysate was tried in the assays, in case there was an issue with permeability of the membrane to the substrate.

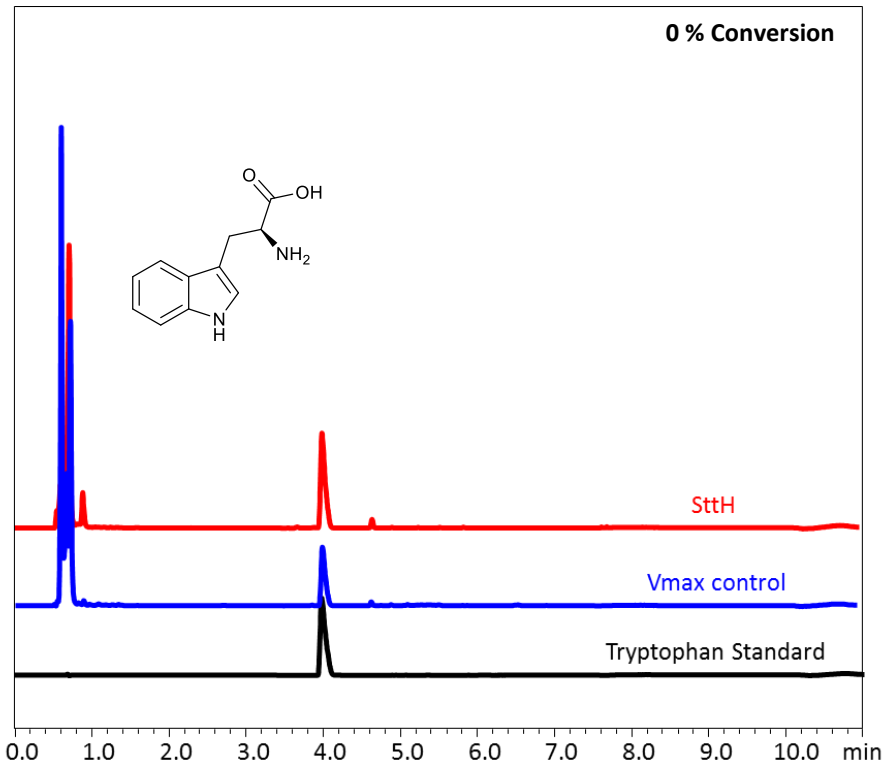


Figure 16: HPLC chromatogram showing the SttH halogenase Lysate reaction assay. Vmax™ control = Lysed Vmax™ cells not containing SttH.

The HPLC chromatograms from the whole cells reaction and lysate reaction did not present any conversion of the tryptophan. It was thought that Vmax™ cells could provide a favourable environment for the whole cell reaction, contrary to the results obtained. If the whole cell reaction had been worked, that would be a great tool for further reactions, as it does not require the use of purified enzymes or any cofactors. In the future, expression of halogenases in Vmax™ could be optimised to improve yield, perhaps then whole cell assays may become possible. Due to the low yield of purified enzymes, this could indicate that perhaps the level of active halogenase is not high enough for detectable activity in whole cell assays.

4.2 Chromo Halogenase

4.2.1 Expression and purification of Chromo-Hal and Fre

The gene encoding for Chromo-Hal had previously been cloned into pET-28a within the Micklefield group, *E. coli* (DE3) Arctic express cells were then chemically transformed with the plasmid and expressed under reported conditions. Purification was performed via Nickel NTA chromatography and analysis via SDS-PAGE shows a characteristic band at 57 kDa. As the $FADH_2$ is unstable and that can lead to degradation via oxidation, there is a requirement of a coenzyme for the reduction of FAD to $FADH_2$ *in situ*. For this enzymatic reduction, *E. coli* produced Flavin Reductase (Fre) has been shown to be optimal for reduction when coupled with use of the halogenase enzymes. This enzyme utilises nicotinamide adenine dinucleotide (NADH) as hydride source for the in vitro formation of $FADH_2$. Both these enzymes were successfully expressed and purified.

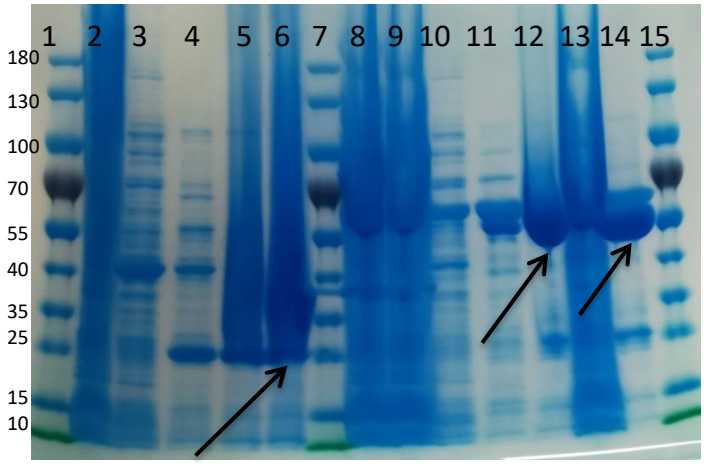


Figure 17: Image of SDS-PAGE showing the expression of Chromo-Hal (57 kDa) and Fre (26.6 kDa) using Arctic Express cells where from left to right; 1, 7, 15: Protein ladder, 2: Fre supernatant, 3: Fre Wash 1, 4: Fre Wash 2, 5: Fre Wash 3, 6: Fre elution, 8: Chromo-Hal Colony1 supernatant, 9: Chromo-Hal Colony1 flow through, 10: Chromo-Hal Colony1 Wash 1, 11: Chromo-Hal Colony1 Wash 2, 12: Chromo-Hal Colony1 elution, 13: Chromo-Hal Colony2 supernatant, 14: Elution Colony2.

Table 2: Concentration of Chromo-Hal and Fre produced using *E. coli* BL21 (DE3) Arctic express and BL21 cells.

Enzymes	Concentration (mg/mL)
Chromo-Hal	10
Fre	26

4.2.2 Chromo-Hal Assays with Tryptophan (Trp)

With both the halogenase and the coenzymes expressed (Figure 17), attention was focused on testing Chromo-Hal with a wide range of assay conditions to identify the optimal condition for its activity. Firstly, a range of pHs were tested with the otherwise standard halogenase assay conditions (Figure 18, Figure 19 and Table 3).

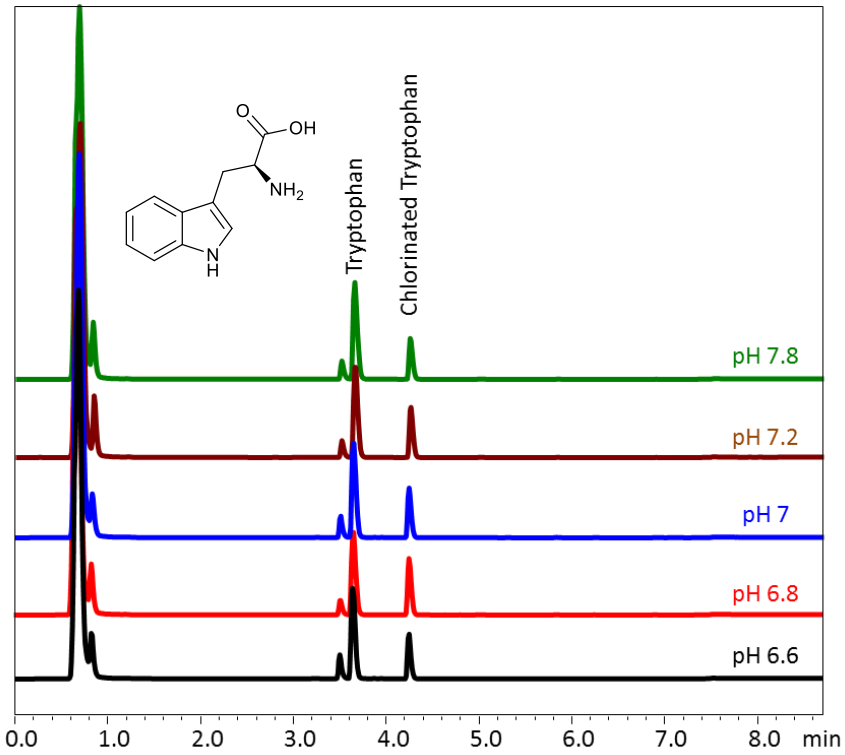


Figure 18: HPLC chromatogram showing the Chromo-Hal assay with Trp at 30 °C and variable buffer pH for 2 hours.

Table 3: Average conversion of Trp using different buffer pH and the % error calculated after assays in triplicate under the same conditions.

pH	Temperature (˚C)	Average conversion (%)	Error (+/-)
6.6	30.0	26.0	0.9
6.8	30.0	34.0	0.9
7.0	30.0	30.0	0
7.2	30.0	27.0	0.5
7.8	30.0	24.0	0.5

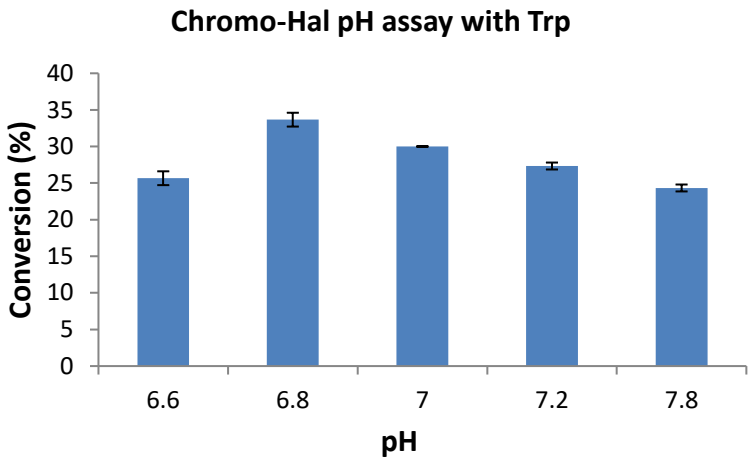


Figure 19: Bar chart which shows the % conversion of Trp using different buffer pH and the error of the assays.

It is clear from the pH data (Figure 18, Figure 19 and Table 3) that buffer pH 6.8 provides the highest conversion (average 34 %) while the use of pH above and below 6.8 drops the conversion. Concurrently, the temperature of the standard assay conditions was varied in order to find the optimal temperature.

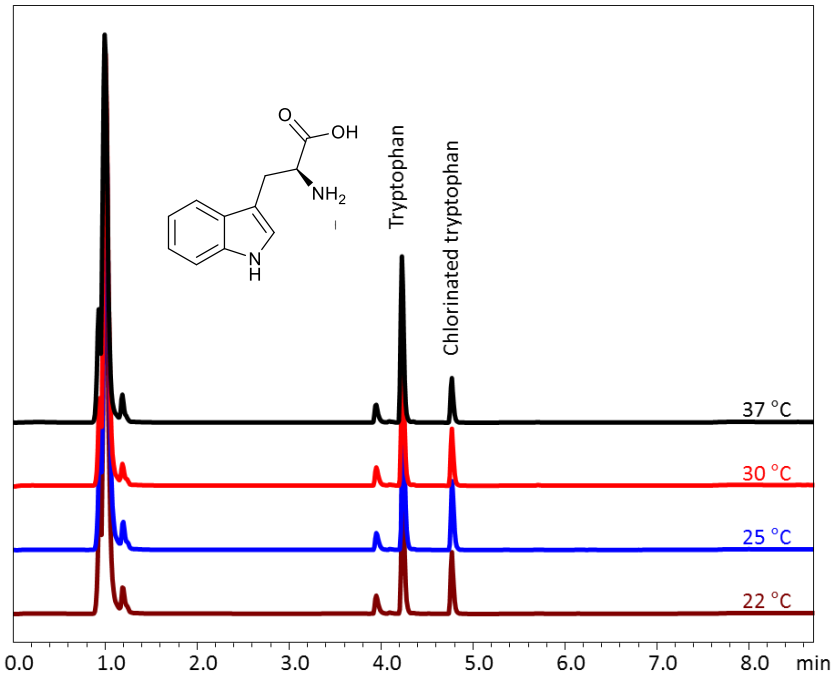


Figure 20: HPLC chromatogram showing the Chromo-Hal assay with Trp at variable T and buffer pH 7.2 for 2 hours

Table 4: Average conversion of Trp using different temperatures and the % error calculated from assays in triplicate under the same conditions.

pH	Temperature (˚C)	Average Conversion (%)	Error (+/-)
7.2	22.0	33.0	3.9
7.2	25.0	39.0	5.8
7.2	30.0	31.0	2.8
7.2	37.0	23.0	2.5

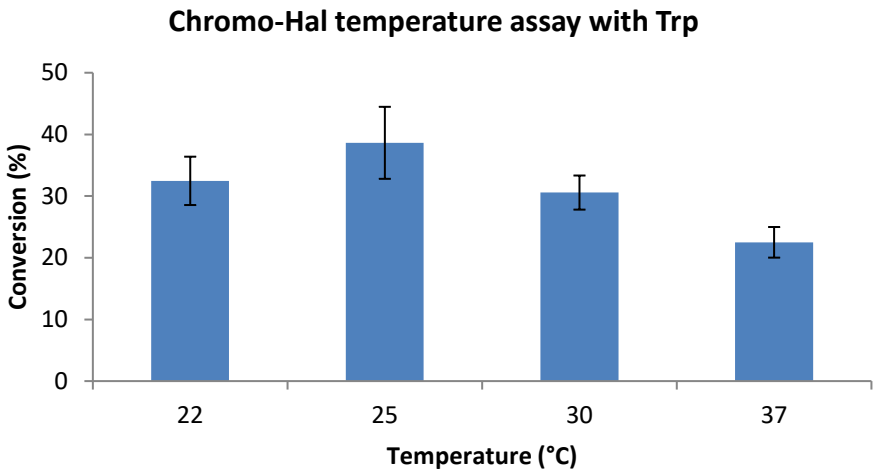


Figure 21: Bar chart which shows the % conversion of Trp using different temperatures.

From the data obtained (Figure 20, Figure 21 and Table 4) it can be seen that 25 ˚C is the optimum temperature as it gives 39 % average conversion which is considerably higher than conversions at 22, 30 or 37 ˚C.

Following optimisation of the temperature and pH of the assay, a comparison of the incorporation of either chlorine or bromine as a halogen source was tested (Figure 22).

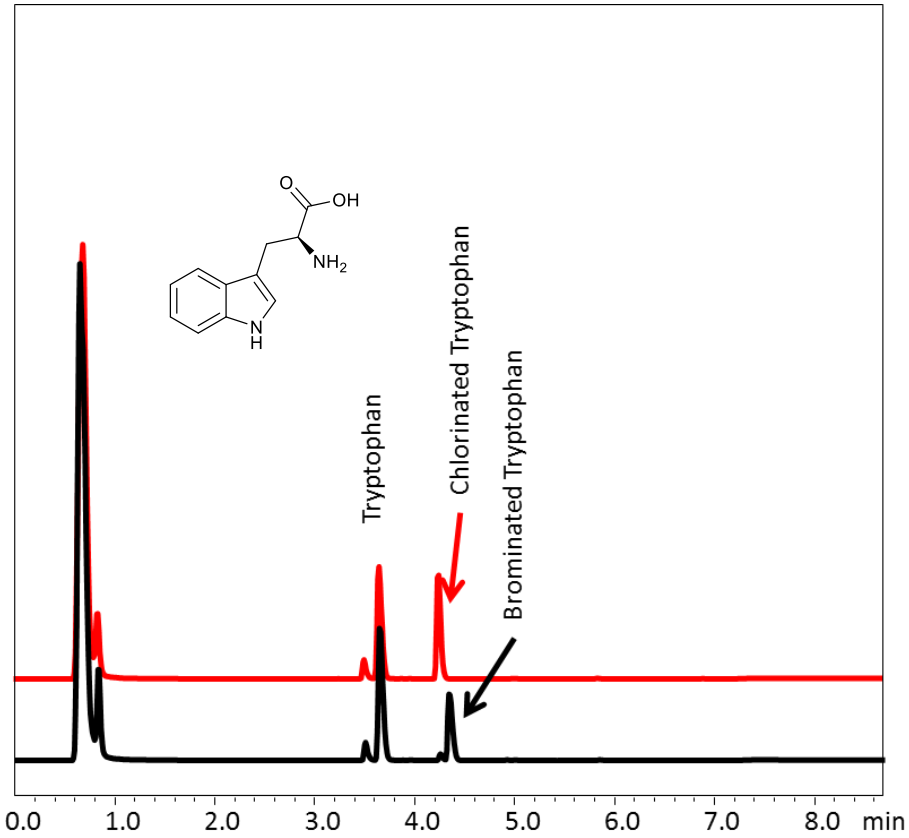


Figure 22: HPLC chromatogram showing the Chromo-Hal assay with Trp using Cl or Br source.

Table 5: Average conversion of Trp using different halogen source and the % error calculated with assays in triplicate under the same conditions.

Conditions	Average Conversion (%)	Error (+/-)
Chlorination	45.0	2.6
Bromination	27.0	5.8

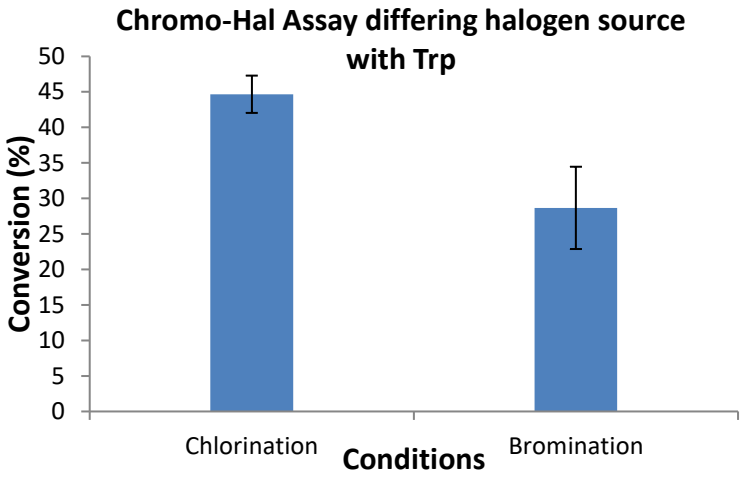


Figure 23: Bar chart which shows the conversion of Trp using different temperatures.

Comparing chlorination and bromination of tryptophan as it is shown in Figure 23 and Table 5, chlorination presents higher conversion. Consequently, it was decided that the further assays would be performed at 25 °C, buffer pH 6.8 and with a Cl source. A time course assay was performed to monitor the conversion over time (Figure 24 and Table 6).

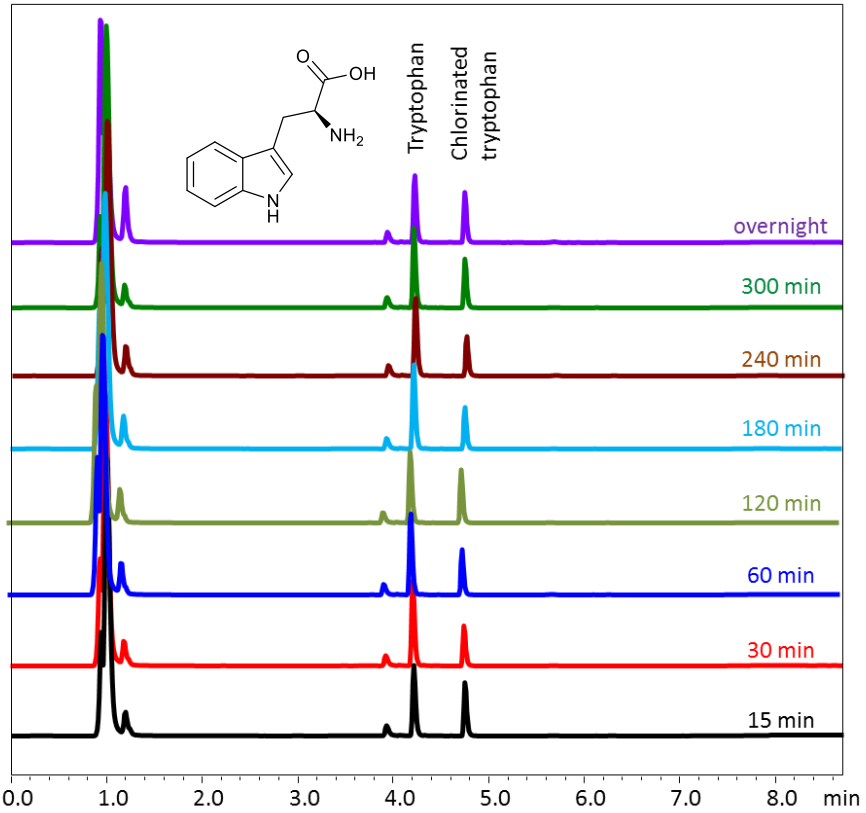


Figure 24: HPLC chromatogram showing the time course Chromo-Hal Assay with Trp, buffer pH 6.8 and T 25 °C.

Table 6: Conversions of Trp in a time course Chromo-Hal assay, the % error was not calculated.

Time (min)	Conversion (%)
15	42
30	30
60	34
120	42
180	31
240	32
300	37
360	38
Overnight	41

It is indicative from the data (Figure 24, and Table 7) that the conversion in the first 15 min is higher than the conversion after hours. What is interesting from this assay is that the conversion is constantly fluctuating. It might be some errors in the assays likely. An additional assay for shorter time was essential, trying to monitor the conversion of Trp in the first 15

minutes (Figure 25, Figure 26 and Table 7). The conversion progress presents the same trend as before. It was not only changing but also decreasing which is unusual and unexpected.

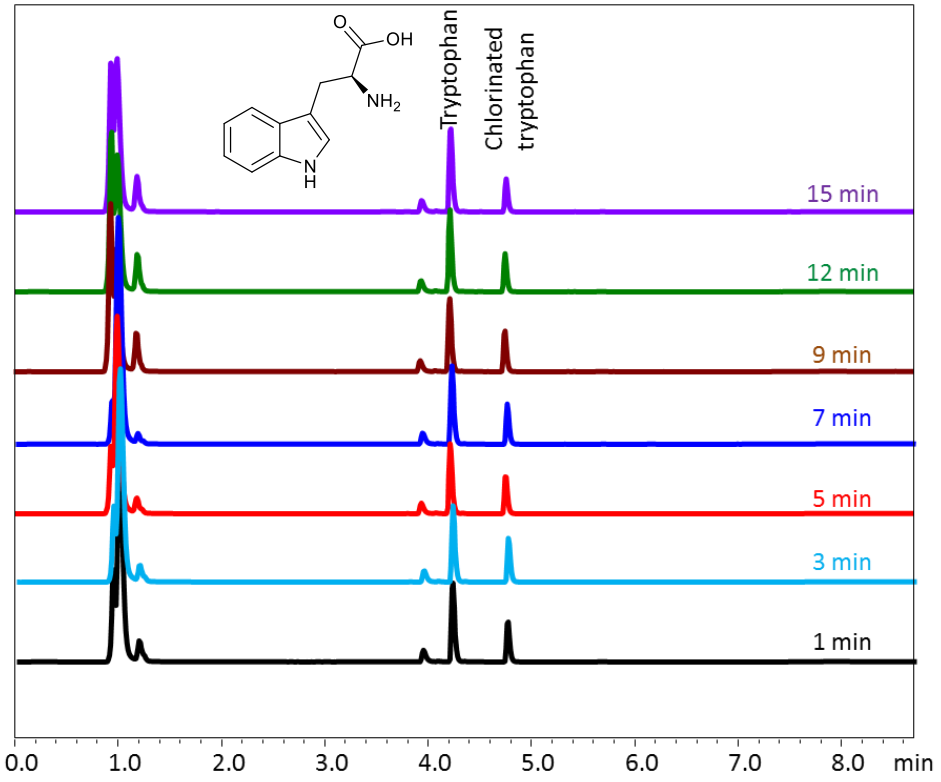


Figure 25: HPLC chromatogram showing the 15 minute time course Chromo-Hal Assay with Trp, buffer pH 6.8 and T 25 °C.

Table 7: Conversions of Trp in a 15 minute time course Chromo-Hal assay, the % error was not calculated.

Time (min)	Conversion (%)
1	33
3	36
5	34
7	32
9	34
12	30
15	26

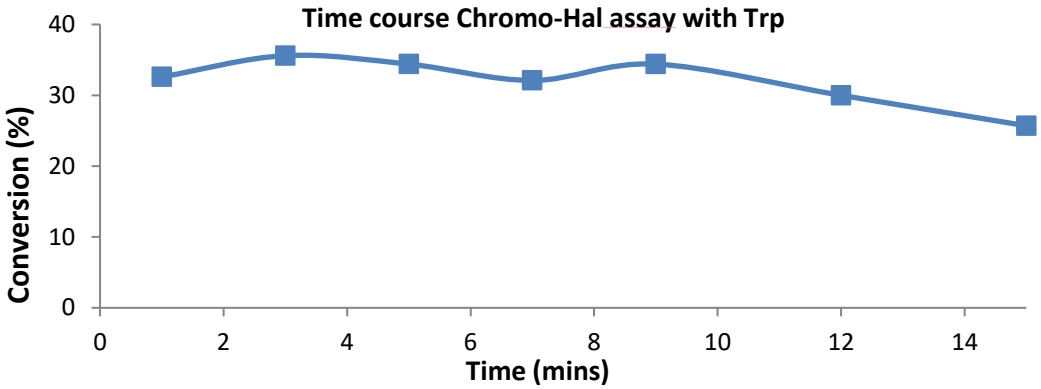


Figure 26: Chart which shows the conversion of Trp in a time course Chromo-Hal assay.

The activity of the enzyme peaked within the first minute, making it hard to monitor the trend in activity during this time. Trying to solve the issue, the assay was repeated reducing Fre concentration which was initially 5 µM. It was considered that high Fre concentration could affect the enzyme activity, and therefore it could be the reason that the assay showed high conversion of trp in the first minute. A short 30-minute assay was performed with different Fre concentrations, lower than 5 µM, to determine a concentration giving adequate conversion within a longer time frame. The data showed that Fre 3 µM gives better conversion (Figure 27, Figure 28 and Table 8). Following assays were then performed with Fre 3 µM.

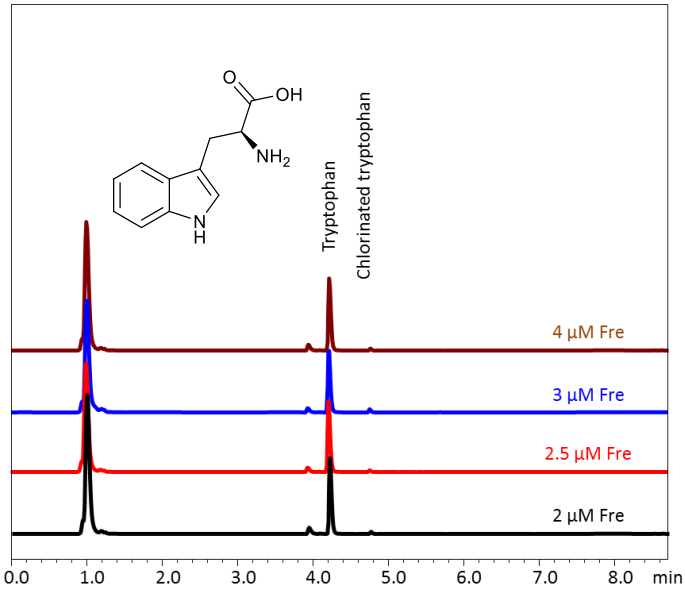


Figure 27: HPLC chromatogram showing the Chromo-Hal assay using different Fre concentrations.

Table 8: Conversions of Trp in a Chromo-Hal assay using different Fre concentrations, the % error was not calculated.

Fre C (µM)	Conversion (%)
2	2.5
2.5	2.1
3	4.2
4	2.4

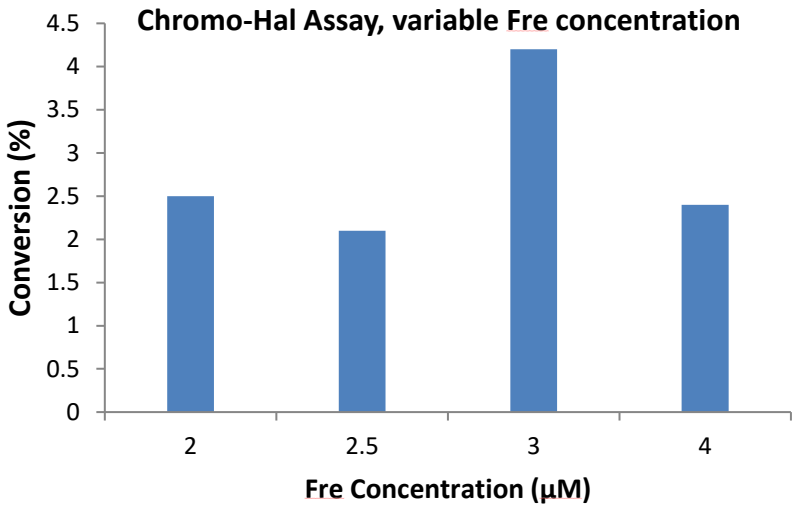


Figure 28: Bar chart which shows the conversion of Trp in a Chromo-Hal assay using different Fre concentrations.

Under the optimum condition (pH, T, Fre concentration), an additional assay (Figure 29, Figure 30 and Table 9) was performed trying to monitor the conversion of the substrate. Initially, the enzyme concentration was 20 mM, so 10 mM enzyme concentration was chosen for a Chromo-Hal assay with Trp. By decreasing the concentration of the enzyme, it was hoped that conversion could be monitored better. Decreasing the enzyme concentration, the conversions presented also a decreasing trend in the first mins but the issue of fluctuation still remained. The conversion over the time showed an increasing trend for the first 10 minutes and then it started fluctuating, as reported in Table 9.

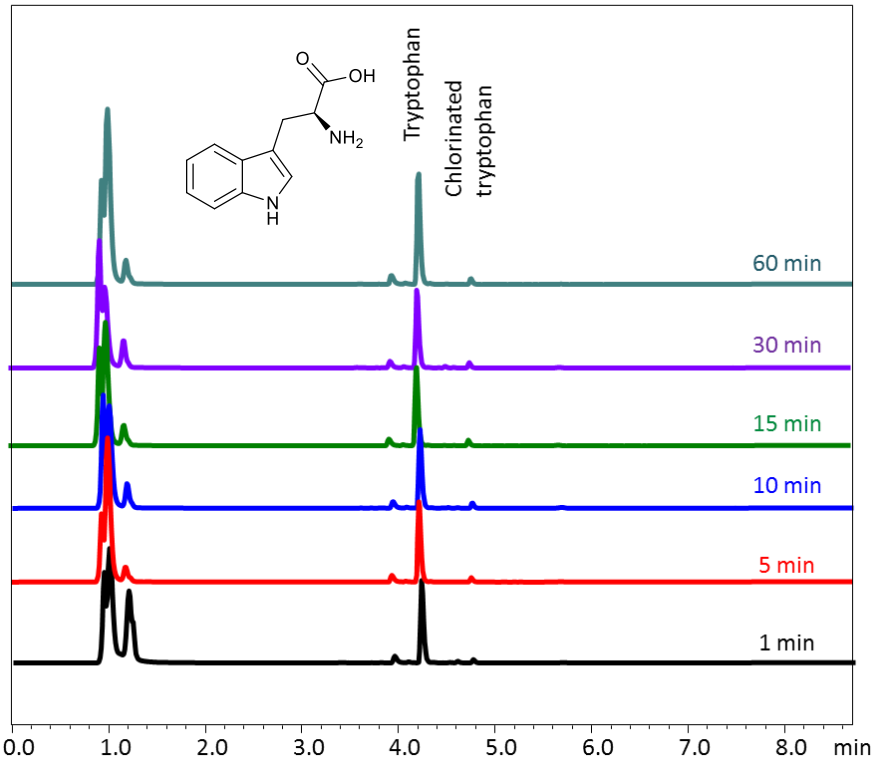


Figure 29: HPLC chromatogram showing time course Chromo-Hal assay with Trp (10mM Chromo-Hal concentration), using Fre 3 µM.

Table 9: Conversions of Trp in a 30 minute time course Chromo-Hal assay with reduced chromo-hal concentration (10 mM). The % error was not calculated.

Time (min)	Conversion (%)
1	3.4
5	3.7
10	5.7
15	5.5
30	5.5
60	4.1

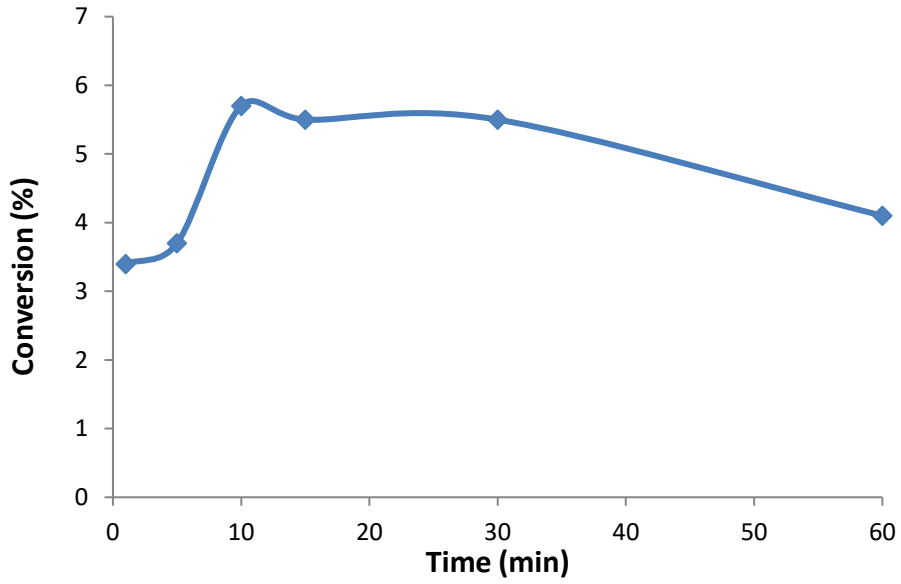


Figure 30: Bar chart which shows the conversion of Trp in a time course Chromo-Hal assay using reduced chromo-hal concentration (10 mM).

Despite the changes in enzyme and Fre concentration, the conversion was not increasing but still fluctuating and the enzyme seemed to lose its activity after 15 – 30 min. Undoubtedly, there was a limitation that blocked the enzyme activity. It was considered that higher concentration of NADH or addition of alcohol dehydrogenase (ADH) could improve the issue. So, additional assays were carried out including these enzymes (Figure 31Figure 32, Figure 33, Table 10). ADHs are a family of enzymes which are responsible for the oxidation of alcohols to the corresponding aldehyde or ketone by the transfer of a hydride anion to NAD^+ with release of a proton. The reduction of NAD^+ returns the system to NADH to participate in the FAD reduction cycle. Purified ADH was provided by Dr. Sarah Shephard.

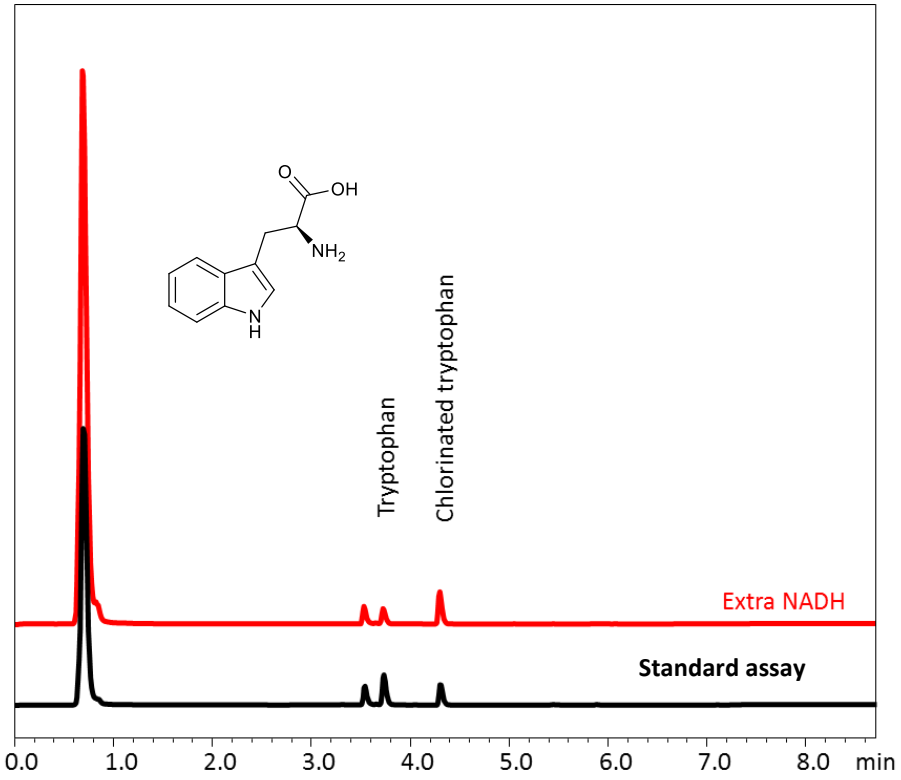


Figure 31: HPLC chromatogram showing Chromo-Hal assay with trp with extra NADH.

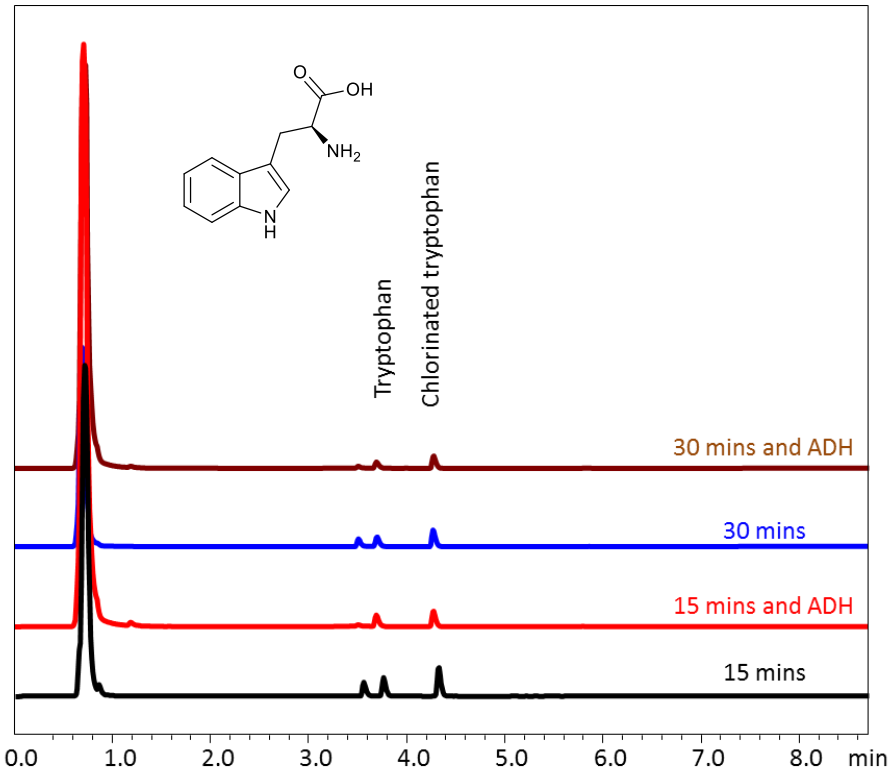


Figure 32: HPLC chromatogram showing Chromo-Hal assay with trp with addition or absence of ADH.

Table 10: Conversions of Trp in a Chromo-Hal assay using ADH, the % error was not calculated.

Time	Conversion (%)
15 mins	62
15 mins and ADH	58
30 mins	63
30 mins and ADH	68

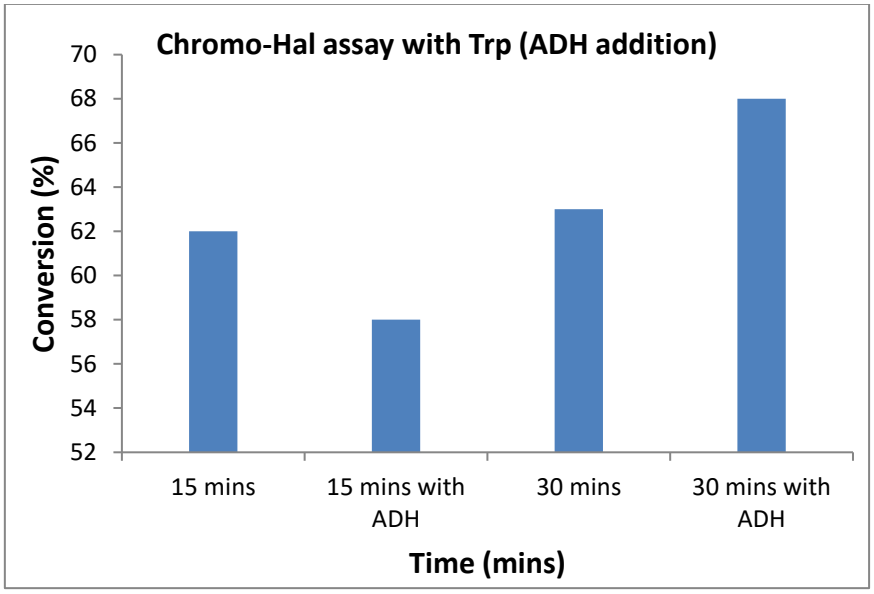


Figure 33: Bar chart which shows the conversion of Trp in a a Chromo-Hal assay using additional ADH.

It was interesting that both extra NADH and adding ADH improved the conversion of Trp. With extra NADH a conversion of 68 % was observed after 30 minutes (Figure 31). With the addition of ADH the conversion was also observed to be 68 % after 30 minutes (Figure 32, Figure 33 and Table 10). Both of these conversions were massively increased compared to previous experiments. However, a limitation in the assays was still present. It is noticeable that the conversion of trp was high in the first 15 min which is unusual, and it did not increase much further (Figure 32, Table 10) within the given period.

When trying to understand the decrease in conversion, it was speculated that the protein might be damaged by reactive oxygen species (ROS) generated *in situ* and the enzyme activity was blocked in 15 to 30 min.[77] Further assays were then carried out with the addition of catalase and superoxide dismutase (SOD) (Figure 34, Figure 36 and Table 11). SOD is an

enzyme which catalyses the dismutation of the superoxide (O_2^-) radical into either ordinary molecular oxygen or H_2O_2.[79] Catalase, an enzyme which is present in all living organisms exposed to oxygen and catalyses the decomposition of hydrogen peroxide (H_2O_2) to water and oxygen.[78] It is an important enzyme which protects the cells from oxidative damage by ROS.

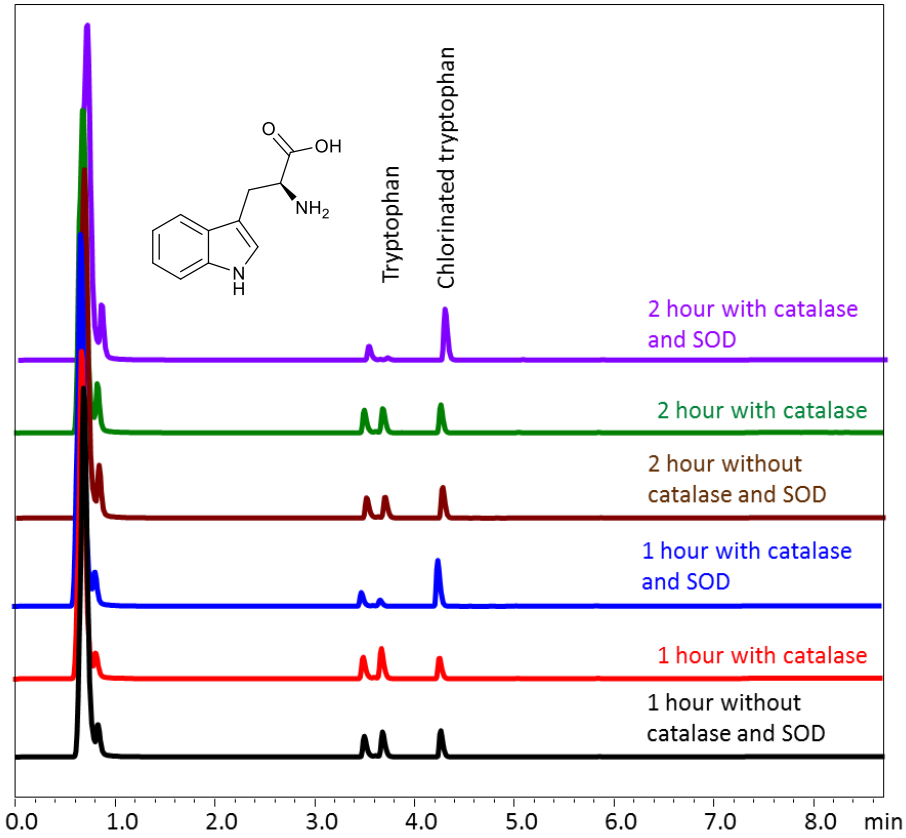


Figure 34: HPLC chromatogram showing Chromo-Hal assay with Trp with addition of catalase and SOD.

Table 11: Conversions of Trp in a Chromo-Hal assay using catalase and SOD, the % error was not calculated.

Conditions	Conversion (%)
1 hour	51
1 hour with catalase	41
1 hour with catalase and SOD	89
2 hours	59
2 hours with catalase	52
2 hours with catalase and SOD	98

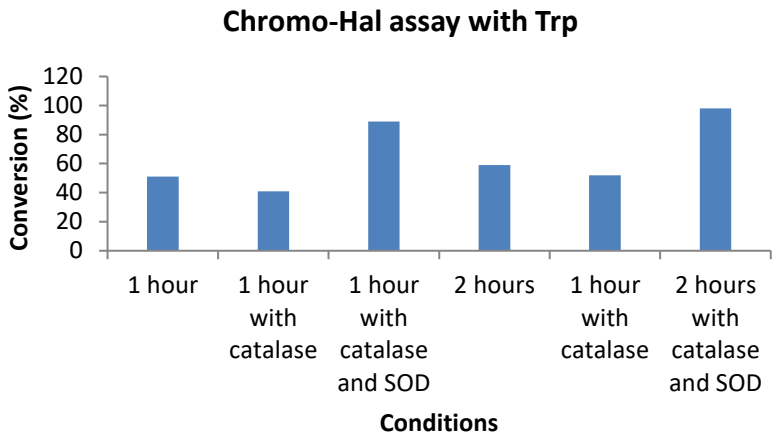


Figure 35: Bar chart which shows the conversion of Trp in a a Chromo-Hal assay using catalase and SOD

In the data shown above, containing SOD and catalase, the conversion was seen to significantly increase. In order to confirm this increase in conversion, assays were performed in triplicate over 1 hour (Figure 35, Table 12).

Table 12: Average conversion of Trp in a Chromo-Hal assay with catalase and SOD with the % error calculated after assays in triplicate under the same conditions.

Conditions	Average conversion (%)	Error (+/-)
1 hour	49.0	2.8
1 hour with catalase and SOD	79.3	7.8

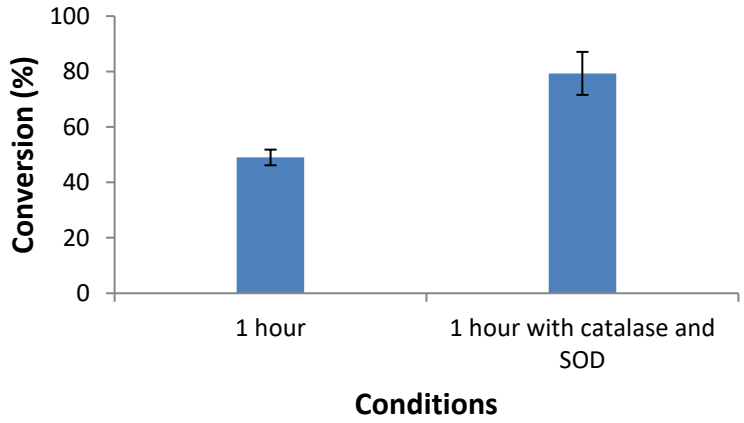


Figure 36: Bar chart which shows the conversion of Trp in a Chromo-Hal assay performed in triplicate using catalase and SOD.

Interestingly the results of the Chromo-Hal chlorination assay showed that with the addition of catalase and SOD, the conversion was impressively 98 % as reported in Table 11. With the addition of catalase and SOD the conversion was not fluctuating as before but presenting an increasing slope. In these assays ADH or extra NADH were not used.

4.2.3 Scaled up Biohalogenation

Previous studies clearly illustrate the poor stability, substrates and enzyme inhibition, low turnover and kinetic rates of using halogenase enzymes. Therefore, there is a need to circumvent these issues. In order to demonstrate the application of tryptophan halogenases in synthetic chemistry, an improvement of enzyme stability and recyclability is necessary. As this was achieved for the wild type RebH via the formation of combiCLEA, experimentation of the same protocol to other halogenating enzymes was envisioned. From previous work it has been shown that long term stability and therefore the efficiency of RebH can be significantly improved by immobilizing RebH with Flavin reductase and ADH, which provided continuous regeneration of the co-factors required for the reaction.[5] The CLEA is formed by saturation of the cell lysate with addition of excess ammonium sulfate $(NH_4)_2SO_4$, which causes the formation of non-covalently bound protein aggregates. Addition of bifunctional glutaraldehyde causes the cross linkage of aggregated moieties by the formation of imine bond between protein surface lysine residues and the aldehyde groups on the glutaraldehyde. The resulting cross-linking of these renders and the addition of glutaraldehyde form an insoluble homogenous protein mass which retains the proteins tertiary structure and therefore its activity.[80] (Figure 37)

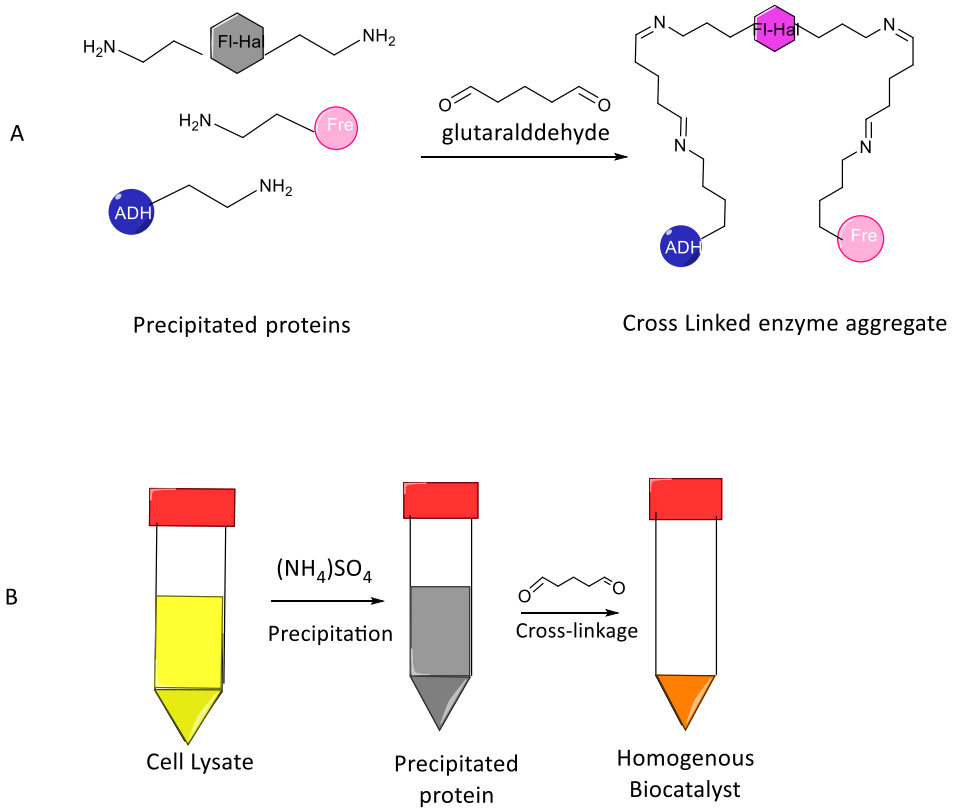
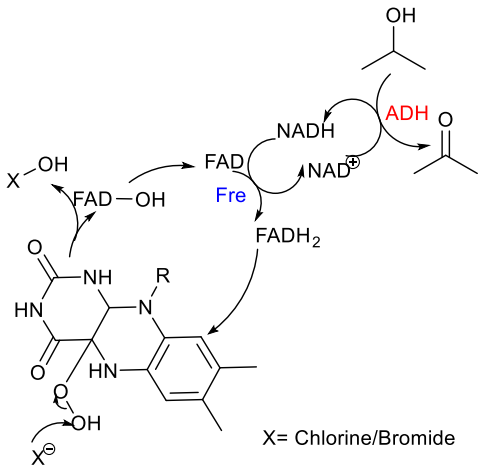


Figure 37: Schematic drawing illustrating the cross linkage between enzymes and glutaraldehyde resulting in Schiff base formation (A) and (B) the steps involved in CLEAs preparation.

Previous work within the Micklefield group showed that co-expressing halogenases with chaperone proteins such as pGro7 increases yield. pGro7 is a plasmid utilised in the TAKARA Chaperone Competent Cell BL21 Series to aid folding of proteins. As a result, the CLEA formation was prepared from halogenases co-expressed in BL21 cells containing this pGro7. As already reported by Frese *et al,* ADH is a coenzyme which removes the need for excess of NADH in the biohalogenation.[45] The ADH works the catalytic cycle to reduce NAD^+ to NADH by oxidation of isopropanol (IPA) to acetone with the use of NAD^+. The reduction of NAD^+ returns the system to NADH participate the FAD reduction cycle (Scheme 13).



Scheme 13: CLEAs recycling system using Chromo-Hal.

An active Chromo-Hal-Fre-ADH CLEA was generated, following the protocol as reported by Frese *et al*. and then used to test for activity against Trp. The enzyme activity gave a conversion of 24 % and purification via BondElut isolated the halogenated product (Figure 38 and Table 13).

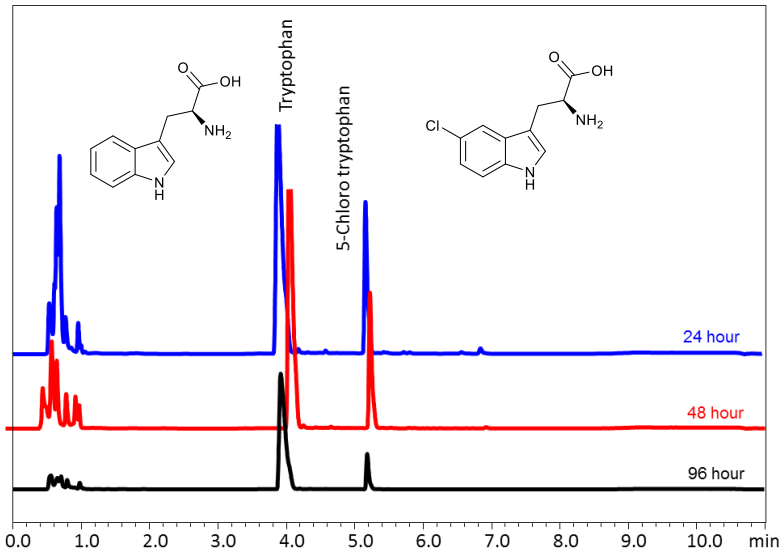


Figure 38: HPLC chromatogram showing Chromo-Hal assay with Trp in a CLEA.

Table 13: Conversions of Trp in a Chromo-Hal CLEA system.

Time (hour)	Conversion (%)
24	23
48	23
96	24

The NMR data of the halogenated product revealed that the halogenation was directed to the C5 position. This was confirmed from the proton (^{1}H), Correlation spectroscopy (COSY) and Nuclear Overhauser effect spectroscopy (NOESY). From the aromatic splitting patterns seen in ^{1}H NMR, it can be deduced that the regioselectivity of the chromo-Hal with trp is directed either to the 5 or 6 position. From COSY experiments W-coupling can be seen between peak 9 and 11, indicating that this singlet is found in the phenyl ring, it would then stand to reason that peak 6 is incorporated into the pyrrole ring as this displays no W-coupling and is downshifted in comparison to the other aromatics, due to its neighbouring nitrogen. Due to the symmetrical nature of the indole ring, it was necessary to perform further NMR experiments to identify the true regioselective nature of the enzyme, for this a NOESY experiment was performed. From the NOESY spectrum, it can be seen that there is coupling between the singlets found in the aromatic region and those found in the aliphatic carbon chain of the molecule. This indicates that these two hydrogens are within a 5 Å radius of the hydrogen atoms in the saturated chain, it can therefore be concluded that the Chromo-Hal is directing halogenation to the 5 position. If halogenation to the 6 position were to be true this coupling wouldn't be seen, as the position of the phenyl singlet would be beyond the 5 Å radius required to see coupling to these hydrogens. This is incredibly interesting, despite low conversion, as there is only one other fully characterised halogenase regioselective to the 5 position, namely PyrH.[36] This expands the collection of enzymes available to diversify future halogenated pharmaceuticals and the expression/assay could be optimised further in the future to be more applicable in industry.

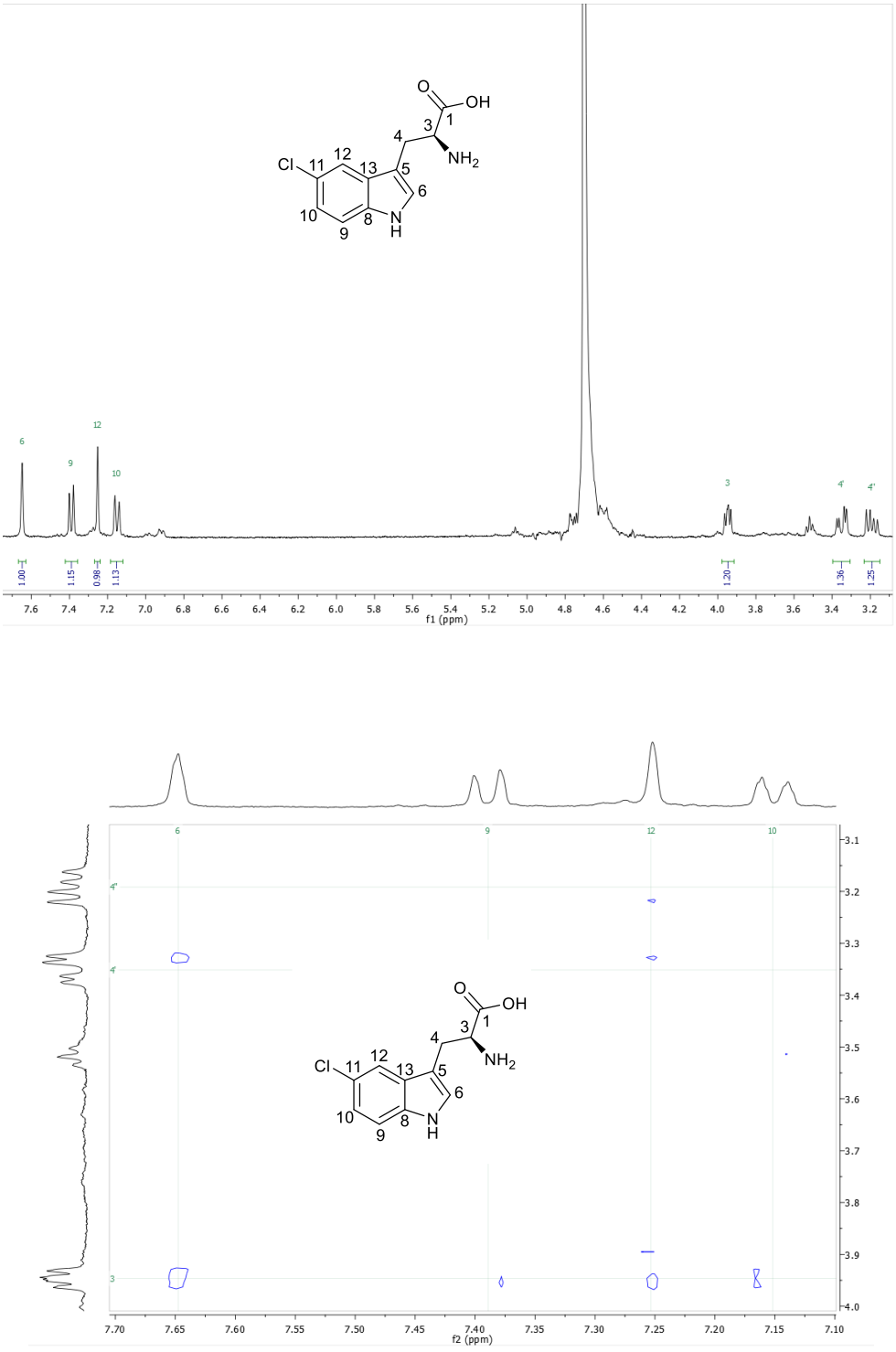


Figure 39: ^{1}H and NOESY NMR spectra of the halogenated product.

4.2.4 Chromo-Hal assays with new substrates

Chromo-Hal halogenase activity has been confirmed with Trp and the assay conditions optimised, therefore a range of new substrates were tested under these conditions in order to diversify its substrate scope. The new substrates to test consisted of tryptophan analogues – in order to increase the likelihood of successful conversion. These substrates have previously been shown to be accepted by other tryptophan halogenases. The new substrates to be tested included; anthranilic acid (Figure 40), tryptophol (Figure 41), Anthranilamide (Figure 43), with Phenylpiperazine (Figure 45), and with 2-amino-4-methylbenzamide (Figure 47).

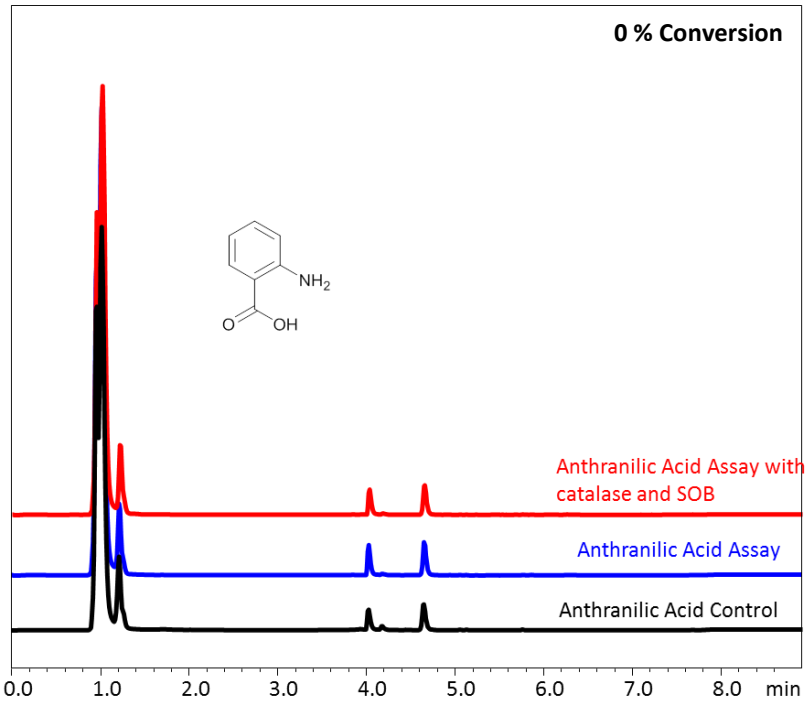


Figure 40: HPLC chromatogram showing Chromo-Hal halogenase assay with Anthranillic acid. Control contained no chromo-hal.

The chromo-hal did not accept the anthranilic acid as a substrate, as 0 % conversion was observed. The other substrates however, showed some success.

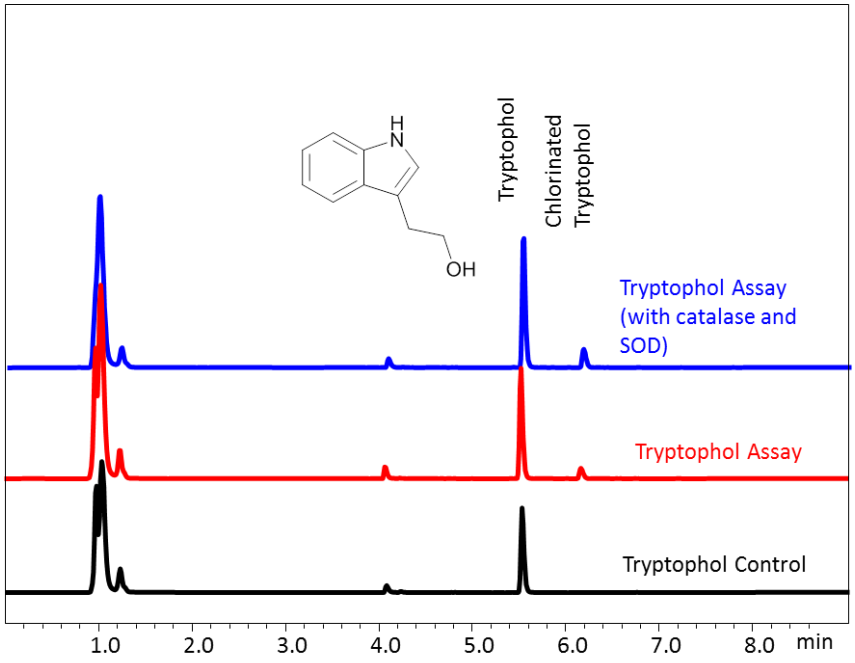


Figure 41: HPLC chromatogram showing Chromo-Hal assay with tryptophol. Control contained no chromo-hal.

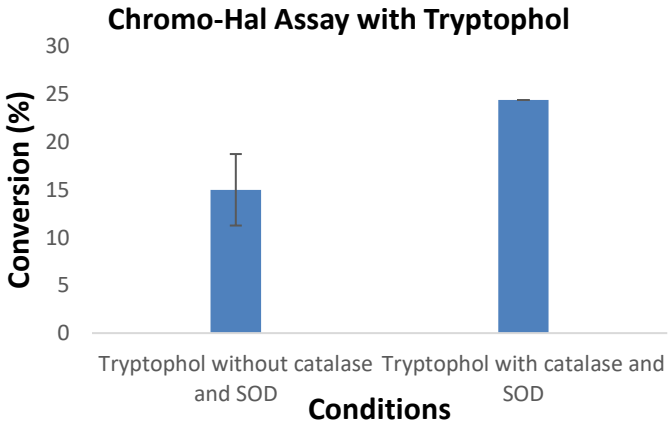


Figure 42: Bar chart which shows the conversion of tryptophol using different conditions.

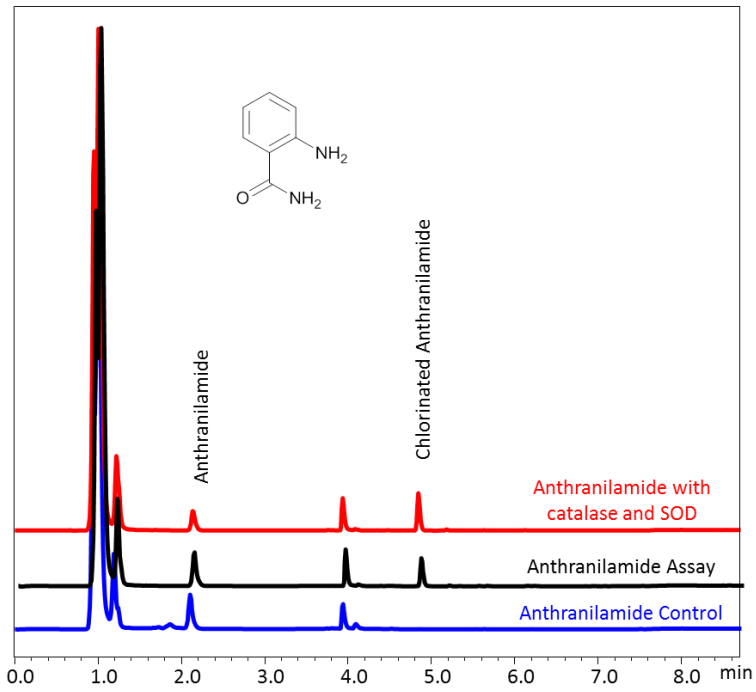


Figure 43: HPLC chromatogram showing Chromo-Hal assay with Anthranilamide. Control contained no chromo-hal.

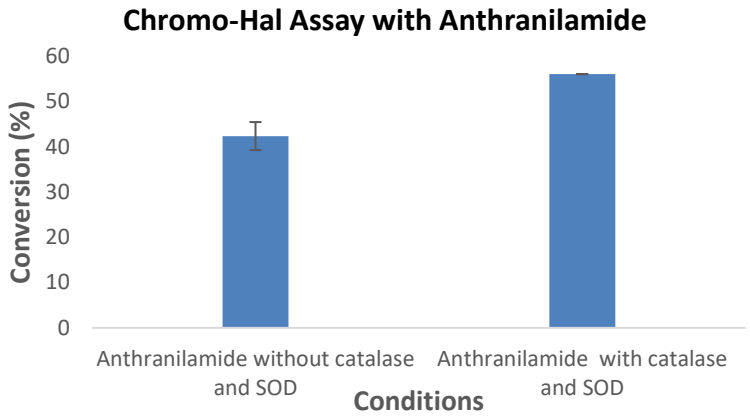


Figure 44: Bar chart which shows the conversion of anthranilamide using different conditions.

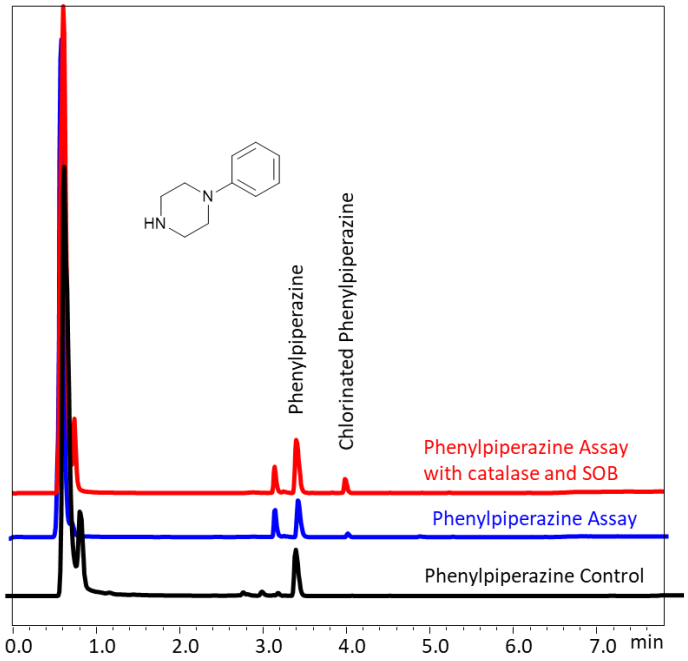


Figure 45: HPLC chromatogram showing chromo-hal activity with Phenylpiperazine. Control contained no chromo-hal.

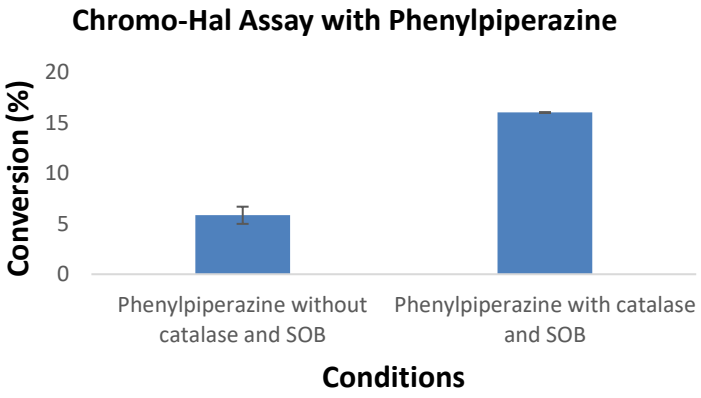


Figure 46: Bar chart which shows the conversion of phenylpiperazine using different conditions and the error bars when available.

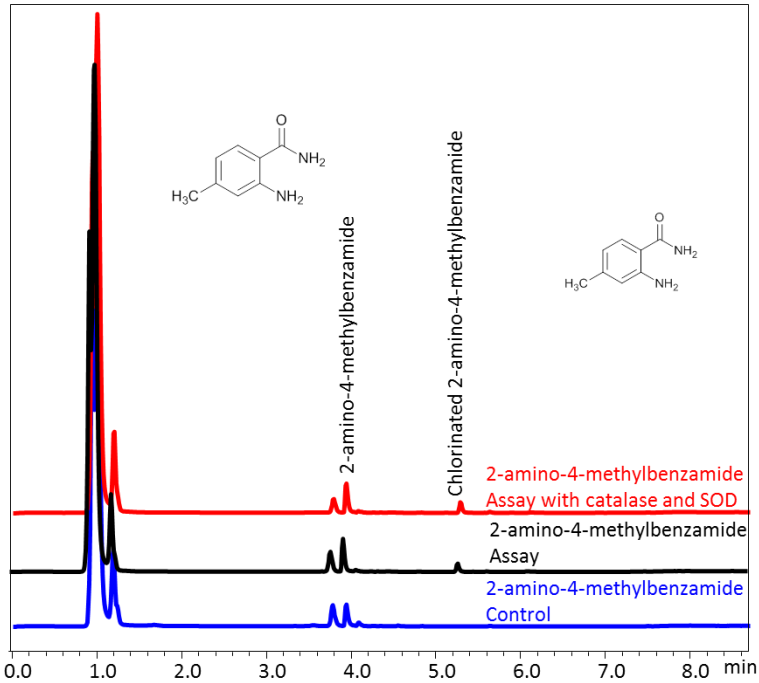


Figure 47: HPLC chromatogram showing Chromo-Hal assay with 2-amino-4-methylbenzamide. Control contained no chromo-hal.

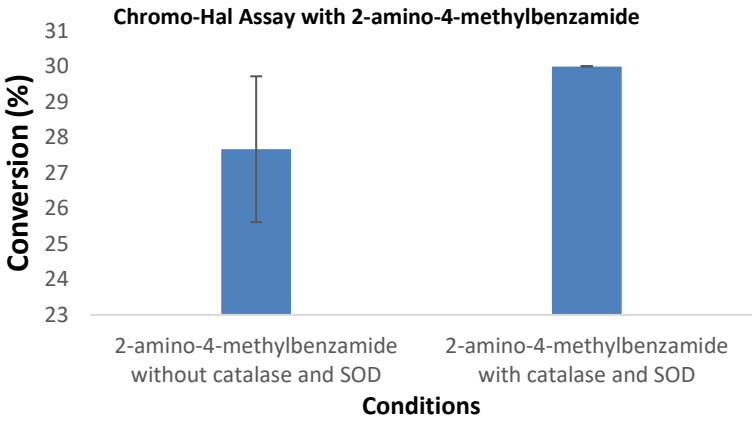


Figure 48: Bar chart which shows the conversion of 2-amino-4-methylbenzamide using different conditions.

All new substrates tested, apart from anthranillic acid, were accepted by the chromo-hal. The highest conversion was noticed in the assay with anthranilamide which was 56 % with catalase and SOD. The addition of catalase and SOD in all assays improved conversion. Conversions with other substrates included 24 % with tryptophol, 16 % with phenylpiperazine and 30 % with 2-amino-4-methylbenzamide. Even with these lower conversions, the acceptance of other substrates with this novel halogenase is a big step in the area of halogenase enzymes – especially due to the rare 5 regioselective nature of the enzyme. Both

the optimisation of expression of chromo-hal and further expansion of substrate scope in the future will increase the chance of implementation of this novel enzyme in the production of pharmaceuticals in the future.

4.3 Ochratoxin Halogenase

4.3.1 Expression of Ochratoxin Halogenase using *E. coli* BL21 competent cells

Previous work within the Micklefield group involved the ochratoxin halogenase, so pET-28a plasmid containing the ochratoxin halogenase was already available. The first attempt for the Ochratoxin halogenase expression was performed by transforming *E. coli* BL21 cells with this pET-28a plasmid containing the gene of interest. Protein expression and purification was performed using the Ni-NTA Affinity chromatography. The first attempt to express the protein of interest (ochratoxin halogenase) was not successful, however it was unclear whether the protein was not expressed or was insoluble. At the same time, expression and purification of Fre (Flavin Reductase) cofactor was attempted successfully, using BL21 competent cells (Figure 49).

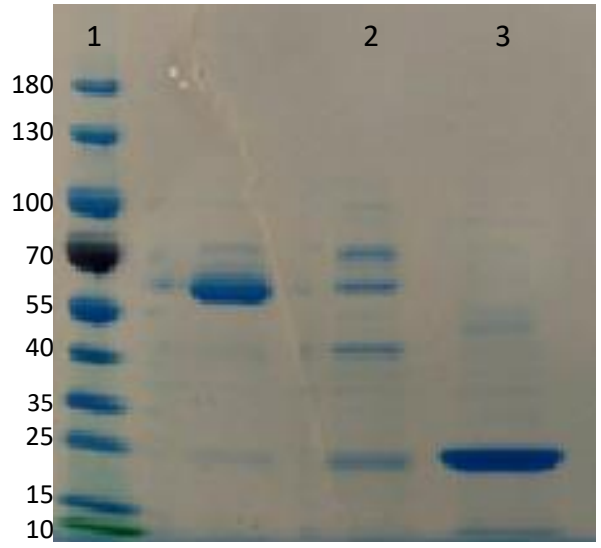


Figure 49: Image from the SDS-PAGE showing the unsuccessful production of the Ochra-Hl and the production of Fre, from left to right; 1: protein ladder, 2: Ochra-Hl (53 kDa), 3: Fre (26.2 kDa).

4.3.2 In vitro-transcription of pET-28a-Ochra-Hl

An *in vitro*-transcription assay using the PURExpress® In Vitro Protein Synthesis Kit was carried out to help determine if the issue was with the plasmid itself or the expression / purification method.

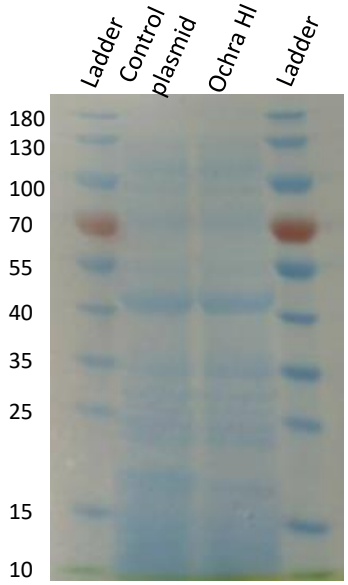


Figure 50: Image from SDS-PAGE showing the *in vitro*-transcription assay where from left to right; 1: protein ladder, 2: control plasmid (pET-28a), 3: *In vitro*-transcription assay, 4: protein ladder

The data showed that there was no expression *in vitro* (Figure 50). Screening tests with different *E. coli* strains and temperatures were performed, as this would be useful in-case there was an issue with protein expression.

4.3.3 Expression of Ochratoxin Halogenase using different *E. coli* strains and temperatures

The next attempt was made using various bacterial strains and different temperatures to monitor and identify if the protein remained in the insoluble fraction or was not expressed. So, a collection of different fractions was necessary. The choice of the competent cells was one of the determinants. More specifically, *E. coli* BL21, Arctic Express, BL21 pGro7, BL21 pKJE7, BL21 pTf16 and Rosetta were transformed with pET-28a-Ochra-Hl and incubated at four different temperatures, 12 °C, 20 °C, 30 °C, 37 °C. Arctic Express (DE3) were chosen as they increase the yield of soluble protein. This strain co-expresses the cold-adapted chaperonins which help and stabilise the unfolded or partially folded proteins. *E. coli* BL21 (DE3) pGro7 and pTf16 are from the TAKARA chaperone competent cell BL21 series. They are

used for the large-scale expression of tryptophan halogenases as they co-express the target protein with chaperones, increasing recovery of proteins in the soluble fraction. *E. coli* Rosetta 2 (DE3) is suitable for large scale expression of eukaryotic proteins that contain codons rarely used in *E. coli*. The BL21 pGro7 did not provide any colonies after the transformation so it was excluded from the screening test.

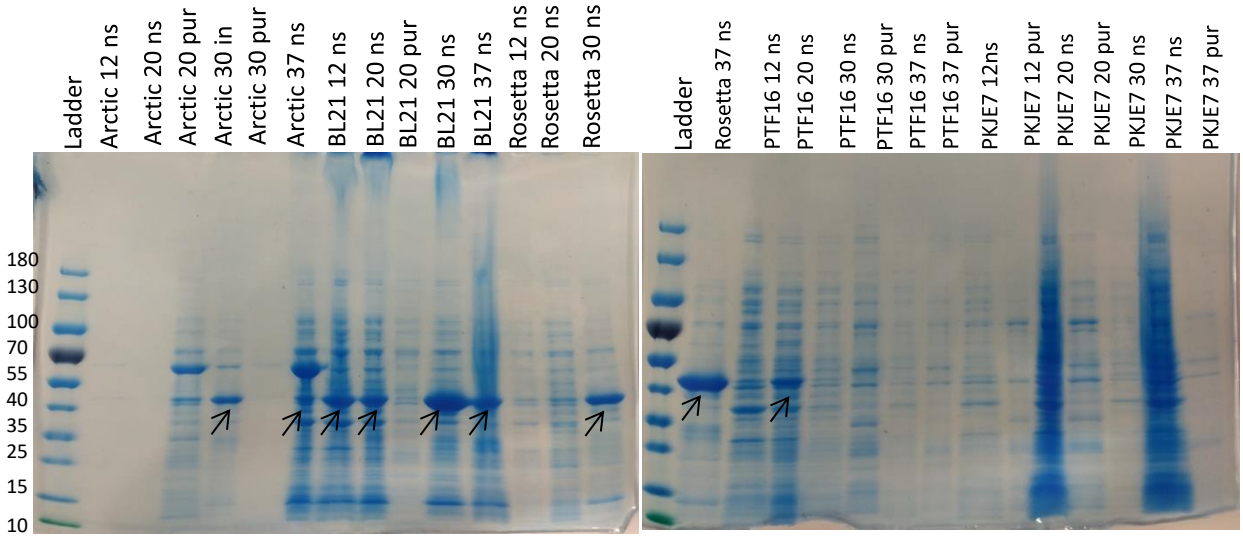


Figure 51: Images from SDS-PAGE showing the evolution of the Ochra Hl (53 kDa) expression, where ns refers to non-soluble part and pur to pure (soluble) part. The bands of interested are indicated in the non-soluble fractions with arrows. Strain and temperature are indicated in lane heading (eg 12 = 12 °C).

The protein gel after purification revealed that using BL21, Rosetta and Arctic express strains the enzyme was expressed in the insoluble fraction (Figure 51). Upon reanalysis of the gene sequence, to determine the persistent issue with insolubility, several options were identified to try to resolve this.

Firstly, a different version of the ochra-Hl sequence was discovered that had been manually dissected and analysed more thoroughly via bioinformatics. The exons and introns present due to its fungal origin had been analysed thoroughly, to remove any nonsense sequence that could have been causing folding issues. This sequence was more highly cited in the literature and upon analysis seemed more likely to be the correct sequence. Furthermore, a different codon-optimisation technique was implemented. The previous codon-optimisation technique used implemented the most common codon used in *E. coli* for each amino acid irrespective of the codons surrounding each. However, the new technique uses a more sophisticated technique of considering the nearby codons – as *E. coli* do not tend to include repeated

codons. A synthetic gene for the ochratoxin halogenase was ordered taking these new factors into account.

4.3.4 Restriction digest and ligation of the synthetic genes

The plasmid containing the synthetic gene (pUC57-Ochra-Hl) was restriction digested with XhoI and NdeI in order to transfer the gene of interest into vector of choice, pET-28a, via ligation. The digested genes and plasmid were run on a DNA agarose gel and extracted using the MinElute® gel extraction Kit (Figure 52).

MCherry is a red fluorescent protein which is usually used as a marker when tagged to molecules and cell components. The protein is a 28.8 kDa monomer and creates an in-frame fusion to track the protein of interest. Also, mCherry has shown improved brightness and photostability compared to other markers. It is possible that the protein of interest can be co-expressed with another protein such as mCherry.[74] In addition, maltose-binding protein (MBP) is a protein part of the maltose-maltodextrin system of *E. coli* and has a molecular weight of 42.5 kDa. Both of these proteins fused to a protein of interest can be used to increase the solubility of recombinant proteins by facilitating the proper folding of the proteins expressed in *E. coli*. Usually, the MBP is expressed as a MBP-fusion protein, preventing possible aggregation of the protein of interest.[75, 76]

In order to attempt such protein fusions as mentioned, vectors pET-28a-MBP and pET-28a-mCherry were also digested with the same enzymes as shown in Figure 52.

A synthetic gene of the Ochra-P450 was also ordered which was optimised for *E. coli*. The same procedure repeated for the Ochra-P450 gene.

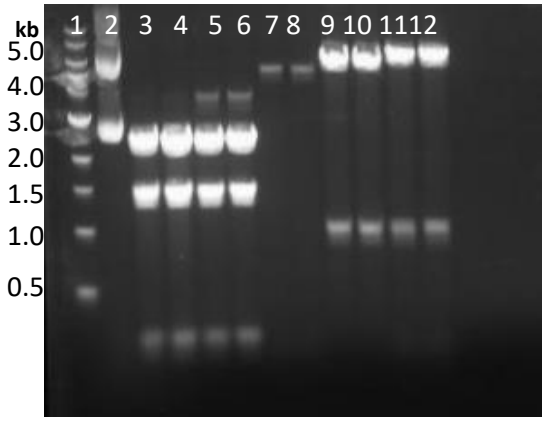


Figure 52: Image of the DNA gel showing restriction digest of the synthetic gene and the vector where from left to right; 1: DNA ladder (1 kb DNA Ladder, NEB), 2: Control, 3-6: pUC57-Ochra-Hl, 7-8: pET-28a, 9-10: pET-28a-MBP, 11-12: pET-28a-mCherry. Digested pUC57 plasmid expected ~2.7 kb and the ochratoxin halogenase gene at ~1.4 kb. pET-28a vectors were expected to produce bands at ~5.3kb, pET-28a-mCherry at ~5.9 kb and pET-28a-MBP at 6.5 kb.

The pET-28a was excluded from further work as it was not successfully digested (Figure 52). It is shown that the digested pUC57 appears on the gel at 2.7 kb and the genes for ochre-Hl and ochra-P450 at 1.4 kb and 1.5 kb, respectively, as expected (Figure 53).

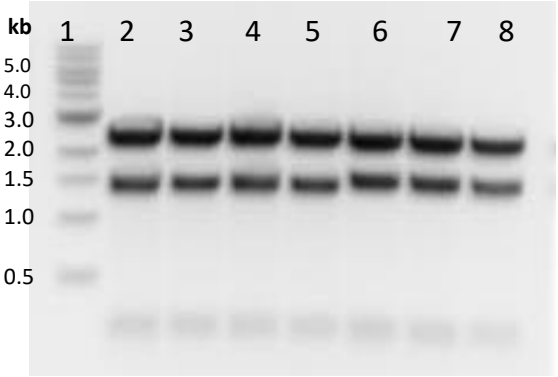


Figure 53: Image of the DNA gel showing restriction digest of the synthetic genes where from left to right; 1: DNA Ladder (1 kb DNA Ladder, NEB), 2-5: pUC57-Ochra-Hl, 6-7: pUC57-Ocha-P450. Digested pUC57 appears on the gel at 2.7 kb and the genes for ochre-Hl and ochra-P450 at 1.4 kb and 1.5 kb, respectively, as expected. Digested pUC57 plasmid expected ~2.7 kb.

Once successfully recovered and extracted from the DNA gel, ligation was performed and *E. coli* DH5α competent cells were transformed with the ligated vectors. After the overnight incubation at 37 °C on agar plates, positive colonies were inoculated following the same protocol that already mentioned above. The ligation pET-28a-mCherry with Ochra-P450 gene

was not successful and did not produce any colonies when transformed, so it was excluded from further work.

The next day, following the plasmid extraction, restriction digest was carried out to confirm the successful ligation. The ochratoxin Hl gene size is 1.4 Kb, Ochra-P450 1.5 kb, and pET-28a vector 5.4 Kb, as it shown in Figure 54.

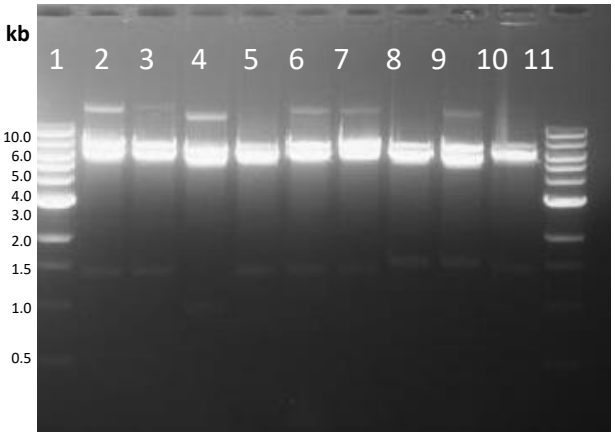


Figure 54: Image of the DNA gel showing the restriction digest of plasmid containing the genes of interest where from left to right; 1, 11: DNA ladder (1 kb DNA Ladder, NEB), 2-4: pET-28a-MBP-OchraHl, 5-7: pET-28a-mCherry-Ochra-Hl, 8-10: pET-28a-MBP-Ochra-P450. The ochratoxin halogenase gene at ~1.4 kb and the ochra-P450 at 1.5 kb. pET-28a vectors were expected to produce bands at ~5.3kb, pET-28a-mCherry at ~5.9 kb and pET-28a-MBP at 6.5 kb.

4.3.5 Expression of Ochratoxin Halogenase using different vectors

After confirmation, E. coli BL21 competent cells were transformed with extracted plasmids for Ochra-Hl for protein expression. Isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.5 mM final concentration) was added to induce the protein expression. The two different cultures (MBP and mCherry fusions) were used and incubated for 3 hours and overnight respectively (Figure 55, Figure 56).

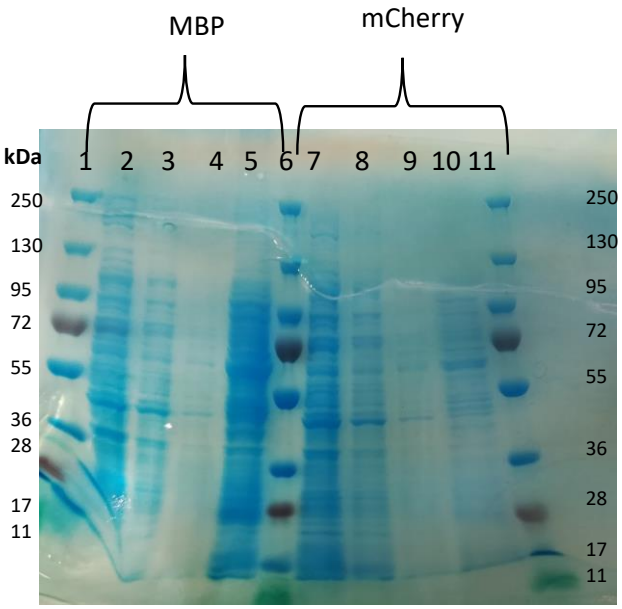


Figure 55: Image of SDS-PAGE showing the attempt of Ochratoxin Hl (53 kDa+28k Da for mCherry and 42 kDa for MBP) expression as a MBP or mCherry fusion after 3 hours incubation where from left to right; 1, 6, 11: Protein ladder, 2: flow through, 3: wash 1 (20mM imidazole), 4: elution (500mM imidazole), 5: Wash 2 (60 mM imidazole), 7: Flow through, 8: Wash 1, 9: wash 2, 10: elution. Fusion used is indicated above lanes.

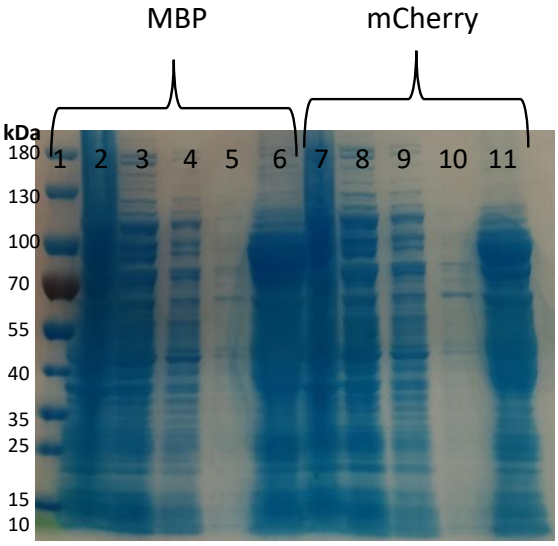


Figure 56: Image of SDS-PAGE showing the attempt of Ochratoxin Hl (53 kDa+28k Da for mCherry and 42 kDa for MBP) expression as an MBP or mCherry fusion after overnight incubation where from left to right; 1: Protein ladder, 2: insoluble fraction 3: flow through, 4: wash 1 (20mM imidazole), 5: Wash 2 (60 mM imidazole), 6: elution (500mM imidazole), 7: insoluble fraction, 8: Flow through, 9: Wash 1, 10: elution, 11: wash 2. Fusion used is indicated above lanes.

From the SDS-gel there is no sign of protein expression, therefore, cloning into the pET-28a vector alone was attempted again as the expression as fusion proteins had not been successful.

4.3.6 Restriction digest and ligation of the synthetic genes into pET-28a

The whole procedure of restriction digest and ligation was repeated with the synthetic genes and empty pET-28a (Figure 57). Overnight digestion of pET-28a resulted in incorrect digestion/star activity, therefore a digestion over two hours was performed (Figure 58).

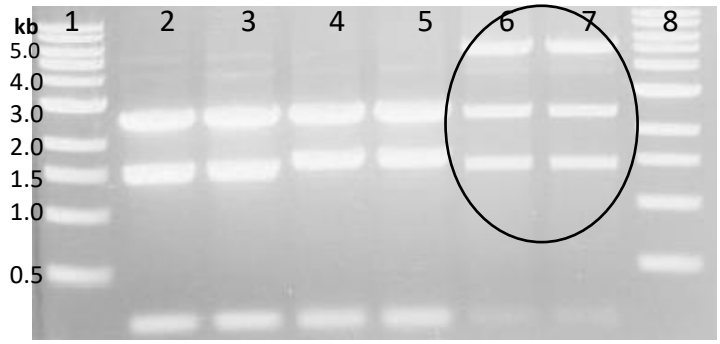


Figure 57: Image of the DNA gel showing the restriction digest of the synthetic genes (pUC57-Ochra-Hl and pUC57-Ochra-P450) and vector (pET-28a) where from left to right; 1, 8: DNA ladder (1 kb DNA Ladder, NEB), 2-3: pCU57-Ochra-Hl, 4-5: pUC57-Ochra-P450, 6-7: pET-28a. Incorrect restriction digest bands for pET-28a are highlighted with a circle. Digested pUC57 plasmid expected ~2.7 kb and the ochratoxin halogenase gene at ~1.4 kb, and ochra-P450 at 1.5 kb. pET-28a vectors were expected to produce bands at ~5.3kb.

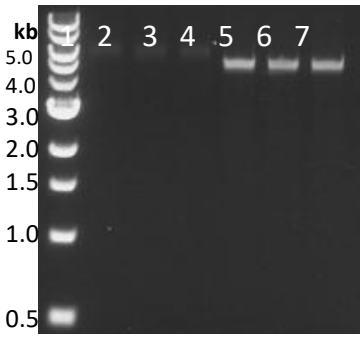


Figure 58: Image of the DNA gel showing the restriction digest of the pET-28a for 2 hours where from left to right; 1: DNA ladder (1 kb DNA Ladder, NEB), 2-7: pET-28a. The successfully digested plasmids (5-7) were recovered from the gel. pET-28a vectors were expected to produce bands at ~5.3kb.

The digested pUC57 vector is shown at 2.5 kb, while the ochra-P450 and halogenase at 1.5 kb and 1.4 kb. Following the same protocol, ligation was performed and for confirmation of the presence of the genes in the *E. coli* DH5α cells, colony PCR (for 3 selected colonies) was also carried out. Unfortunately, ligation of the pET-28a-Ochra-Hl was not successful so experiments were carried forward with the P450 only. Colonies 1 and 2 showed positive results as the digested pET-28a is shown at 5.4 kb and the Ochra-P450 gene at 1.5 kb (Figure 59).

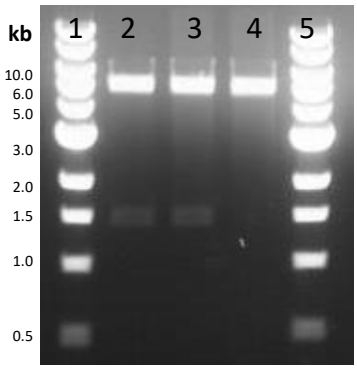


Figure 59: Image of DNA gel showing the colony PCR where from left to right; 1, 5: DNA Ladder (1 kb DNA Ladder, NEB), 2-4: pET-28a-Ochra-P450. The ochra-P450 was expected to produce bands at 1.5 kb.

The following day, after the plasmid extraction, BL21 competent cells were transformed with the ligated vector for expression and purification of the protein (Ochra-P450).

4.3.7 HiFi DNA Assembly

The ochra-Hl gene could not be successfully ligated with pET-28a, so HiFi DNA assembly was used instead of ligation. PCR amplified gene (ochratoxin Hl) is shown in Figure 60, where the gene of interest appears on the gel at 1.4 kb.

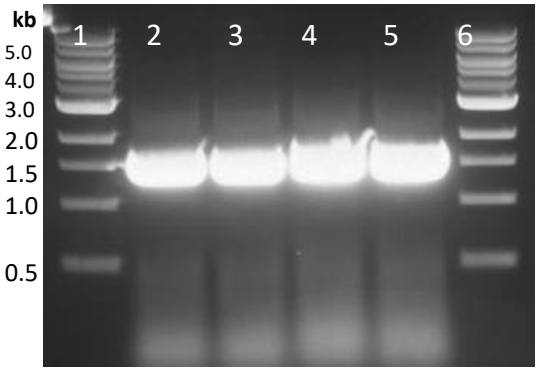


Figure 60: Image of DNA gel showing the PCR of the Ochra-Hl gene where from left to right; 1, 6: DNA ladder (1 kb DNA Ladder, NEB), 2-5: pET-28a-Ochra-Hl. The ochra-Hl was expected to produce bands at 1.4 kb.

Following HiFi assembly of the pET-28a-Ochra-Hl, transformation and colony PCR were performed to confirm the presence of the gene where the genes of interest (ochre-hl and ochre-P450) are shown at 1.4 kb and 1.5 kb (Figure 61). For the Ochra-Hal *E. coli* DH5α cells were used and for the P450 *E.* coli BL21 colonies were tested. For the halogenase control sample synthetic gene (pUC57-Ochra-Hl) was used and for the Ochra-P450 control pET-28a-Ochra-P450 previously extracted was used, respectively.

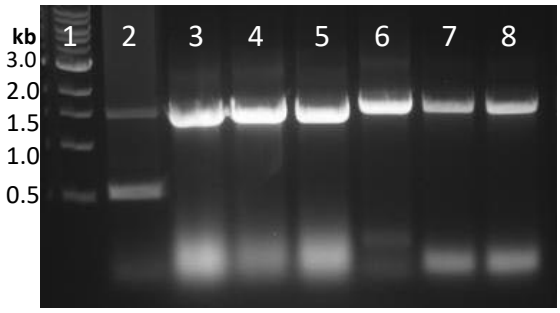


Figure 61: Image of DNA gel showing the Colony PCR where from left to right 1: DNA Ladder (1 kb DNA Ladder, NEB), pUC57-Ochra-Hl control, 3-5: Ochra-Hl colonies, 6: pET-28a-Ochra-P450 control (from previous extracted plasmids), 7-8: Ochra-P450 colonies. The ochra-Hl was expected to produce bands at 1.4 kb and the ochre-P450 at 1.5 kb.

Following confirmation by colony PCR, in order to be certain that the gene had successfully been cloned, an analytical restriction digest was performed. The restriction digest was performed on the extracted plasmids for both pET-28a-Ochra-Hl and pET-28a-Ochra-P450 (Figure 62).

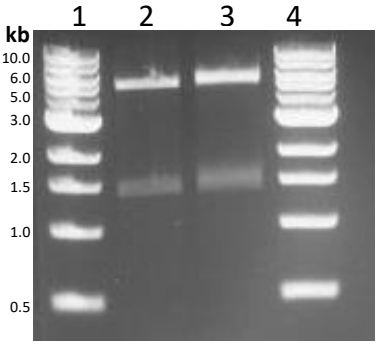


Figure 62: Image of DNA gel showing the restriction digest for the genes confirmation where from left to right; 1, 4: DNA ladder (1 kb DNA Ladder, NEB), 2: pET-28a-Ochra-Hl, pET-28a-Ochra-P450. The ochratoxin halogenase gene was expected at ~1.4 kb, and ochra-P450 at 1.5 kb and pET-28a vectors were expected to produce bands at ~5.3kb.

The restriction digest for the genes of interest was successful as it is shown in Figure 62 where there are characteristic bands at 1.5 and 1.4 kb.

4.3.8 Expression and purification of Ochra-P450 and Ochratoxin Hl

Expression and purification of the enzymes were attempted using *E. coli* BL21 cells. Despite the image of the protein gel which is not clear as it is overloaded, the Ochra-P450 expression does not seem to be successful. If it had been successful a distinct band would be observed at ~58.4 KDa. It is also indicative from the SDS-gel that the Ochra-Hl enzyme (53 KDa) was not expressed either.

Despite the successful ligation of the genes a possible reason of the unsuccessful expression of the proteins could be the non-natural environment (*E. coli)*. The ochratoxin halogenase and the Ochra-P450 are both from fungal species, and their expression in bacteria organism might not be favourable. However, it was considered that perhaps a N terminal transmembrane region identified in the P450 enzyme could be causing issues with its folding/expression. Therefore, as a final attempt to try to express and characterise the P450 enzyme, truncations were attempted using various designed primers.

4.3.9 Expression of the truncated Ochra-P450

A final attempt for expressing the Ochra-P450 was made, designing primers to remove the N terminal trans-membrane domain of the protein. It was thought that the trans-membrane would impede the expression of the protein or result in protein aggregation and misfolding due to the hydrophobic nature of transmembrane segments. Three different combinations of primers were designed and named as Primers A, B, C, D.

- Primers A and B (reaction 1) were aimed to amplify the full length of the gene (Ochra-P450), with C-Terminal 6 His-Tag and cloned into pET-21b plasmid.
- Primers C and B (reaction 2) were aimed to amplify the gene for the truncated protein without the N-terminal trans-membrane domain, with C-terminal 6 His-Tag and cloned into pET-21b.
- Primers C and D (reaction 3) were aimed to amplify the gene for the truncated protein without the N-terminal transmembrane domain, with N-terminal 6 His-Tag and cloned into pET-28a.

PCR reaction was carried out as it is shown in Figure 63 where the truncated genes were amplified and appeared on the gel at approximately 1.5 kb. The genes were recovered from the DNA gel, extracted, digested using XhoI and NdeI. There are 3 characteristic bands at ~1.5 kb in Figure 64. Ligation/transformation was performed and confirmed by colony PCR (Figure 65), following the same protocol. The ligation of the reaction 3 was successful, so the attempt for expression of the Ochra-P450 was made with the reaction 3.

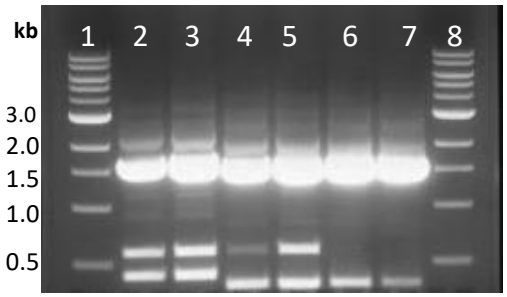


Figure 63: Image of DNA gel showing the PCR reactions of truncated Ochra-P450, where from left to right; 1, 8: DNA ladder (1 kb DNA Ladder, NEB), 2-3: Reaction 1, 4-5: Reaction 2, 6-7: Reaction 3. The ochra-P450 was expected at ~1.5 kb.

Digested plasmid pET-21b was provided by Dr. Brian Law.

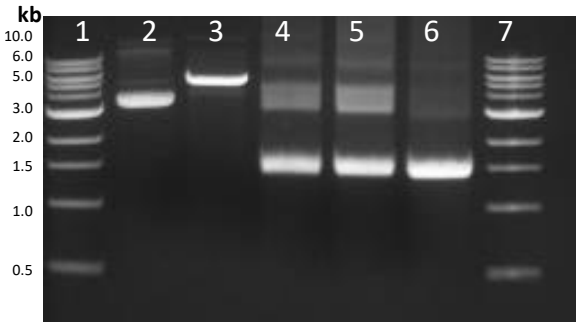


Figure 64: Image of DNA gel showing the restriction digest where from left to right; 1, 7: DNA ladder (1 kb DNA Ladder, NEB), 2: control (pUC57-Ochra-P450), 3: pET-28a, 4: Reaction 1, 5: Reaction 2, 6: Reaction 3. The ochra-P450 was expected at ~1.5 kb pET-28a vectors were expected to produce bands at ~5.3kb.

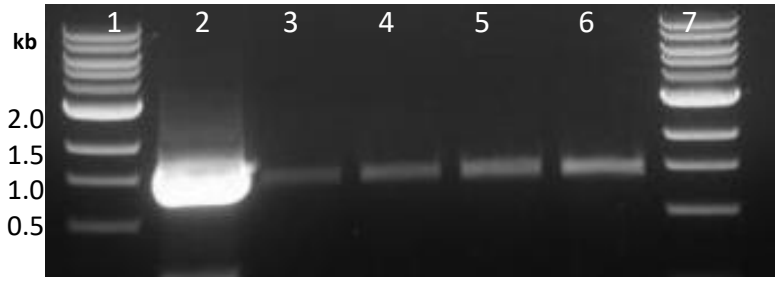


Figure 65: Image of DNA gel showing the Colony PCR, where from left to right; 1, 7: DNA ladder (1 kb DNA Ladder, NEB), 2-6: colonies from reaction 3 ligation (colonies 1-5). The ochra-P450 was expected at ~1.5 kb.

From the protein gel it was indicative that the Ochra-P450 enzyme was not expressed. The difficulty in expressing these genes was likely due to unfavourable expression in *E. coli* cells. The origin of these enzymes are fungal, therefore, despite codon-optimisation and other attempts to aid folding *in vivo*, the system is too different from their natural host. Perhaps in the future expression in a more similar organism could be attempted, such as in Yeast, as these contain similar structured DNA/protein expression systems.

5 Conclusions

In conclusion, showing the activity of halogenases in whole cell assays was explored – despite a lack of activity being seen. *E. coli* cells have been shown to be unsuccessful in whole cell halogenase assays, therefore Vmax™ cells were explored as an option. Due to the marine origin of the organism, it was thought that they would be a more suitable host for halogenase assays. Unfortunately, the whole cell assays were unsuccessful – however, this could be due to the low yield of the halogenases. The fact that detectable activity was achieved in these Vmax™ expressed halogenases at all is a novel finding, as use of this halophile has not been reported with expression of halogenases previously. If the expression was optimised further in this organism, perhaps the yield could be increased and this may enable detectable activity in whole cell assays.

An attempt was also made to extend the repertoire of halogenases available for production of pharmaceuticals. A novel tryptophan halogenase identified from bioinformatics analysis of the *S. chromofuscus* genome was expressed and screened against a library of substrates and conditions. With the tryptophan natural substrate, halogenation regioselective to the 5 position was confirmed. This observed activity is a significant step in expanding the panel of tryptophan halogenases, as currently this is limited to only one other wild-type enzyme selective to this position, namely PyrH. This enzyme was tested with several other substrates, with detectable conversion achieved in most. Future work could be performed on this enzyme to expand its substrate scope and improve the yield and/or conversion. Improvements in enzyme activity have already been achieved in this project through addition of SOD and catalase. Mutations could potentially be introduced into the enzyme to direct the activity against new substrates and potential targets for the production of pharmaceuticals in the future.

The attempt to express and characterise new halogenases was explored further with the Ochratoxin halogenase identified in the Ochratoxin A mycotoxin biosynthesis pathway. Not only would the successful characterisation of this enzyme expand the repertoire of known halogenases, however it would aid in the understanding of a biosynthesis pathway involved in the production of one of the major polyketide mycotoxins faced in agriculture. In relation to this aim, characterisation of the P450 was attempted in parallel. Although expression and characterisation was unsuccessful despite many optimisation and alterations within the gene,

it was established that perhaps the usual method of initially attempting expression in *E. coli* is not suitable. A more similar system to the original fungal host is recommended to achieve soluble, active enzyme. It is predicted that the issues arose from incorrect identification of introns/exons in the sequence published on the databases. This would be avoided through extraction of the mRNA from the host organism, production of cDNA and then introduction of the gene into yeast. The production of cDNA from the host mRNA would ensure the removal of all the unwanted introns and using yeast will provide a similar host environment to increase the chance of correct folding and avoid insolubility issues. Due to the timescale of the project, this will be explored in future work.

6 Experimental Section

6.1 Buffer preparations

6.1.1 Preparation of 100 mM KPi buffers

500 mL bottles were filled with 1M K_2HPO_4 aqueous solution, 1 M KH_2PO_4 aqueous solution, 500 mM NaCl and the desired mM of imidazole was added, the mixture was pH adjusted to pH=7.2 and was made up to a final volume of 500 mL with distilled water. All buffers were stored at 4˚C until required.

Table 14: Quantity of imidazole for buffer preparation.

Desired imidazole concentration (mM)	Imidazole quantity (g)
0	0
10	0.34
80	2.72
300	10.21

6.1.2 Preparation of 10 mM KPi buffer

In a 100mL bottle 100mM KPi buffer was added and the solution was diluted 1:10 (or tenfold up to you) with distilled water and the pH was adjusted to pH=7.2 and the buffer was stored at 4 ˚C until required.

6.1.3 Preparation of 100 mM sodium phosphate buffer

A 500 mL bottle was charged with $Na_2HPO_4.12H_2O$ (12.25 g), $NaH_2PO_4.2H_2O$ (2.46 g) and NaCl (14.61 g), to this was added 500 mL of distilled water. After complete dissolution the pH was adjusted to pH=7.2 and the buffer was stored at 4 ˚C until required.

6.1.4 Protein storage buffer

20 mM Hepes (0.95 g), 50 mM NaCl (0.58 g) was made up to 200 mL containing glycerol solution (10% v, v). The pH was then adjusted to pH=7.2.

6.2 Preparation of Growth Media

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

6.2.1 Lysogeny Broth (LB)

10 g tryptone, 5 g yeast extract and 10 g NaCl were made up to 1L with distilled water. The media was sterilised by autoclaving at 121°C for 20 minutes prior to use.

6.2.2 Enhanced 2xYT medium

20 g Yeast Extract, 32 g tryptone, 17 g NaCl, 0.2% glucose and 17.6mM Na_2HPO_4 were made up to 1L with distilled water. The pH of the media was adjusted to pH=7.4 prior to sterilisation by autoclaving at 121°C for 20 minutes prior to use.

6.2.3 Lysogeny Broth Agar

10 g tryptone, 5 g yeast extract, 10 g NaCl and 15 g agar were made up to 1L with distilled water. The media was sterilised by autoclaving at 121°C for 20 minutes prior to use.

6.3 Preparation of agar plates

LBA was placed in the microwave until fully melted. 25 mL of LBA was transferred into a 50 mL sterile falcon tube under sterile conditions. 25 µL of the appropriate antibiotic solution was added to the solution. The mixture was poured into clear polystyrene petri dish in a laminar flow cabinet and left at room temperature until the LBA solidified.

Table 15: Target proteins and their appropriate antibiotics.

Target protein of plasmid	Appropriate antibiotic solution
PrnA, PyrH, RebH, SttH, Chromo-Hal, Ochra-Hl	Kanamycin (50mg/mL)
pGro7	Chloramphenicol (20mg/mL)
Fre, ADH	Ampicillin (50mg/mL)

6.4 Competent cell preparation

6.4.1 Competent cell preparation solution (KMES I)

A solution of MES (25 mM) and CaCl2 (60 mM) in distilled water was adjusted to pH 5.8 with addition of KOH. $MgCl_2$ and $MnCl_2$ were added to reach a final concentration of 5 mM each the mixture was then sterilised by autoclaving at 121°C for 20 minutes. KMES I buffer was stored at 4˚C until required.

6.4.2 Competent cell storage solution (KMES II)

A solution of MES (25 mM) and $CaCl_2$ (60 mM) in a mixture of glycerol and distilled water (10% v/v) was adjusted to pH 5.8 with addition of KOH. $MgCl_2$ and $MnCl_2$ were added to reach a

final concentration of 5 mM each and the mixture was then sterilised by autoclaving at 121°C for 20 minutes. KMES II buffer was stored at 4˚C until required.

6.4.3 Competent *E. coli* cells preparation

An *E. coli* strain was grown overnight in 10 mL LB with incubation at 37 ˚C, 180 rpm with appropriate antibiotic. The culture was then diluted fifty-fold into 100mL autoclaved LB broth and incubated at 37˚C and 180 rpm until an optical density (OD_{600}) of 0.4 was reached. The culture was chilled on ice for 10 min and then centrifuged at 4 ˚C and 1500 rpm for 10 min. The cell pellet was resuspended in 100 mL of chilled KMES I buffer. The mixture was chilled on ice for 1 h and centrifuged at 4 ˚C and 1500 rpm for 10 min. The cell pellet was again resuspended in 10 mL of chilled KMES II buffer. Competent cells were finally split into 100 µL aliquots and stored at -80 ˚C until required.

6.4.4 General method for cell transformation and gene inoculation

The suitable stored competent cells were defrosted on iced and a 50 µL aliquot was taken and added to a sterile 1.5 mL eppendorf. Purified DNA plasmid (3 – 5 µl) was then added and the mixture was incubated on ice for 30 mins. The competent cell solution was then heat shocked for 40 seconds at 42 ˚C in a heat block and then incubated on ice for further 5 mins. 500 µL of autoclaved LB media was added and the cells were incubated for 1 hour at 37 ˚C with shaking at 180 rpm. 150 - 200 µL of the cells were spread onto a prepared agar plate which was incubated overnight at 37 ˚C.

6.5 SDS-PAGE electrophoresis

6.5.1 X10 SDS running buffer

In 1 L bottle Tris-base (30 g), glycine (144 g) and SDS (10 g) were added. The mixture was dissolved in 1 L of distilled water. The resulting buffer was stored at room temperature and diluted 10-fold prior to use.

6.5.2 X4 Loading dye

In a mixture of glycerol and distilled water (40% v/v) β-mercaptoethanol (0.5% m/v), bromophenol blue (0.04% m/v) SDS (8% m/v) and Tris HCl (240 mM final concentration) were added. The mixture was stored at -20 ˚C until needed.

6.5.3 SDS-PAGE analysis

X4 loading dye (5 µL) was added to protein samples (15 µL). The samples were heated at 95˚C for 10 min and centrifuged at 13000 rpm for 5 min. 12 µL of the mixtures were loading onto the SDS-PAGE gel (8-16% gradient SDS gels, purchased from VWR International). Electrophoresis was carried out according to the supplier's instructions (220 V for 45 mins) and the gels were then stained with Instant Blue™ with orbital shaking at room temperature. This was later washed with distilled water to allow bands to be visualised.

6.5.4 General method for Flavin Reductase (Fre) Expression

The gene of interest for Fre expression was already ligated into pET45b (+) plasmid as part of previous work. The plasmid was transformed into *E. coli* BL21 (DE3) as previously described using ampicillin antibiotic (50 mg/mL) for selection. After the overnight incubation at 37 °C, single colonies from the agar plate were picked and inoculated in autoclaved LB medium with amplicillin (50 mg/mL), this was incubated overnight at 37 °C with shaking at 180 rpm. Under sterile condition in 8 x 2 L conical flasks containing 400 mL LB medium, 4 mL of the starter culture were transferred with 400 µL amplicillin (50 mg/mL). The flasks were incubated at 37 °C with shaking at 180 rpm until the OD_{600} reached 0.6. Once the desired OD_{600} was reached, IPTG (0.1 mM final concentration) was added to induce protein expression and the flasks were incubated for a further 3 hours at 30 °C with shaking at 180 rpm. Cells were then harvested by centrifugation (4000 rpm, 20 min, and 4 °C) and the pellets were stored at 4 °C prior to protein purification.

6.6 General method for halogenases expression in *E. coli* Arctic Express (DE3)

The gene of interest for halogenases expression was already ligated into pET-28a (+) plasmid as part of previous work. The plasmid was transformed into *E. coli* Arctic express (DE3) as previously described using kanamycin antibiotic (50 mg/mL) for selection. After the overnight incubation at 37 °C, single colonies from the agar plate were picked and inoculated in autoclaved LB medium with kanamycin and gentamycin antibiotics (50 mg/mL and 20 mg/mL), this was incubated overnight at 37 °C with shaking at 180 rpm. Under sterile condition in 8 x 2 L conical flasks containing 400 mL LB medium, 4 mL of the starter culture were transferred without antibiotic selection. The flasks were incubated at 30 °C with shaking (180 rpm) for about 3 hours. Then, cultures were cold shocked for at least 30 min and IPTG (0.1 mM final concentration) was added to induce protein expression, the flasks were further

incubated at 16 °C with shaking at 180 rpm overnight. Cells were then harvested by centrifugation (4000 rpm, 20 min, and 4 °C) and the pellets were stored at -20 °C prior to protein purification.

6.7 General method for halogenases expression in *E. coli* BL21 (DE3) pGro7

The gene of interest for halogenases expression was already ligated into pET-28a (+) plasmid as part of previous work. The plasmid was transformed into *E. coli* BL21 (DE3) pGro7 as previously described using kanamycin and chloramphenicol antibiotics (50 mg/mL and 20 mg/mL) for selection. After the overnight incubation at 37 °C, single colonies from the agar plate were picked and inoculated in autoclaved LB medium with kanamycin and chloramphenicol antibiotics (50 mg/mL and 20 mg/mL); this was incubated overnight at 37 °C with shaking (180 rpm). Under sterile condition in 8 x 2 L conical flasks containing 400 mL LB medium, 4 mL of the starter culture were transferred with kanamycin and chloramphenicol antibiotics (50 mg/mL and 20 mg/mL). The flasks were incubated at 37 °C with shaking (180 rpm) until the OD_{600} reached 0.6. Once the desired OD_{600} was reached, IPTG (0.1 mM final concentration) and L-(+)-arabinose (1 mg/ mL final) were added to induce protein expression and the flasks were further incubated at 20 °C with shaking at 180 rpm overnight. Cells were then harvested by centrifugation (4000 rpm, 20 min, and 4 °C) and the pellets were stored at -20 °C prior to protein purification.

6.8 General method for Ochra-Hl expression in *E. coli* Rosetta 2 (DE3)

The gene of interest for Ochratoxin halogenase was already ligated into pET-28a (+) plasmid. The plasmid was transformed into *E. coli* Rosetta 2 (DE3) as previously described using kanamycin antibiotic (50 mg/mL) for selection. After the overnight incubation at 37 °C, single colonies from the agar plate were picked and inoculated in autoclaved LB medium with kanamycin antibiotic (50 mg/mL), this was incubated overnight at 37 °C with shaking at 180 rpm. Under sterile condition in 8 x 2 L conical flasks containing 400 mL LB medium, 4 mL of the starter culture were transferred with kanamycin antibiotic (50 mg/mL) to allow 1:1000 culture dilution. The flasks were incubated at 37 °C with shaking at180 rpm until the OD_{600} reached 0.6. Once the desired OD_{600} was reached, IPTG (0.1 mM final concentration) was added to induce protein expression and the flasks were further incubated at 20 °C with shaking at 180 rpm overnight. Cells were then harvested by centrifugation (4000 rpm, 20 min, 4 °C) and the pellets were stored at -20 °C prior to protein purification.

6.9 Gateway Cloning

A PCR reaction, as it is explained in 6.22 section, was taken place for the pET-28a plasmid containing the gene of interest.

The PCR reactions were performed as follow.

Table 16: PCR reaction protocol for gateway cloning.

Components	100 µl reaction
5X Q5 Reaction Buffer	20 µl
10 mM dNTPs	8 µl
100 µM Forward Primer	1 µl
100 µM Reverse Primer	1 µl
Template DNA	1 µl
Q5 High-Fidelity DNA Polymerase	1 µl
5X Q5 High GC Enhancer	20 µl
DMSO	5 µl
Nuclease-Free Water	43 µl

The reaction mixture was divided into 4 (25 µl) PCR tubes and transferred to the PCR machine. The thermocycling conditions are as follow.

Table 17: Thermocycling conditions for PCR reaction in getaway cloning

Step	Temperature	Time
Initial Denaturation	98 °C	30 sec
	98 °C	10 sec
30 Cycles	65 °C (annealing)	10 sec
	72 °C (extension)	30 sec
Final extension	72 °C	10 min
Hold	4-10 °C	

The mixture from the PCR reaction was loaded to a DNA gel for agarose electrophoresis. The genes of interest were the extracted from the gel, using the MinElute® gel extraction kit from Quigen. The BP reaction was then performed.

Table 18: BP reaction for getaway cloning

Components	Volume (µl)
Halogenase gene	2
pDONR vector	1
Eb buffer	5
PB clonase 2	2

The mixture was incubated at 25 °C for 1 hour. 1µl of proteinase K was added to the reaction and incubated at 37 °C for 10 min. 5 µl of mixture was transformed and placed onto ampicillin LB plates as already described. Colonies were selected and placed in 10 mL fresh and autoclaved LB with ampicillin and grown overnight at 37 °C. Plasmids were extracted by plasmid miniprep KIT (QIAprep® spin miniprep KIT from Quigen). The LR reaction was then performed as below.

Table 19: LR reaction for getaway cloning

Components	Volume (µl)
pDONR plasmid	1
pDEST17 vector	1
EB buffer	6
LR clonase	2

The mixture was incubated at 25 °C for 1 hour. 1µl of proteinase K was added to the reaction and incubated at 37 °C for 10 min. 5 µl of mixture was transformed and placed onto kanamycin LB plates as already described. Colonies were selected and placed in 10 mL fresh and autoclaved LB with kanamycin and grown overnight at 37 °C. Plasmids were extracted by plasmid miniprep KIT (QIAprep® spin miniprep KIT from Quigen).

6.10 General method for halogenases expression in Vmax™ Express cells

The genes of interest for the halogenases were already ligated into pDEST17 plasmid. The plasmid was transformed through electroporation into Vmax™ competent cells. 1 µl of plasmid was added to the Vmax™ competent cells, and then the mixture was transferred to ice cold cuvettes for electroporation at 0.9 kV for 3.5 to 4.5 ms. The whole mixture was then transferred to 500 µl of enhanced 2xYT media at 30 °C and incubated for 2 hours at 30 °C. All the cells were plated onto LB agar supplemented with ampicillin (12.6 µg/ml) and incubated at 30 °C overnight. The next day, single colonies from the agar plate were picked and

inoculated in 10 ml of autoclaved enhanced 2xYT medium with ampicillin antibiotic (25 µg/mL); this was incubated overnight at 30 °C with shaking (180 rpm). Under sterile condition 500 µL of the starter culture was added to a 50 ml solution of autoclaved enhanced 2xYT medium with ampicillin antibiotic (25 µg/mL) to allow 1:1000 culture dilution. The flask was incubated at 30 °C with shaking (180 rpm) until the OD_{600} reached 0.6. Once the desired OD_{600} was reached, IPTG (1 mM final concentration) was added to induce protein expression and the flasks were further incubated at 30 °C with shaking at180 rpm overnight. Cells were then harvested by centrifugation (4000 rpm, 20 min, 4 °C).

6.11 Protein Purification

Cell pellets were thawed on ice and fully resuspended in 10 mM imidazole KPi buffer. Cells were then lysed via sonication (8 min, 30% pulse, 50% power) and the lysate was centrifuged (10000 rpm, 45 min, 4 °C). All protein purifications were performed using Bio-Rad Econo-Pac® gravity flow chromatography columns where 1 ml Ni-NTA agarose gel was added. Lysate was then loaded onto a Ni-NTA resin column, which had previously been equilibrated using 10 mM imidazole KPi buffer. The Ni-NTA resin was then washed with 3 column volumes (CV) of 60 mM imidazole KPi buffer and the protein of interest was eluted using 6 CVs of 300 mM imidazole KPi buffer, which was then transferred to a vivaspin 20 centricon (30,000 MWCO) for further protein concentration. The protein sample was then subjected to buffer exchange using protein storage buffer and snap frozen in liquid nitrogen at -80 °C. Protein concentration was measured using the ThermoScientific 2000 Nanodrop.

6.12 General method for cross-linked enzyme aggregates formation (CLEA)

ADH enzyme was expressed in *E. coli* BL21 (DE3) and purified by Dr. Sarah Shepherd. The amounts required for ADH and Fre were determined based on a plate reader assay by SS or EC. Cross-linked enzyme aggregates, were produced from 1.5L culture of *E. coli* BL21 (DE3) pGro7. Cell pellets were produced as described above. The cell pellet was resuspended in 100 mM sodium phosphate buffer and lysed via sonication (8 min, 30% pulse and 50% power), the lysate was then clarified via centrifugation (60 min, 10000 rpm, 4 °C) and the supernatant transferred to a clean 50 mL falcon tube. Purified Fre (2.5 U/mL) and ADH (1 U/mL) were added to the lysate and mixed by inversion before protein precipitation by the addition of ammonium sulfate (16.2 g) with mixing at 4 °C for 3 hours. To this was added 50% v/v aqueous glutaraldehyde solution and the mixture was left for a further 1 to 2 hours at 4°C. The protein

aggregate was then collected via centrifugation (20 min, 10,000 rpm, 4°C) and washed thrice with sodium phosphate buffer (100 mM), the resulting cross linked enzyme aggregate was stored at 4 ˚C prior to use.

6.13 Assay conditions with purified WT halogenases

The substrate (0.5 mM), halogen salt ($MgCl_2$ or NaBr – 50 mM), FAD (1 µM), NADH (5 mM), purified Fre (1 µM) and purified halogenase enzyme (20 mM) were added to 10 mM KPi buffer to a reaction volume of 100 µL. If required, 1 µM superoxide dismutase and 5000 U/mL catalase (1 µL) were also included. The mixture was then shaken at a desired temperature at 750 rpm in a thermomixer. Assays were quenched by boiling at 98 °C for 10 min prior to centrifugation (14000 rpm, 5 mins) and the supernatant was analysed by HPLC.

6.14 General method for analytical scale biohalogenation with pure enzymes (ADH recycling system)

The substrate (0.5 mM), halogen salt ($MgCl_2$ or NaBr – 50 mM), FAD (1 µM), NADH (5 mM), purified Fre (1 µM), isopropyl alcohol (2.5% of final volume), purified ADH (6 µM) and purified halogenase enzyme (20 mM) were added to 10 mM KPi buffer to a reaction volume of 100 µL The mixture was then shaken at a desired temperature and 750 rpm in a thermomixer. Enzymes were precipitated by boiling at 98 °C for 10 min prior to centrifugation (14000 rpm, 5 mins) and the supernatant was analysed by HPLC.

6.15 General method for preparative scale biohalogenation with cross-linked enzyme aggregates

The substrate (3 mM), halogen salt ($MgCl_2$ 30 mM), FAD (10 µM), NADH (0.1 mM) and isopropanol (5% v/v) were added to 15 mM sodium phosphate buffer in a reaction volume of 30 mL. The mixture was shaken at room temperature and 180 rpm overnight. A sample of the reaction was taken and centrifuged at 14000 rpm for 10 min and the supernatant was analysed by HPLC. The biohalogenation was concentrated in vacuo and purification was done via C18 BondElut chromatography with the column being activated with 3 CV water, 3 CV ACN and 3 CV water. Product was eluted using a solvent mix of H_2O: ACN 50:50 to yield an off-white solid.

6.16 Whole cells reaction in Vmax™ cells

Following the expression of the halogenases in Vmax™ cells as described in section 1.9; the whole cells were collected and resuspended in tris-buffer. The resuspended cells were transferred into a dialysis membrane in a conical flask and 50 ml of tris buffer and 1 mM of tryptophan were added. The flask was incubated at 30 °C overnight. The next day, the membrane was removed and a sample from the supernatant was spun down and analysed via HPLC.

6.17 Lysate reaction in Vmax™ cells

Following the expression of the halogenases in Vmax™ cells as described in section 1.9 ; the whole cells were collected, resuspended in tris-buffer and sonicated (8 min, 30% pulse, 50% power) and the lysate was centrifuged (10000 rpm, 45 min, 4 °C). The lysate was transferred into a dialysis membrane and transferred in a conical flask with 50 ml of tris buffer, 1 mM of tryptophan and the necessary cofactors, FAD (1 µl), $MgCl_2$ (100 mM) and NADH (2.5 mM). The flask was incubated at 30 °C overnight. The next day, the membrane was removed and a sample from the supernatant was spun down and analysed via HPLC.

6.18 HPLC Method

50 µL of supernatant from the 100 µL enzyme assays were injected and separated using a Phenomenex Kinetex C18 column (5 µm, 150 x 4.6 mm). Two solvents were used to generate a gradient elution in which solvent A = water + 0.05% TFA and solvent B = ACN + 0.05% TFA. Flow rate was 1.5 mL/min and the column temperature was maintained at 40 °C throughout. UV absorbance was detected at 280 nm. The gradients established for the analysis of biotransformations are tabulated below:

Table 20: HPLC gradient conditions for biotransformation analysis.

Start time (min)	End time (min)	Initial solvent ratio (A: B)	Final solvent ratio (A:B)
0.01	1.50	95:5	95:5
1.50	6.00	95:5	25:75
6.00	6.10	25:75	5:95
6.10	7.50	5:95	5:95
7.50	7.60	5:95	95:5
7.60	8.70	95:5	95:5

6.19 Restriction Digest

The synthetic genes were purchased from Genewiz®. The reaction mixture contained 3-5 µg of DNA, 1 µL of each restriction enzyme, 5 µL CutSmart (10 x) up to 50 µL with nuclease free (nf) water. The restrictions enzymes were added last. The reaction mixture was incubated at 37 °C for 4 to 5 hours.

10 µL of DNA dye was added and DNA gel electrophoresis was carried out.

6.20 DNA gel – Agarose electrophoresis

For the preparation of one gel, 0.5 g of agarose was dissolved in 50 mL of TAE (a buffer solution containing Tris base, acetic acid and EDTA) with a use of a microwave. Once agarose was dissolved, 3 µL of ethidium bromide (EB) was added and the solution was poured into a cast allowing the mixture to get solidified. A comb is used to create wells for loading sample when the mixture is still in liquid form.

The DNA gel was run for 40 min at 95 Volt.

6.21 Plasmid ligation

The Quick Ligation™ Kit was purchased from New England Biolabs. According to the supplier's instructions, 10 µL Quick Ligase Reaction Buffer (2X), 50 ng of the DNA vector, 37.5 ng of the insert DNA, 1 µL Quick ligase were mixed with nuclease-free water up to 20 µL reaction volume. The Quick ligase was added last. The reaction mixture was gently mixed by pipetting up and down, and then was incubated at room temperature for about 30 min. The mixture was placed on ice and 3 µL of the reaction mixture was added to 50 µL competent cells DH5α as described above.

6.22 Polymerase Chain Reaction (PCR)

The PCR kit was purchased from New England Biolabs. According to the supplier's instructions, 10 µl 5X Q5 Reaction Buffer, 20 µl 5X Q5 High GC Enhancer, 2 µl 10 mM dNTPs, 1 µl 10 µM Forward Primer and 1 µl 10 µM Reverse Primer, 1 µl template DNA which, 1 µl Q5 High-Fidelity DNA Polymerase, and 5 µl DMSO, were added to NF water in a final reaction volume of 100 µl. The mixture was mixed well but gently by pipetting up and down and the

split in 4 PCR tubes. The PCR tubes were transferred to the PCR machine. The thermocycling conditions for a routine PCR are presented to the table below.

Table 21: Thermocycling conditions for a routine PCR reaction

Step	Temperature	Time
Initial Denaturation	98 °C	30 sec
	98 °C	5-10 sec
25-30 Cycles	50-72 °C (annealing)	10-30 sec
	72 °C (extension)	20-30 sec/kb
Final extension	72 °C	2 min
Hold	4-10 °C	

6.23 Colony PCR

Colony PCR is a convenient high-throughput method for determining the presence or absence of insert DNA in plasmid constructs. The template DNA that is used for the colony PCR is the actual colony that is appeared after the ligation with the vector.

Table 22: Colony PCR reaction

Component	25 µl reaction
5X Q5 Reaction Buffer	5 µl
10 mM dNTPs	0.5 µl
25 µM Forward Primer	0.5 µl
25 µM Reverse Primer	0.5 µl
Template DNA	colony
Q5 High-Fidelity DNA Polymerase	0.25 µl
5X Q5 High GC Enhancer	5 µl
Nuclease-Free Water	Up to 25 µl

The reaction mixture was mixed well but gently by pipetting up and down. The PCR tube was transferred to the PCR machine. The thermocycling conditions for the colony PCR are presented to the table below.

Table 23: Thermocycling conditions for a colony PCR reaction

Step	Temperature	Time
Initial Denaturation	98 °C	30 sec
	98 °C	10 sec
25-30 Cycles	65 °C (annealing)	20 sec
	68 °C (extension)	50 sec
Final extension	72 °C	2 min
Hold	4-10 °C	

6.24 PCR reaction HiFi Assembly

PCR reaction is performed according to the following protocol.

Table 24: PCR Reaction for HiFi assembly

Component	100 µl reaction
HiFi polymerase	50 µl
25mM Forward primer	2 µl
25 mM Reverse primer	2 µl
Template DNA	45 µl
Nuclease-Free Water	Up to 100 µl

The reaction mixture was mixed well but gently by pipetting up and down. The PCR tube was transferred to the PCR machine. The thermocycling conditions for the colony PCR are presented to the Table 21. After the gene extraction, HiFi DNA Assembly was performed. The reaction was set up on ice.

Table 25: HiFi DNA Assembly Protocol

Component	20.5 µl reaction
Vector	0.5 µl
Template DNA	10 µl
HiFi DNA Assembly Master	10 µl

The sample was incubated at 50 °C for 30 minutes Following incubation, the sample was placed on ice. *E. coli* DH5α competent cells were transformed with 5 µl of the assembled product.

7 Appendix

Protein sequences

Ochra P450

RISPREIHIHDPNYYETLYATNSPRNKDPWFTAHFGVDESAFSTLDYRLHRSRRAMTIAPFFAKARMTDG
AQPLIKANLAKLIRHLDAHASSQSVLKVEVAYNSFTGDVITGYTSYRSFEYLETTDMTVPWSETVRNLVES
GMTLSRHLPGFFPLLAWAGSRCIAAVYPKLLPVIAFRMKCAQEVNFMTWTNSEEGKKEAIQSGCSEPAL
FPELVSRASSAPDITEERLLHEFITIVAAGTETTAHTMTTVCTFHIAHNQSILDKLREELDNRFHSNADMTD
LQTLEQLPYLTGIIYEGLRLSYGLSHRLQRISPIDPLKYKDIAIPPNTPVGMTSSALMTHHDESIFPQSHEFIP
ERWTDPNDRRRLNKYMTVSFSKGSRQCIGMNLAFAELYLGLFRRYDMTKLHDTTLDDVQLHGDKMLP
RAKNGSKGVRVTLRRAQSVG

Ochratoxin Halogenase

MEIPHKATVLVIGGGPGGSYTASALAREGIDIVLLEADVFPRFIDLDETFVNYGFVRKNASHGSLADTDFIL
QPGADTFAWNVVRSECDDLMFKHASNSGARAFDGVKVTAIEFDPLDESAIDDHPGRPVSASWKAKDG
RTGSISFDYLVDASGRAGIASTKYLKSRTYNSYLKNVASWGYWRGATPYGVGTSVEGQPYFEALQDGSG
WVWFIPLHNGTTSVGVVMNQELATQKKKSSTVTSSRAFYLESVEGARVISQLLQPANLDGEIKQASDWS
YNASSYGSPYLRIVGDAGAFIDPYFSSGVHLAVSGGLSAAVSIAASIRGDCPEHTAWQWHSQGVANRYG
RFLLVVLGATKQIRARDSPVLNSEGQDGFDDAFAVIRPVIQGTADVQGKVSAREVLDAVTFSTNAVRPSA
GGQNVVLEESSRSLRSQVEQEMGDVANSLAKAYKDTDVYEGLMARLERGSLGLKAVGVMG

PyrH

MERRKRERLGSLGRPTKKELRMIRSVVIVGGGTAGWMTASYLKAAFDDRIDVTLVESGNVRRIGVGEAT
FSTVRHFFDYLGLDEREWLPRCAGGYKLGIRFENWSEPGEYFYHPFERLRVVDGFNMAEWWLAVGDRR
TSFSEACYLTHRLCEAKRAPRMLDGSLFASQVDESLGRSTLAEQRAQFPYAYHFDADEVARYLSEYAIARG
VRHVVDDVQHVGQDERGWISGVHTKQHGEISGDLFVDCTGFRGLLINQTLGGRFQSFSDVLPNNRAVA
LRVPRENDEDMRPYTTATAMSAGWMWTIPLFKRDGNGYVYSDEFISPEEAERELRSTVAPGRDDLEAN
HIQMRIGRNERTCHROMO-
HALCVAVGLSAAFVEPLESTGIFFIQHAIEQLVKHFPGERWDPVLISAYNERMAHMVDGVKEFLVLHYKG
AQREDTPYWKAAKTRAMPDGLARKLELSASHLLDEQTIYPYYHGFETYSWITMNLGLGIVPERPRPALLH
MDPAPALAEFERLRREGDELIAALPSCYEYLASIQ

Number of amino acids: 532

Molecular weight: 60726.71

Theoretical pI: 5.87

PrnA

MNKPIKNIVIVGGGTAGWMAASYLVRALQQQANITLIESAAIPRIGVGEATIPSLQKVFFDFLGIPEREW
MPQVNGAFKAAIKFVNWRKSPDPSRDDHFYHLFGNVPNCDGVPLTHYWLRKREQGFQQPMEYACYP
QPGALDGKLAPCLSDGTRQMSHAWHFDAHLVADFLKRWAVERGVNRVVDEVVDVRLNNRGYISNLLT
KEGRTLEADLFIDCSGMRGLLINQALKEPFIDMSDYLLCDSAVASAVPNDDARDGVEPYTSSIAMNSGW

TWKIPMLGRFGSGYVFSSHFTSRDQATADFLKLWGLSDNQPLNQIKFRVGRNKRAWVNNCVSIGLSSCF
LEPLESTGIYFIYAALYQLVKHFPDTSFDPRLSDAFNAEIVHMFDDCRDFVQAHYFTTSRDDTPFWLANRH
DLRLSDAIKEKVQRYKAGLPLTTTSFDDSTYYETFDYEFKNFWLNGNYYCIFAGLGMLPDRSLPLLQHRPE
SIEKAEAMFASIRREAERLRTSLPTNYDYLRSLRDGDAGLSRGQRGPKLAAQESL

Number of amino acids: 538

Molecular weight: 61075.28

Theoretical pI: 6.33

SttH (B)

MNTRNPDKVVIVGGGTAGWMTASYLKKAFGERVSVTLVESGTIGTVGVGEATFSDIRHFFEFLDLREEE
WMPACNATYKLAVRFQDWQRPGHHFYHPFEQMRSVDGFPLTDWWLQNGPTDRFDRDCFVMASLC
DAGRSPRYLNGSLLQQEFDERAEEPAGLTMSEHQGKTQFPYAYHFEAALLAEFLSGYSKDRGVKHVVDE
VLEVKLDDRGWISHVVTKEHGDIGGDLFVDCTGFRGVLLNQALGVPFVSYQDTLPNDSAVALQVPLDM
EARGIPPYTRATAKEAGWIWTIPLIGRIGTGYVYAKDYCSPEEAERTLREFVGPEAADVEANHIRMRIGRS
EQSWKNNCVAIGLSSGFVEPLESTGIFFIHHAIEQLVKHFPAGDWHPQLRAGYNSAVANVMDGVREFLV
LHYLGAARNDTRPWKDTKTRAVPDALAERIERWKVQLPDSENVFPYYHGLPPYSYMAILLGTGAIGLRPS
PALALADPAAAEKEFTAIRDRARFLVDTLPSQYEYFAAMGQRV

Number of amino acids: 523

Molecular weight: 58720.34

Theoretical pI: 5.45

Chromo-Hal

MLDKVVIVGGGTSGWMSASYLKAAFGDRISVTVVESDRVPTIGVGEATFSTVRHFFEFLGLKETDWMPA
CNASYKLGIRFQDWSAPGTHFYHPFERLRVVDGFPLTDWWLKLRPSDRFDADNFVMAAMGDSMRSP
RYLDGTIFERSENDLGPGSAAFRSTLSEQDTQFPYAYHFDAALLAKYLTEYGVSRGVRHVVDDVVDVRLD
DRGWIKAVATRDHGDLEGDLFVDCTGFKGLLVNEALEEPFVSYQDTLPNDRAVALRVPVDTAEVGMPP
YTRATAQEAGWIWTIPLFDRVGTGYVYASDYCGPEEAERTLRAFVGPAAEGMEANHIRMRIGRSRNSW
TKNCVAIGLASGFVEPLESTGIFFIQHGIEQLVKHFPQQDWSPRLRDSYNRAVNNCLDGVREFLVLHFKG
AARDDNAYWRDAKTRAVPDGLAARLEQWQHKTPDAENIFPHYHGFEAYSYVAMLMGLQGIPLRAPG
ALDLMDPTAAAEEFARVRATAAELVRSLPGQYEYLAQMR

Number of amino acids: 516

Molecular weight: 57849.23

Theoretical pI: 5.28

RebH

MNNRLQKITIVGGGTAGWMTALLLEMVFSRTSATKDRPRICLIESPNIATVGVGEATVPRMPVTLRQAG
ISERDFFRETNASFKLGVKFCNWNKDAKGKRIDYVNPFAHGQMLEGLEAAEYFLRFGNGDRDFTQSISH
DDLARLCKGARLLGQPEFEQRFGYAYHLDAVKFAGMLTRVCTQRGVEHIRDEVQSVELNEQGNVSHLM
LERAGRHDIEMVVDCTGFRGLIINQALGEPFMDYSDYLPNDRAMALQIEHPNPRKIESVTRSTALGAGW
TWRVPLYNRVGTGYVYSSAHRTDDQAADEYLEWLGDSGKGATPRVIPMRIGRVRNAWVKNCVAIGLA
GGFIEPLESTAIHMVDHAVRWFAEHLPTRDIEPSLRDRYNRQMNKLYDEVLDFICLHYKLNNRTDDQYW
IDARTEMKVPDRLAENLELWKNRLPMQHDVEFATLFDHRVYQTVLLGKQVYDTGYGTGIRDRIRPLRKD
IWFQWWKGAKVDLAQILKAMPDHKTLLRDIRGELKQPTFAMAKAAQPTVPMPGAAPEPVVIQNMPN
LAEIQSGKKDLQLF

Number of amino acids: 561

Molecular weight: 63922.15

Theoretical pI: 8.06

Fre

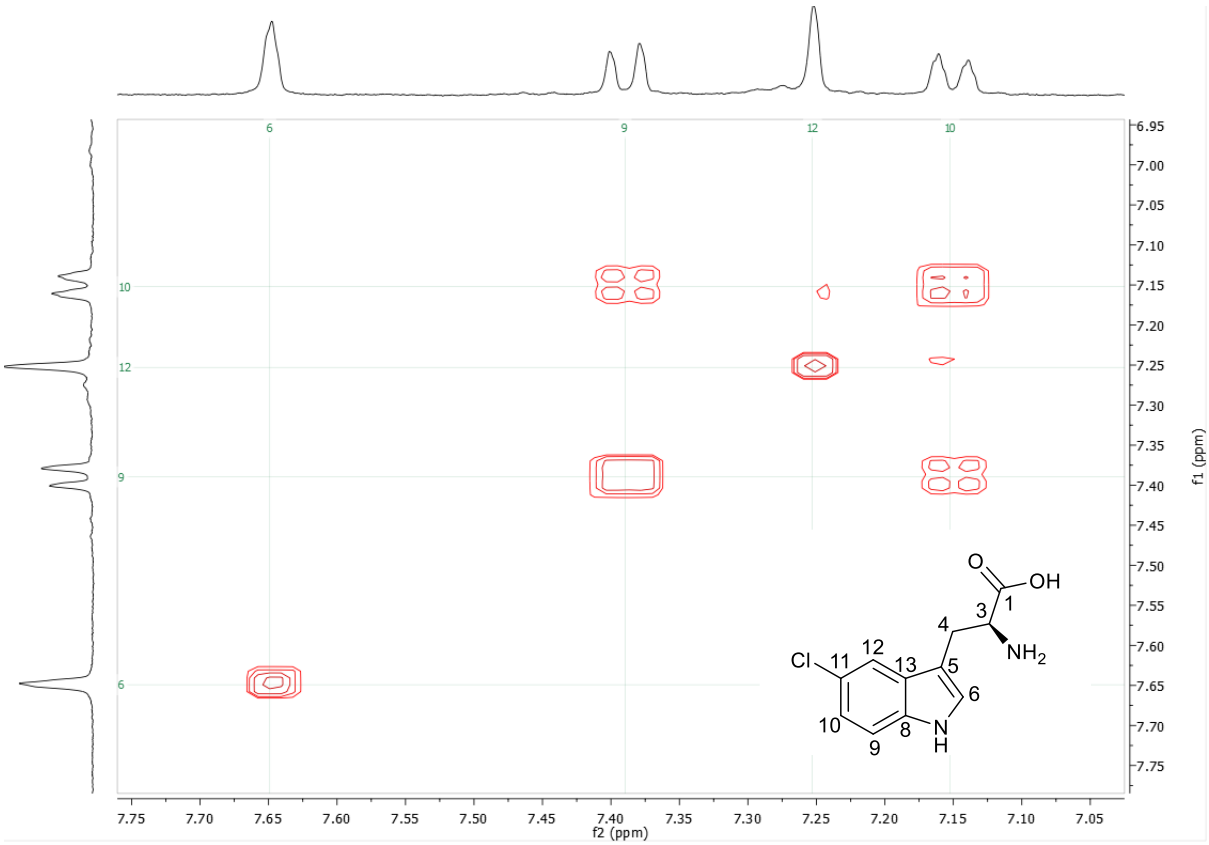
MTTLSCKVTSVEAITDTVYRVRIVPDAAFSFRAGQYLMVVMDERDKRPFSMASTPDEKGFIELHIGASEIN
LYAKAVMDRILKDHQIVVDIPHGEAWLRDDEERPMILIAGGTGFSYARSILLTALARNPNRDITIYWGGRE
EQHLYDLCELEALSLKHPGLQVVPVVEQPEAGWRGRTGTVLTAVLQDHGTLAEHDIYIAGRFEMAKIAR
DLFCSERNAREDRLFGDAFAFI

Number of amino acids: 233

Molecular weight: 26241.99

Theoretical pI: 5.29

NMR



Primers

Primers for Ochratoxin Hal and P450 (Hi-Fi primers)

Forward Ochra 5’GCCTGGTGCCGCGCGGCAGCCATATGGAAATCCCGCACAAAG’3

Reverse Ochra 5’CAGTGGTGGTGGTGGTGGTGCTCGAGTTAACCCATCACGC’3

Forward P450 5’ GCCTGGTGCCGCGCGGCAGCCATATGAGCATTGATACTTTACTG ‘3

Reverse P450 5’ CAGTGGTGGTGGTGGTGGTGCTCGAGTTAACCAACACTCTG ‘3

Primers for gateway cloning

Forward PyrH 5’GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGATTCGTAGCGTTGTTATTGT’3

Reverse PyrH 5’GGGGACCACTTTGTACAAGAAAGCTGGGTTTACTGAATGCTTGCCAGATATTC’3

Forward SttH 5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAATACCCGCAATCCGGATA ‘3

Reverse SttH 5’ GGGGACCACTTTGTACAAGAAAGCTGGGTT TACACACGCTGACCCATTG ‘3

Forward RebH 5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTCCGGCAAGATTGAC’3

Reverse RebH 5 GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGCGGCCGTG’3

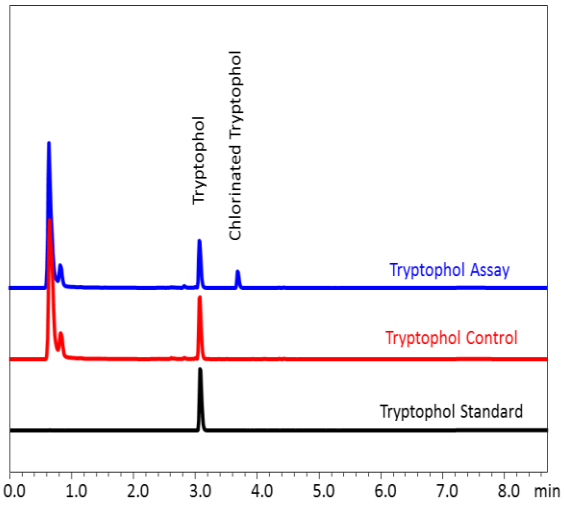
Forward PrnA 5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAACAAGCCGATCAAG’3

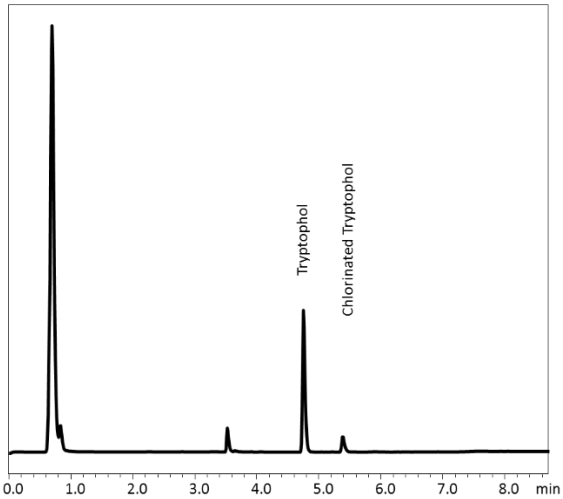
Reverse PrnA 5’ GGGGACCACTTTGTACAAGAAAGCTGGGTCTACAGGCTTTCCTCG’3

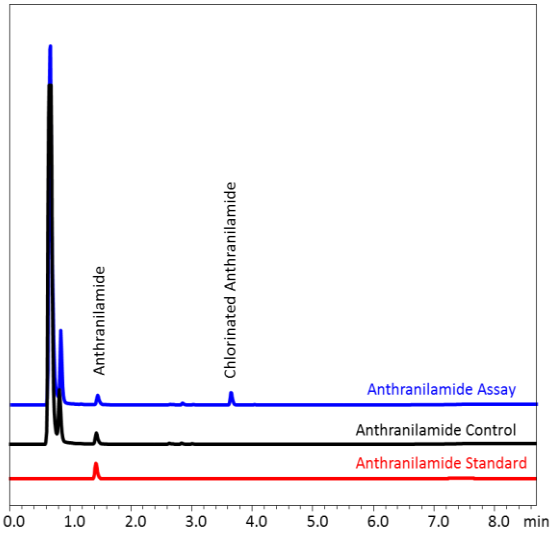
Primers for the truncated P450

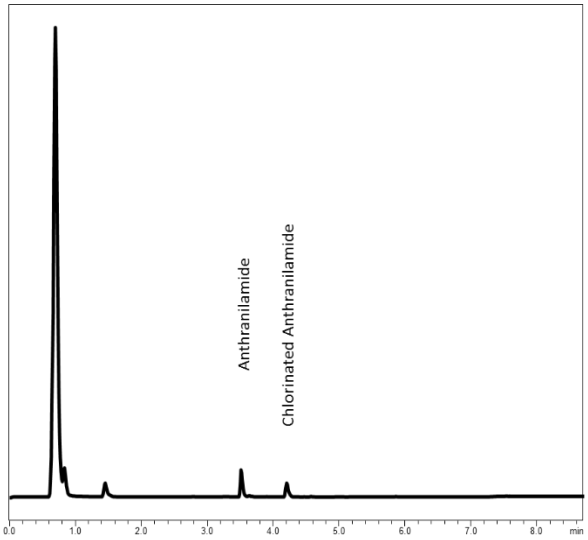
Primer A forward	5’ AAAAAACATATGAGCATTGATACTTTACTGTATACCAGCT ‘3
Primer B reverse	5’ AAAAAACTCGAGACCAACACTCTGTGCGCGAC ‘3
Primer C forward	5’ AAAAAACATATGTGGAGCCCGCTGGCCCATAT’ 3
Primer D reverse	5’AAAAAACTCGAGTTAACCAACACTCTGTGCGCGACGTAAG’3

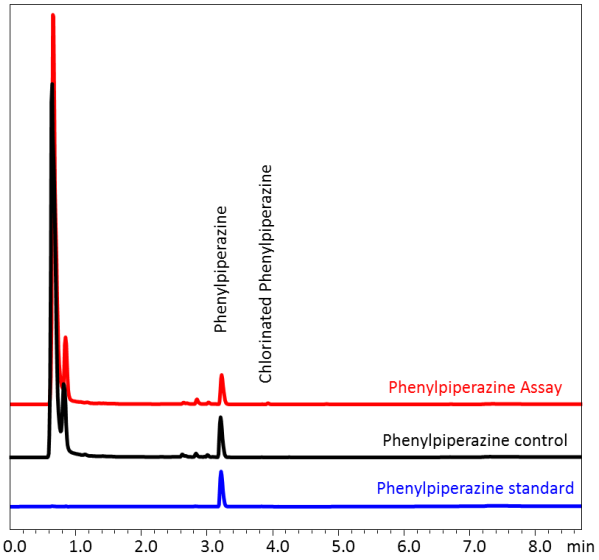
Additional HPLC

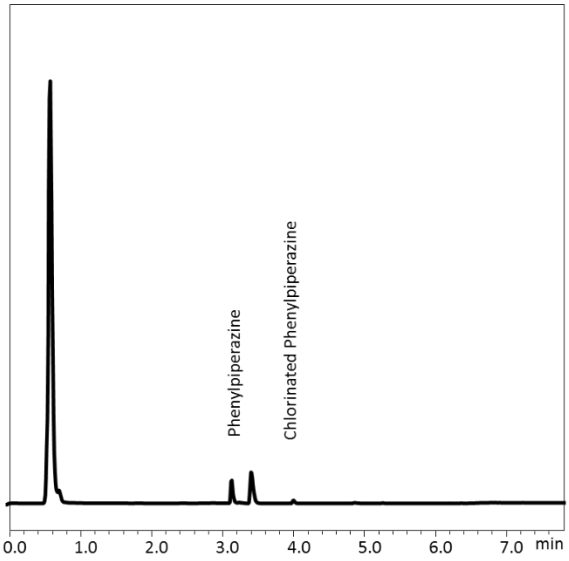


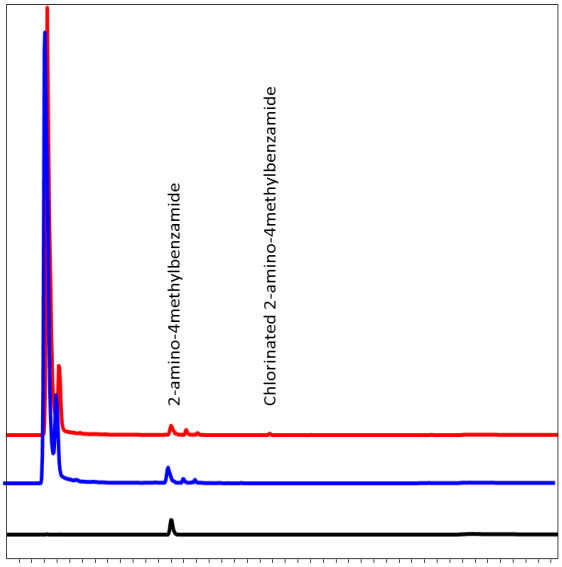


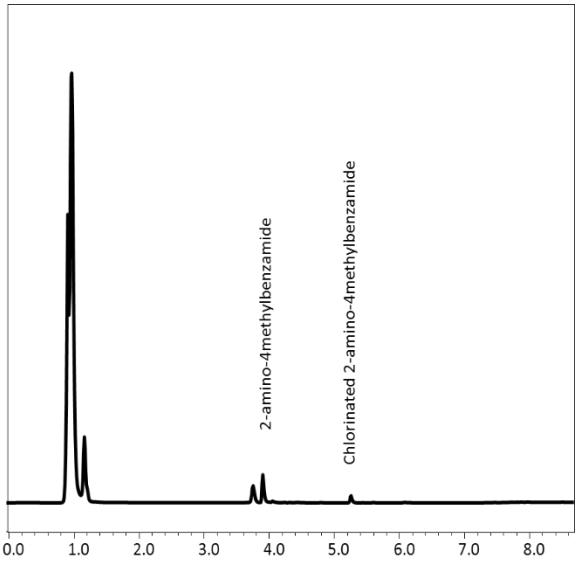









2-amino-4methylbenzamide
Chlorinated 2-amino-4methylbenzamide


2-amino-4methylbenzamide
Chlorinated 2-amino-4methylbenzamide
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0

8 References

1. 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**, 3454-3460.
2. J. Latham, E. Brandenburger, S. A. Shepherd, B. R. K. Menon and J. Micklefield, *Chem Rev*, 2017, **118**, 232-269.
3. G. W. Gribble, *Journal of Chemical Education*, 2004, **81**, 1441.
4. F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova and C. T. Walsh, *Chem Rev*, 2006, **106**, 3364-3378.
5. M. Frese and N. Sewald, *Angew Chem Int Ed Engl*, 2015, **54**, 298-301.
6. K. Tanemura, T. Suzuki, Y. Nishida, K. Satsumabayashi and T. Horaguchi, *Chemistry Letters*, 2003, **32**, 932-933.
7. L. Voskressensky, N. Golantsov and A. Maharramov, *Synthesis*, 2016, **48**, 615-643.
8. D. R. Morris and L. P. Hager, *J Biol Chem*, 1966, **241**, 1763-1768.
9. A. Taurog, E. M. Howells and H. I. Nachimson, *Journal of Biological Chemistry*, 1966, **241**, 4686-+.
10. J. R. B. a. L. P. Hageri, *Journal Of Biological Chemistry*, 1963, **238**, 3091-3094.
11. M. Sundaramoorthy, J. Terner and T. L. Poulos, *Chem Biol*, 1998, **5**, 461-473.
12. A. Taurog and E. M. Howells, *J Biol Chem*, 1966, **241**, 1329-1339.
13. R. Vazquez-Duhalt, M. Ayala and F. J. Marquez-Rocha, *Phytochemistry*, 2001, **58**, 929-933.
14. E. E. Coupe, M. G. Smyth, A. P. Fosberry, R. M. Hall and J. A. Littlechild, *Protein Expr Purif*, 2007, **52**, 265-272.
15. S. Dai, E. Vera and J. H. McNeill, *Pharmacology & Toxicology*, 1995, **76**, 263-268.
16. A. Messerschmidt, L. Prade and R. Wever, *Biol Chem*, 1997, **378**, 309-315.
17. N. Itoh, H. Sasaki, N. Ohsawa, M. S. Shibata and J. i. Miura, *Phytochemistry*, 1996, **42**, 277-281.
18. V. Weichold, D. Milbredt and K. H. van Pee, *Angew Chem Int Ed Engl*, 2016, **55**, 6374-6389.
19. F. H. Vaillancourt, J. Yin and C. T. Walsh, *Proc Natl Acad Sci U S A*, 2005, **102**, 10111-10116.
20. V. Kohler and N. J. Turner, *Chem Commun (Camb)*, 2015, **51**, 450-464.
21. F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. E. O'Connor and C. T. Walsh, *Nature*, 2005, **436**, 1191-1194.
22. L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh and C. L. Drennan, *Nature*, 2006, **440**, 368-371.
23. J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger, Jr. and C. Krebs, *Biochemistry*, 2003, **42**, 7497-7508.
24. M. L. Hillwig and X. Liu, *Nat Chem Biol*, 2014, **10**, 921-923.
25. A. J. Mitchell, Q. Zhu, A. O. Maggiolo, N. R. Ananth, M. L. Hillwig, X. Liu and A. K. Boal, *Nat Chem Biol*, 2016, **12**, 636-640.
26. A. J. Mitchell, N. P. Dunham, J. A. Bergman, B. Wang, Q. Zhu, W. C. Chang, X. Liu and A. K. Boal, *Biochemistry*, 2017, **56**, 441-444.
27. D. Holtmann and F. Hollmann, *Chembiochem*, 2016, **17**, 1391-1398.
28. M. M. Huijbers, S. Montersino, A. H. Westphal, D. Tischler and W. J. van Berkel, *Arch Biochem Biophys*, 2014, **544**, 2-17.

29. S. Visitsatthawong, P. Chenprakhon, P. Chaiyen and P. Surawatanawong, *J Am Chem Soc*, 2015, **137**, 9363-9374.
30. S. Unversucht, F. Hollmann, A. Schmid and K.-H. van Pée, *Advanced Synthesis & Catalysis*, 2005, **347**, 1163-1167.
31. E. Yeh, S. Garneau and C. T. Walsh, *Proc Natl Acad Sci U S A*, 2005, **102**, 3960-3965.
32. S. Keller, T. Wage, K. Hohaus, M. Holzer, E. Eichhorn and K. H. van Pee, *Angew Chem Int Ed Engl*, 2000, **39**, 2300-2302.
33. X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pee and J. H. Naismith, *J Mol Biol*, 2009, **391**, 74-85.
34. C. Dong, A. Kotzsch, M. Dorward, K. H. van Pee and J. H. Naismith, *Acta Crystallogr D Biol Crystallogr*, 2004, **60**, 1438-1440.
35. E. Yeh, L. J. Cole, E. W. Barr, J. M. Bollinger, Jr., D. P. Ballou and C. T. Walsh, *Biochemistry*, 2006, **45**, 7904-7912.
36. E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan and C. T. Walsh, *Biochemistry*, 2007, **46**, 1284-1292.
37. C. Dong, S. Flecks, S. Unversucht, C. Haupt, K. H. van Pee and J. H. Naismith, *Science*, 2005, **309**, 2216-2219.
38. M. L. Mascotti, M. Juri Ayub, N. Furnham, J. M. Thornton and R. A. Laskowski, *J Mol Biol*, 2016, **428**, 3131-3146.
39. D. R. Smith, S. Gruschow and R. J. Goss, *Curr Opin Chem Biol*, 2013, **17**, 276-283.
40. Y. L. Du and K. S. Ryan, *ACS Synth Biol*, 2015, **4**, 682-688.
41. S. Zehner, A. Kotzsch, B. Bister, R. D. Sussmuth, C. Mendez, J. A. Salas and K. H. van Pee, *Chemistry & Biology*, 2005, **12**, 445-452.
42. J. Zeng and J. Zhan, *Biotechnol Lett*, 2011, **33**, 1607-1613.
43. S. A. Shepherd, B. R. Menon, H. Fisk, A. W. Struck, C. Levy, D. Leys and J. Micklefield, *Chembiochem*, 2016, **17**, 821-824.
44. M. E. Welsch, S. A. Snyder and B. R. Stockwell, *Curr Opin Chem Biol*, 2010, **14**, 347-361.
45. M. Frese, P. H. Guzowska, H. Voss and N. Sewald, *Chemcatchem*, 2014, **6**, 1270-1276.
46. B. R. K. Menon, E. Brandenburger, H. H. Sharif, U. Klemstein, S. A. Shepherd, M. F. Greaney and J. Micklefield, *Angew Chem Int Ed Engl*, 2017, **56**, 11841-11845.
47. J. T. Payne, M. C. Andorfer and J. C. Lewis, *Angew Chem Int Edit*, 2013, **52**, 5271-5274.
48. A. Lang, S. Polnick, T. Nicke, P. William, E. P. Patallo, J. H. Naismith and K. H. van Pee, *Angew Chem Int Ed Engl*, 2011, **50**, 2951-2953.
49. M. C. Andorfer, H. J. Park, J. Vergara-Coll and J. C. Lewis, *Chem Sci*, 2016, **7**, 3720-3729.
50. A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik and R. Milo, *Biochemistry*, 2011, **50**, 4402-4410.
51. J. Q. Yu and Z. Shi, *Top Curr Chem*, 2010, **292**, xi-xiii.
52. B. R. Menon, J. Latham, M. S. Dunstan, E. Brandenburger, U. Klemstein, D. Leys, C. Karthikeyan, M. F. Greaney, S. A. Shepherd and J. Micklefield, *Org Biomol Chem*, 2016, **14**, 9354-9361.
53. 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**, 2761-2766.
54. C. Schnepel and N. Sewald, *Chemistry*, 2017, **23**, 12064-12086.
55. K. C. Nicolaou, P. G. Bulger and D. Sarlah, *Angew Chem Int Ed Engl*, 2005, **44**, 4442-4489.
56. A. Suzuki, *Angew Chem Int Ed Engl*, 2011, **50**, 6722-6737.

57. A. Deb Roy, S. Gruschow, N. Cairns and R. J. Goss, *J Am Chem Soc*, 2010, **132**, 12243-12245.
58. K. J. van der Merwe, P. S. Steyn, L. Fourie, D. B. Scott and J. J. Theron, *Nature*, 1965, **205**, 1112-1113.
59. A. Chakrabortti, J. Li and Z. X. Liang, *Genome Announc*, 2016, **4**.
60. M. Ferrara, G. Perrone, L. Gambacorta, F. Epifani, M. Solfrizzo and A. Gallo, *Appl Environ Microbiol*, 2016, **82**, 5631-5641.
61. E. Petzinger and K. Ziegler, *J Vet Pharmacol Ther*, 2000, **23**, 91-98.
62. A. Pfohl-Leszkowicz and R. A. Manderville, *Mol Nutr Food Res*, 2007, **51**, 61-99.
63. A. Gallo, K. S. Bruno, M. Solfrizzo, G. Perrone, G. Mule, A. Visconti and S. E. Baker, *Appl Environ Microb*, 2012, **78**, 8208-8218.
64. R. Geisen and M. Schmidt-Heydt, *Mycota*, 2009, **15**, 353-376.
65. A. Gallo, G. Perrone, M. Solfrizzo, F. Epifani, A. Abbas, A. D. W. Dobson and G. Mule, *Int J Food Microbiol*, 2009, **129**, 8-15.
66. J. O'Callaghan, P. C. Stapleton and A. D. Dobson, *Fungal Genet Biol*, 2006, **43**, 213-221.
67. J. Gil-Serna, C. Vazquez, M. T. Gonzalez-Jaen and B. Patino, *Int J Food Microbiol*, 2015, **214**, 102-108.
68. D. Hoffmeister and N. P. Keller, *Nat Prod Rep*, 2007, **24**, 393-416.
69. A. Gallo, B. P. Knox, K. S. Bruno, M. Solfrizzo, S. E. Baker and G. Perrone, *Int J Food Microbiol*, 2014, **179**, 10-17.
70. R. Geisen, M. Schmidt-Heydt and A. Karolewiez, *Mycotoxin Res*, 2006, **22**, 134-141.
71. H. J. Pel, J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann, P. J. Schaap, G. Turner, R. P. de Vries, R. Albang, K. Albermann, M. R. Andersen, J. D. Bendtsen, J. A. Benen, M. van den Berg, S. Breestraat, M. X. Caddick, R. Contreras, M. Cornell, P. M. Coutinho, E. G. Danchin, A. J. Debets, P. Dekker, P. W. van Dijck, A. van Dijk, L. Dijkhuizen, A. J. Driessen, C. d'Enfert, S. Geysens, C. Goosen, G. S. Groot, P. W. de Groot, T. Guillemette, B. Henrissat, M. Herweijer, J. P. van den Hombergh, C. A. van den Hondel, R. T. van der Heijden, R. M. van der Kaaij, F. M. Klis, H. J. Kools, C. P. Kubicek, P. A. van Kuyk, J. Lauber, X. Lu, M. J. van der Maarel, R. Meulenberg, H. Menke, M. A. Mortimer, J. Nielsen, S. G. Oliver, M. Olsthoorn, K. Pal, N. N. van Peij, A. F. Ram, U. Rinas, J. A. Roubos, C. M. Sagt, M. Schmoll, J. Sun, D. Ussery, J. Varga, W. Vervecken, P. J. van de Vondervoort, H. Wedler, H. A. Wosten, A. P. Zeng, A. J. van Ooyen, J. Visser and H. Stam, *Nat Biotechnol*, 2007, **25**, 221-231.
72. L. M. Ferracin, C. B. Fier, M. L. Vieira, C. B. Monteiro-Vitorello, M. Varani Ade, M. M. Rossi, M. Muller-Santos, M. H. Taniwaki, B. Thie Iamanaka and M. H. Fungaro, *Int J Food Microbiol*, 2012, **155**, 137-145.
73. A. Gallo, M. Ferrara and G. Perrone, *Toxins (Basel)*, 2013, **5**, 717-742.
74. N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer and R. Y. Tsien, *Nat Biotechnol*, 2004, **22**, 1567-1572.
75. H. Bedouelle and P. Duplay, *Eur J Biochem*, 1988, **171**, 541-549.
76. P. Rondard, F. Bregegere, A. Lecroisey, M. Delepierre and H. Bedouelle, *Biochemistry*, 1997, **36**, 8954-8961.
77. D. R. M. Smith, A. R. Uria, E. J. N. Helfrich, D. Milbredt, K. H. van Pee, J. Piel and R. J. M. Goss, *ACS Chem Biol*, 2017, **12**, 1281-1287.
78. P. Chelikani, I. Fita and P. C. Loewen, *Cell Mol Life Sci*, 2004, **61**, 192-208.
79. M. Hayyan, M. A. Hashim and I. M. AlNashef, *Chem Rev*, 2016, **116**, 3029-3085.

80. L. M. van Langen, R. P. Selassa, F. van Rantwijk and R. A. Sheldon, *Org Lett*, 2005, **7**, 327-329.