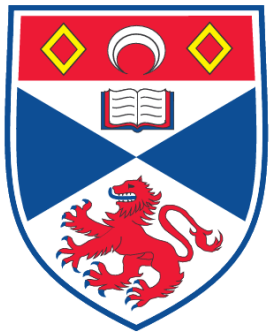
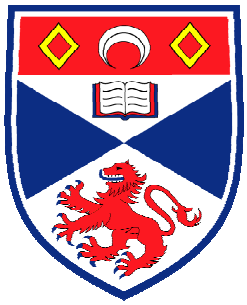
IN VITRO STUDIES OF THE ENZYMES INVOLVED IN FLUOROMETABOLITE BIOSYNTHESIS IN *STREPTOMYCES CATTLEYA*

Stuart Cross

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews



2009



University
of
St Andrews

In vitro studies of the enzymes involved in fluorometabolite biosynthesis in *Streptomyces cattleya*

By
Stuart Cross

A thesis presented for the degree of
Doctor of Philosophy
in the
School of Chemistry
University of St Andrews

December 2008

I, Stuart Cross, hereby certify that this thesis, which is approximately 39,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in October 2005 and as a candidate for the degree of PhD in September 2006; the higher study for which this is a record was carried out in the University of St Andrews between 2005 and 2008.

Date 23/04/09 Signature of Candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 23/04/09 Signature of Supervisor

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Access to Printed copy and electronic publication of thesis through the University of St Andrews.

Date 23/04/09 Signature of candidate

Date 23/04/09 Signature of supervisor

Acknowledgements

I would like to thank Professor David O'Hagan for allowing me this opportunity. I am grateful for all of his help, support and guidance over the last 3 years. I am also thankful to GlaxoSmithKline (GSK), particularly Dr. Antony Gee, and BBSRC for financial support.

I would also like to thank Dr Hai Deng (now at the University of Aberdeen), for being a great source of help and advice, and also for being a great mate and someone I thouroughly enjoyed working with. The same thanks also go to Mayca Onega, who also helped me a lot, and I really appreciate it. It has been a great pleasure to work with both of you, and the same must be said of the O'Hagan group as a whole.

I am grateful to a number of people at the University of St Andrews for technical support and advice. Mrs Melanja Smith and Dr Tomas Lebl for the NMR service, Dr Huanting Liu and Dr Louise Major for E. coli expression vectors and enzymes, Dr Uli Scwarz-Linek for ITC, and finally Dr Catherine Botting and Mr Alex Houston for mass spectrometry. Thanks also go to Dr Jack Hamilton (University of Belfast) for GC-MS analysis.

I would also like to thank Mike Harrison, for being a great mate and flat mate for all the time we have been at St Andrews together.

Finally I would like to thank my Parents and the rest of my family, without whom this would not of been possible.

List of Abbreviations

4-FT - 4-fluorothreonine

5'-FDA - 5'-fluorodeoxyadenosine

5'-FDI - 5'- fluorodeoxyinosine

5-FDRP - 5-fluorodeoxyribose-1-phosphate

5-FDRulP - 5-fluorodeoxyribulose-1-phosphate

DAB - diaminobenzidine

DHAP - dihydroxyacetone-1-phosphate

FAc - fluoroacetate

FAld - fluoroacetaldehyde

FALD - fuculose aldolase

HRP - horseradish peroxidase

ITC - isothermal calorimetry

LAAO - L-amino acid oxidase

L3GP - glycerol-3-phosphate

MTRI - methylthioribose-1-phosphate isomerase

NMR - nuclear magnetic resonance

PET - positron emission tomography

SAM - s-adenosyl L-methionine

SDS-PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis

Abstract

Enzymatic fluorination of natural products is extremely rare. Of the 4000 halogenated natural products identified, only 13 possess a fluorine atom. The C-F bond forming enzyme from the soil bacterium, *Streptomyces cattleya*, remains the only native enzyme to be identified that is capable of such biochemistry. It generates 5'-fluoro-5-deoxyadenosine (5'-FDA) from *S*-adenosyl-L-methionine (SAM) and F^-. The "fluorinase" is the first committed step toward the biosynthesis of the two fluorometabolites, 4-fluorothreonine and fluoroacetate, *via* the common intermediate, fluoroacetaldehyde (FAld). The enzymatic steps responsible for the conversion of 5'-FDA to the fluorometabolites remained to be fully characterised when this project began.

Previously, a purine nucleoside phosphorylase was identified that was capable of generating 5-fluorodeoxyribose-1-phosphate (5-FDRP) from 5'-FDA. 5-FDRP is subsequently isomerised to 5-fluorodeoxyribulose-1-phosphate (5-FDRulP) by an aldose-ketose isomerase enzyme.

Chapter 2 describes the identification of the isomerase gene from the genomic DNA of *S. cattleya* and the corresponding protein product was capable of generating 5-FDRulP from 5-FDRP.

The next intermediate, FAld, is generated from 5-FDRulP by a fuculose aldolase. Attempts to identify the aldolase gene from *S. cattleya* were unsuccessful, however a putative fuculose aldolase from *Streptomyces coelicolor* was isolated that could generate FAld from 5-FDRulP, which is described in Chapter 3.

Following the identification and over expression of a PLP-dependant transaldolase, which generates 4-fluorothreonine (4-FT) from FAld and L-threonine in *S. cattleya*, Chapter 4 details the successful *in vitro* reconstitution of fluorometabolite biosynthesis using five over- expressed enzymes.

In Chapter 5, attempts to develop a novel assay for fluorinase activity was explored. The colorimetric detection of L-methionine produced by the fluorinase in a coupled L-amino acid oxidase and horseradish peroxidase assay, leading to the oxidation of a dye substance. This was carried out with interest in developing a high-throughput assay for fluorinase mutants, generated by random mutagenesis, in order to identify those with increased activity. In the event, it proved unsuccessful.

Table of Contents

1 Introduction

1.1 Natural products in medicine

Plant natural products have been used to treat and cure disease since the very beginning of medicine.[1] Today's advanced screening and detection techniques, and the discovery of a wealth of marine sources of natural products[2] means that therapeutic agents derived from natural sources remain a major avenue toward drug discovery and development. The use of natural products as the source of novel therapeutics peaked in the Western pharmaceutical industry between 1970 and 1980, and between 1981 and 2002 around 49% of the 877 small molecule New Chemical Entities (NCE's) were natural products, semi-synthetic natural product analogues or synthetic compounds based on natural-product pharmacophores.[3] Indeed over 25% of all drugs currently in circulation have their origins in natural products, and more than 80% of the world's population rely on natural extracts for primary healthcare.[4] It is therefore unsurprising that, in the wake of annual worldwide pharmaceutical spending of more than US$30 billion, this is a significant area in modern drug discovery.

Natural products have evolved to complement normal metabolic processes, usually as a defence mechanism after a period of active growth, in a nutrient-deficient environment. Although plants are the most abundant sources of secondary metabolites, bacteria, fungi

and increasingly marine organisms are also useful resources for natural product discovery. The majority of low molecular weight natural products can be categorised into alkaloids, terpenoids, polyketides, glycosides and phenolic compounds. Large natural product molecules include the ribosomal and non-ribosomal peptides.[5, 6]

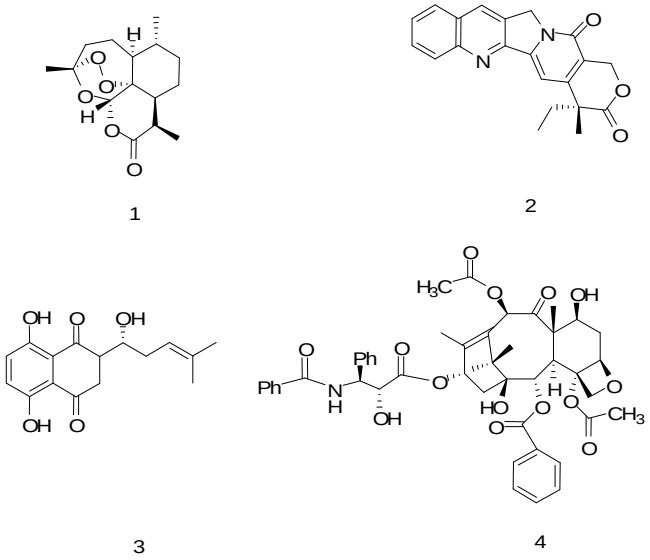


Figure 1.1. Selected natural products; artemisinin **1**, camptothecin **2**, shikonin **3**, taxol **4**.

Intensive studies on these natural products and the enzymatic mechanisms by which they are generated has lead to a greater understanding of the ecological role of these substances, as well as providing frameworks for elaboration in medicinal and organic chemistry. Some of the most powerful anti-cancer and anti-malarial compounds currently in production are natural product compounds. Artemisinin **1**[7], camptothecin **2**[8], shikonin **3**[9], and taxol **4**[10] are significant examples (Figure 1.1).

1.2 Biological halogenation

More than 4000 natural products have been isolated from natural sources[11] that incorporate chlorine, fluorine, bromine or iodine atoms. Representatives of this group display a wide range of biological activities, including anticancer and antibiotic properties.[12] These halogenated products have been isolated from bacteria, fungi, marine algae, lichens, higher plants, mammals and insects.

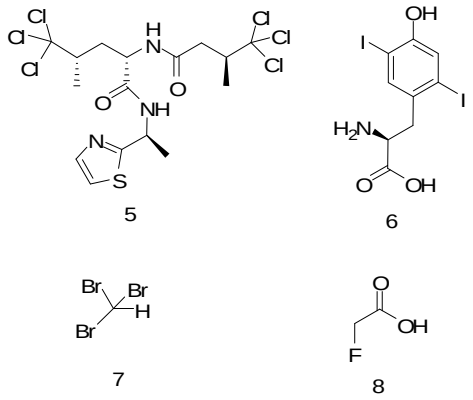


Figure 1.2. Structurally diverse halogenated compounds; nordysidenin **5** (*Lyngbya majuscule*)[13, 14], diiodotyrosine **6** (*G. cavollini*), bromoform **7** (marine algae) and fluoroacetate **8**[17].

Brominated secondary metabolites are most commonly identified from marine organisms and chlorinated metabolites are the more prevalent in terrestrial organisms. In contrast, the generation of fluorinated and iodinated compounds are far less common. Halogenated secondary products exhibit wide structural diversity, as the examples in Figure 1.2 show. The presence of a halogen atom is critical for biological activity for many of these compounds. This has been observed in the antibiotic vancomycin **16**, which requires two chlorine atoms in order to achieve a clinically active conformation[15], and the antitumour

compound rebeccamycin **18**, which loses antimicrobial activity when the chlorines are removed.[16]

1.3 Enzymatic halogenation

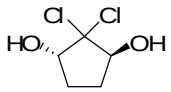
Recently, more details have emerged on the mechanism by which halogens are incorporated into organic compounds. For a long period, haloperoxidases were thought to catalyse all halogenation reactions. However the identification of a greater number of halogenated natural products, combined with more sophisticated techniques to elucidate the mechanisms of halogenation, have led to the recent re-evaluation of enzymatic halogenation into several distinct categories.

1.3.1 Haloperoxidases and halogenases

The generation of a reactive hypohalite species by two-electron oxidation of halide ion is one of the main strategies for enzymatic halogenation. This method is used to halogenate electron-rich carbon centres of natural products and occurs by two distinct mechanisms, which seperates the two classes of halogenating enzymes that perform this biochemistry; the haloperoxidases and the halogenases. Hydrogen peroxide is used by the haloperoxidase class, whereas molecular oxygen is used by the halogenase class.

1.3.1.1 Haloperoxidases

Haloperoxidases are a group of enzymes that catalyse the halogenation of organic compounds in the presence of H_2O_2. The first halogenating enzyme to be characterised was from the bacterium *Caldariomyces fumago*, during an investigation into the biosynthesis of the chlorinated metabolite, caldariomycin **9**. [17] It was observed that the enzyme responsible for this halogenation required a chloride ion and H_2O_2. As a result, this enzyme was named a 'chloroperoxidase'.[18]

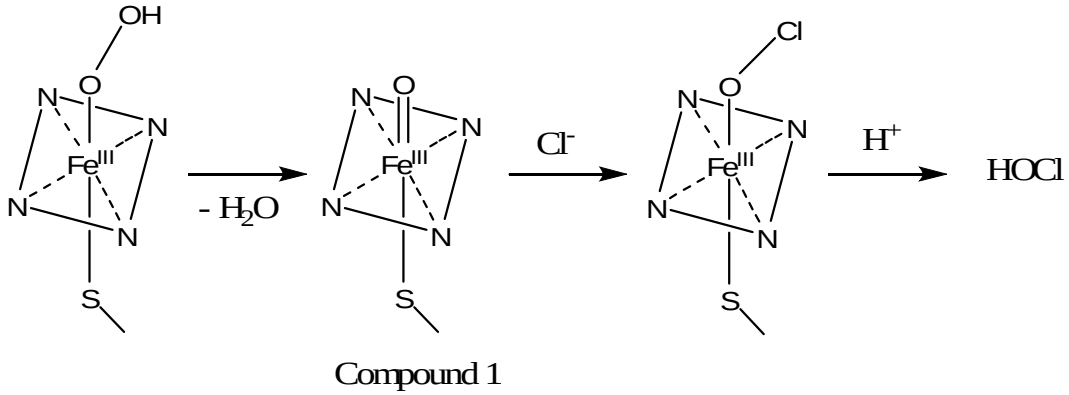


9

Other haloperoxidases have been identified from a wide range of prokaryotes and eukaryotes[19]. They are classified by the halide source that they use; chloroperoxidases are able to incorporate chloride, bromide and iodide, bromoperoxidases use bromide and iodide, and iodoperoxidases only iodide. There are no haloperoxidases capable of incorporating fluoride as a halide source. These enzymes can be further subcategorised according to their catalytic mechanism; those that contain a haem group, those which contain vanadium and those that do not contain metal ions, the perhydrolases.

1.3.1.1.1 Haem containing haloperoxidases (H-HPOs)

The chloroperoxidase from *C. fumago* contains a haem group, and is the prototypical haem-dependant haloperoxidase. During halogenation, hypohalous acid (HOCl) is generated as the halogenating agent in the presence of H_2O_2 and halide ions (Scheme 1.1).

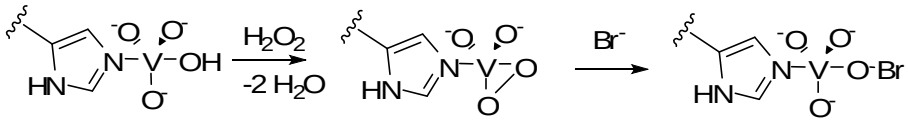


Scheme 1.1. Formation of hydrochlorous acid by haem-dependent haloperoxidases.[12]

Chloroperoxidases are capable of the dichlorination of electron rich carbon centres, for example at the 3- and 5- carbons of the amino acid tyrosine.[18] The reaction is proposed to proceed *via* the binding of H_2O_2 to the axial position of the ground state Fe^{III}-porphyrin complex. This is followed by the removal of water to generate the Fe^{IV}-oxo species, known as compound I (Scheme 1.1), which forms an Fe^{III}-hypohalite species in the presence of halide. This reactive intermediate can directly halogenate substrates at the active site, or free hypohalous acid can be released to cause remote halogenation away from the active site.

1.3.1.1.2 Vanadium containing haloperoxidases (V-HPOs)

Studies on the halogenation of marine natural products have revealed a different haloperoxidase from marine algae which required vanadium instead of iron for halogenation.[20] These enzymes are thought to be responsible for the majority of halogenation events during the biosynthesis of marine natural products. Vanadium-dependant bromoperoxidases are well distributed through seaweeds.[21] Vanadium-dependant chloroperoxidases have also been found in terrestrial fungi and two bacterial species.[22] Like the H-HPOs, the metal centre of V-HPO binds hydrogen peroxide and activates it for attack by halide ion. However unlike H-HPOs, the vanadium is not redox active and maintains its oxidation state (V(V)) throughout the catalytic cycle (Scheme 1.2).

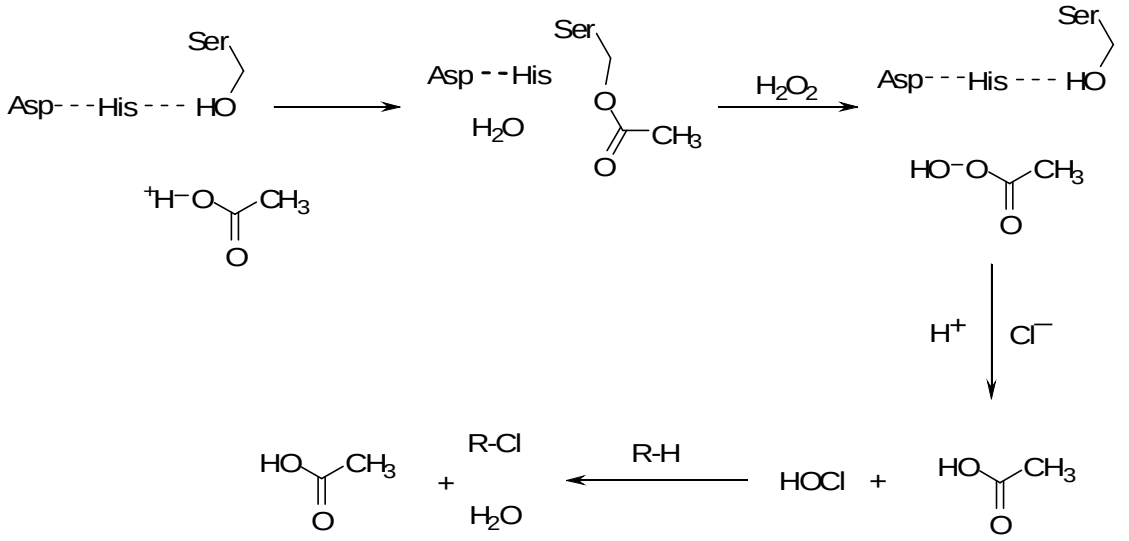


Scheme 1.2. The formation of vanadium-bound hypobromite by vanadium-dependant bromoperoxidases.[12]

1.3.1.1.3 Perhydrolases

Halogenating enzymes that do not possess a haem or metal group, but that are dependant upon hydrogen peroxide for activity are the perhydrolases. These enzymes have been isolated from the soil bacterium *Streptomyces lividans* and the proteobacterium *Pseudomonas fluorescens*.[23] The reaction mechanism of these enzymes proceeds via the

formation of an acyl-enzyme intermediate by the reaction of a short-chained carboxylic acid with a serine residue at the active site.[24] The addition of H_2O_2 causes the perhydrolysis of the acyl-enzyme intermediate, forming a peracid which then in turn oxidises halide ions to hypohalous acids, the halogenating agent (Scheme 1.3).[25]



Scheme 1.3. The proposed enzymatic mechanism of perhydrolases.[25]

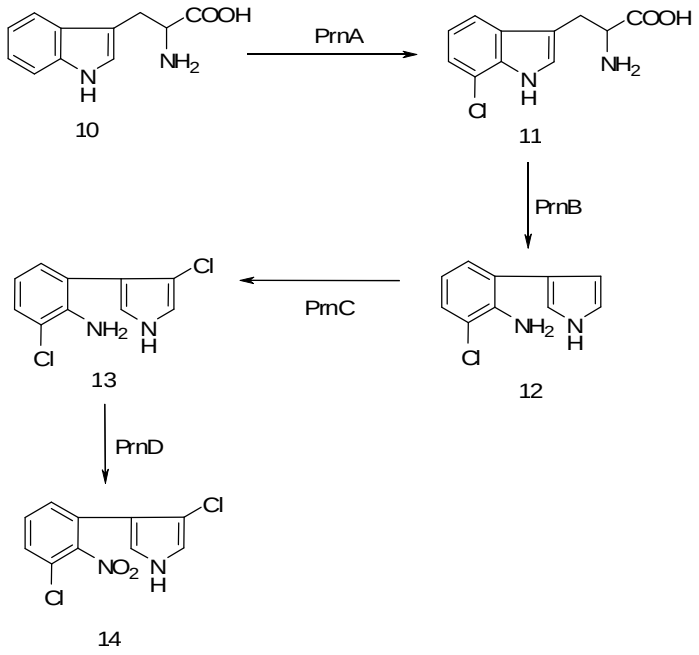
1.3.1.2 Halogenases

Halogenases are a group of enzymes that catalyse the halogenation of organic compounds in the presence of molecular oxygen.

1.3.1.2.1 Flavin-dependant halogenases

Enzymatic halogenation *via* hypohalite can also be catalysed using molecular oxygen as the oxidant and flavin as the redox cofactor. The first halogenase of this nature to be

characterised was PrnA from *P. fluorescens*[26, 27, 28] involved in the production of 7-chlorotryptophan **11** from free tryptophan **10** on the biosynthetic pathway of the anti-fungal compound pyrrolnitrin **14** (Scheme 1.4).



Scheme 1.4. The biosynthetic steps to pyrrolnitrin **14** in *P. fluorescens*.[26-28]

Members of this family of halogenase enzymes have subsequently been identified in the production of chlorotetracyclin **15**[26], vancomycin **16**[29], calicheamicin[30], balhimycin[31], and pyoleuterin **17**[32] biosynthesis (Figure 1.4).

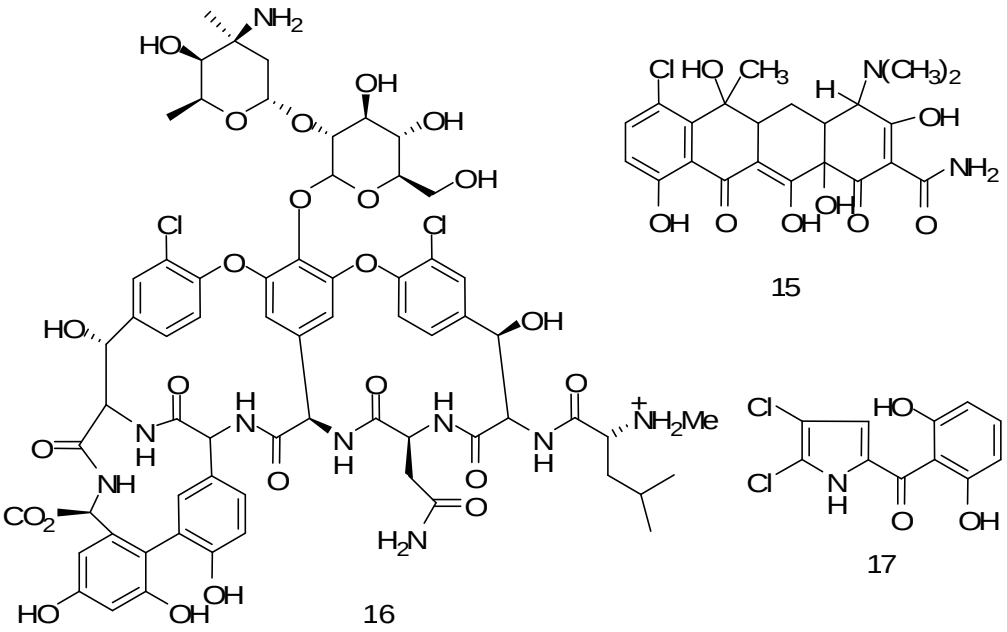
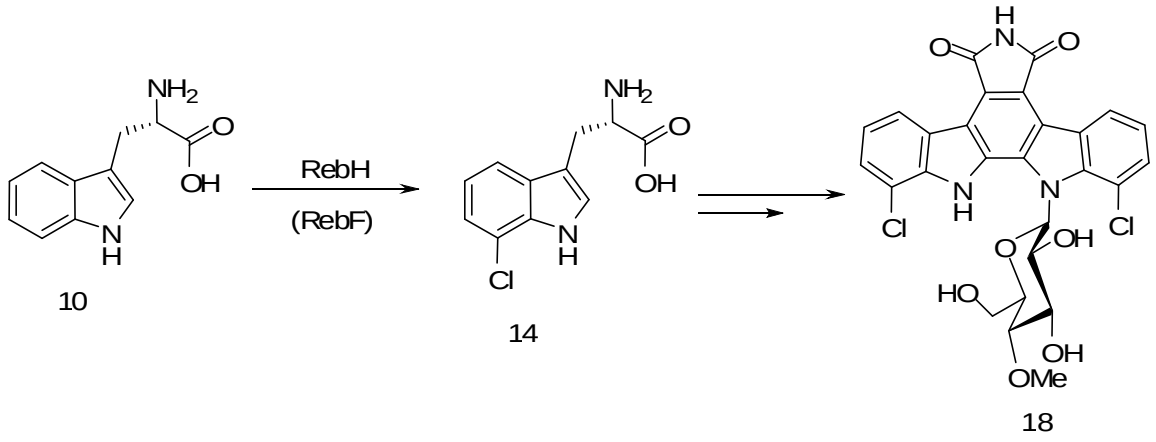


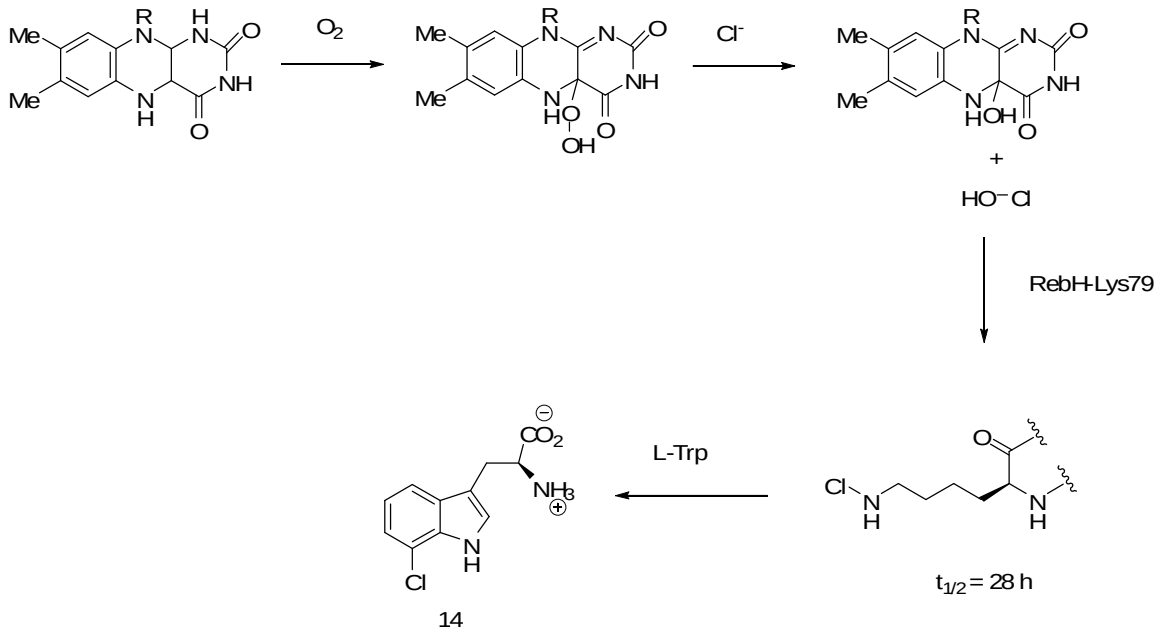
Figure 1.4. Some products of Flavin-dependant halogenases.

The crystal structure of the PrnA halogenase was recently elucidated at St Andrews University by J. Naismith and co-workers[33]. Also C. Walsh and co-workers identified an enzyme capable of the identical transformation to 7-chlorotryptophan (RebH) in the biosynthesis of the natural product, rebeccamycin **18** (Scheme 1.5)[34]. This enzyme was also crystallised.[35]



Scheme 1.5. The chlorination of 7-chlorotryptophan as the first step in rebeccamycin **18** biosynthesis.[34]

Subsequently, a mechanism for the regioselective chlorination of tryptophan **10** by PrnA and RebH was proposed (Scheme 1.5).



Scheme 1.6. Mechanism of halogenation by RebH (and PrnA). [33, 35]

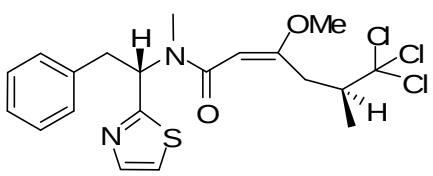
Crystal structures of PrnA and RebH reveal the flavin binding domain with a chloride ion bound in a pocket on the solvent-protected face of the pocket.[33, 35] The tryptophan binding pocket is located 10 Å away from the flavin cofactor, with a narrow channel connecting the two sites which prevents direct interaction between the substrate and oxidised flavin. It is proposed that chloride ion attack on the distal oxygen of the oxidised flavin produces an enzyme trapped HOCl that can diffuse toward the substrate binding site, specifically the side-chain of a lysine residue. A conserved lysine residue in the tryptophan **10** binding site (Lys79, RebH) first reacts with HOCl to generate a less reactive, but more selective, lysine-chloroamine species with a half life of 28 h.[35] In the presence of tryptophan **10**, the selective chlorination at the 7- position occurs to generate the product 7-chlorotrytophan **14**.

Halogenases catalysing chlorination at the 5- and 6- positions of tryptophan **10** have also been described.[36, 37] These homologous enzymes produce a single chlorotryptophan isomer, exemplifying the control of regioselective halogenation by this class of halogenases. These halogenases are also thought to be responsible for the halogenation of aromatic substrates in secondary metabolite biosynthesis.

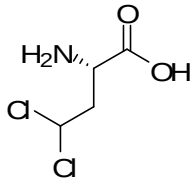
1.3.2 Halogenation using halogen radicals

The identification of halogenated natural products such as the marine molluscicide barbamide **19** from the cyanobacterium *Lyngbya majuscule*,[38] the antibiotic armentomycin **20** from *Streptomyces armentosus*[39] and the plant toxin syringomycin **21**[41] (Figure 1.5) demonstrates that chlorine is incorporated at unactivated carbon centres.

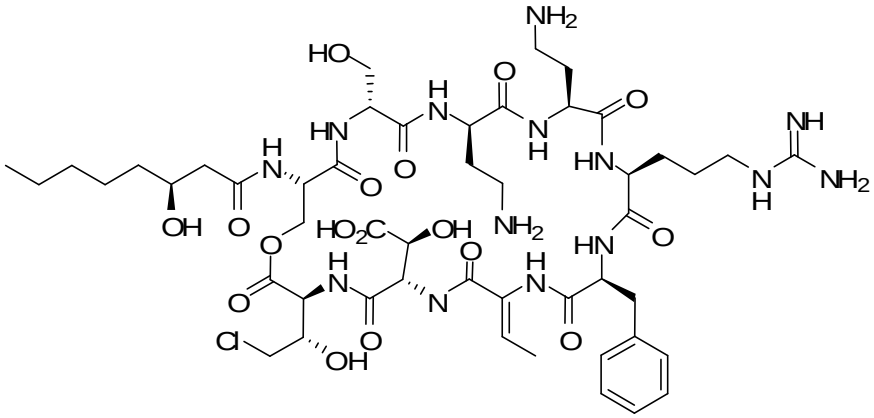
These centres are not obviously ameanable to electrophilic halogenation and therefore a radical mechanism for halogenation was proposed. Recent *in vitro* reconstitution of the barbamide **19**[40], syringomycin **21**[41] and armentomycin **20**[42] biosynthetic pathways revealed that ferrous iron, chloride, oxygen and α-ketoglutarate (αKG) are required for enzymatic activity. These mononuclear non-haem iron halogenases are imbedded in the non-ribosomal peptide synthetase (NRPS) assembly lines, and act on the methyl groups of the thiolation domain-tethered amino acids.



19



20



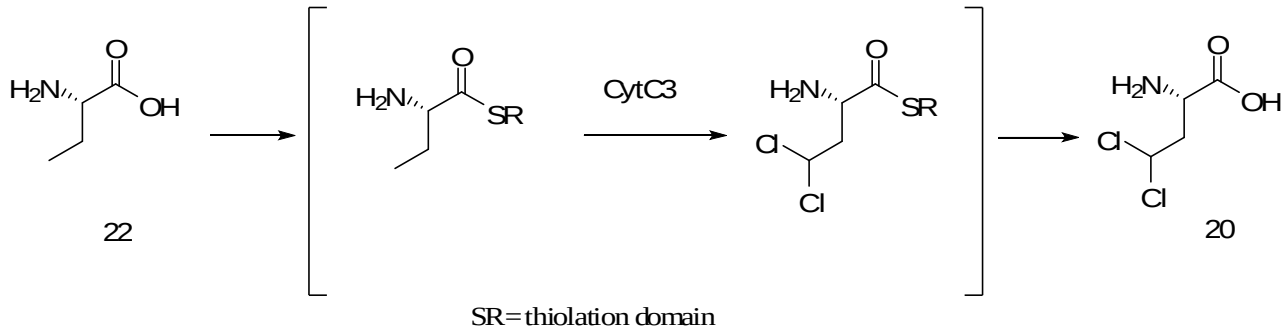
21

Figure 1.5. The chlorinated natural products barbamide **19**[40], armentomycin **20**[42] and syringomycin **21**.[41]

1.3.2.1 Chlorination by Fe(II)/αKG-dependant halogenases

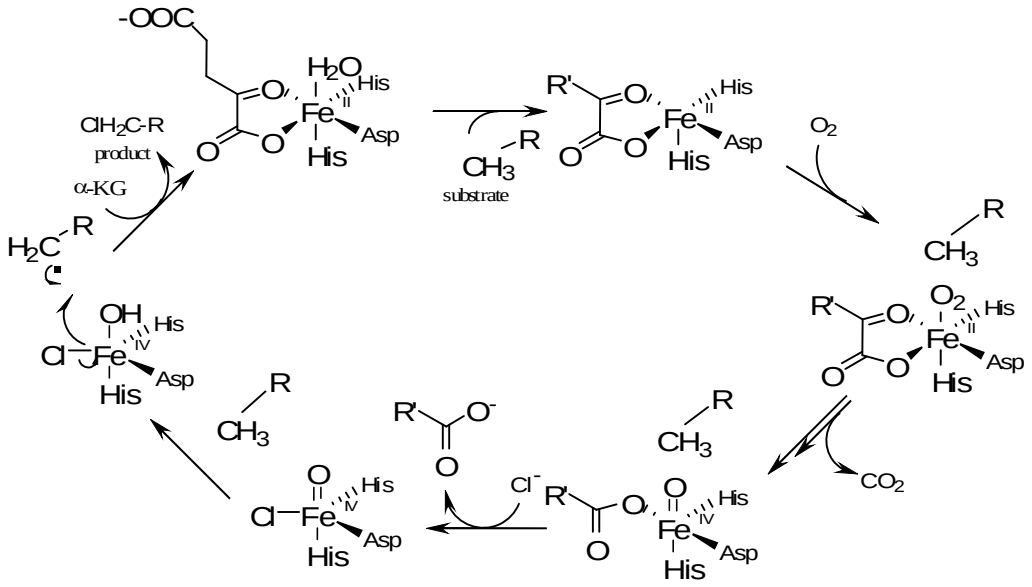
The chlorination step involved in the biosynthesis of armentomycin **20** is carried out by an Fe(II)/αKG-dependant halogenase, CytC3 (Scheme 1.7). This halogenase adds two

chlorine atoms to the terminal methyl group of L-2-aminobutyric acid (Aba) **22**. Chlorination occurs when the amino acid is attached to the thiolation domain (CytC2), during armentomycin **20** biosynthesis.[42]



Scheme 1.7. Halogenation of L-2-aminobutyric acid **22** to armentomycin **20** by CytC3.[42]

Characterisation of the intermediates during the CytC3-catalysed chlorination of the Aba-S-CytC2 complex reveal that halogenation proceeds *via* the formation of a Fe(IV)-oxo species with similar characteristics to Fe(II)- and αKG-dependant dioxygenases. In these dioxygenases, Fe(IV)-oxo is a key catalytic intermediate which removes hydrogen from the substrate molecule to form a substrate radical and an Fe(III)-OH species.[43, 44] Iron is coordinated by two histidines and one carboxylate residue (i.e an aspartic acid or glutamic acid) in a "facial triad".[45, 46, 47] Recently, the X-ray structure of the halogenase responsible for chlorination in syringomycin **21** biosynthesis revealed that the active site iron is coordinated by two histidine residues, and that the carboxylate residue is replaced by chloride ion which coordinates to iron.[48] It was determined that the Fe(IV)-oxo species catalyses C-H cleavage to initiate substrate halogenation[49] (Scheme 1.8).



Scheme 1.8. Proposed mechanism for CytC3. R= L-2-aminobutyric acid.[49]

Binding of dioxygen to Fe(II) in halogenases leads to the formation of a Cl-Fe(IV)-oxo species, activating the substrate for halogenation by abstraction of hydrogen to form the substrate radical and a Cl-Fe(III)-OH intermediate. The oxidative transfer of the chlorine atom to the substrate radical results in product formation and reduction of iron to the Fe(II) oxidation state at the active site.

It appears that the Fe(II)/αKG-dependant halogenases have evolved from dioxygenases as they are analogous in many respects. These enzymes have been developed in order to halogenate non-activated carbon centres and they generate a variety of halogenated natural products.

1.4 Biological fluorination

Although fluorine is the most abundant halogen in the earth's crust (ranging from 270-740 ppm), compared to that of chlorine (10-180 ppm),[50] incorporation into organic compounds is extremely limited.[51, 52] Its ability to form largely insoluble salts (e.g. fluorospar) with inorganic cations, leads to very poor bioavailability. Fluorine is the smallest of all of the halogens, with an atomic radius only slightly larger than hydrogen. However fluorinated natural products are extremely rare and their numbers do not appear to be increasing despite extraction, isolation and screening methods becoming more sophisticated.[53] The low bioavailability coupled with fluoride ion being a very poor nucleophile in water (only 1.3 ppm in sea water), makes it a poor candidate for enzymology. Fluoride ion cannot be oxidised like the other halogens by the haloperoxidases to form an X^+ species because its redox potential is too low (Table 1.1).[54, 55] As a direct result, very few biological systems have evolved to incorporate fluoride into organic compounds.

Halogen, X^-	**Heat of hydration, X^- [KJ mol^{-1}]**	**Standard redox potential (E^0)**
F^-	490	-3.06
Cl^-	351	-1.36
Br^-	326	-1.07
I^-	285	-0.54

Table 1.1. Heat of hydration and standard redox potential for the halogens.[54, 55]

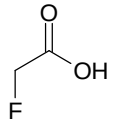
1.4.1 The organo-fluorine compounds

Only seven classes of organic compounds that contain fluorine have been identified. Three of these compounds are found in plants.

1.4.1.1 Organo-fluorine metabolites from plants

1.4.1.1.1 Fluoroacetate 8

Fluoroacetate **8** is the most abundant fluorinated metabolite found to date. It is largely biosynthesised as a toxin by some plants and one bacteria. Fluoroacetate **8** was first isolated from the South African shrub *Dichapetalum cymosum*, where the leaves of this plant were known to be toxic to cattle.[56]

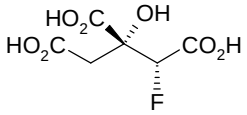


8

Many other species from the *Dichapetalum* genus have been shown to contain high levels of fluoroacetate **8** in their leaves.[57, 58, 59] In Australia, more than forty plant species from the *Leguminosae* genus have been shown to contain traces of fluoroacetate **8**.

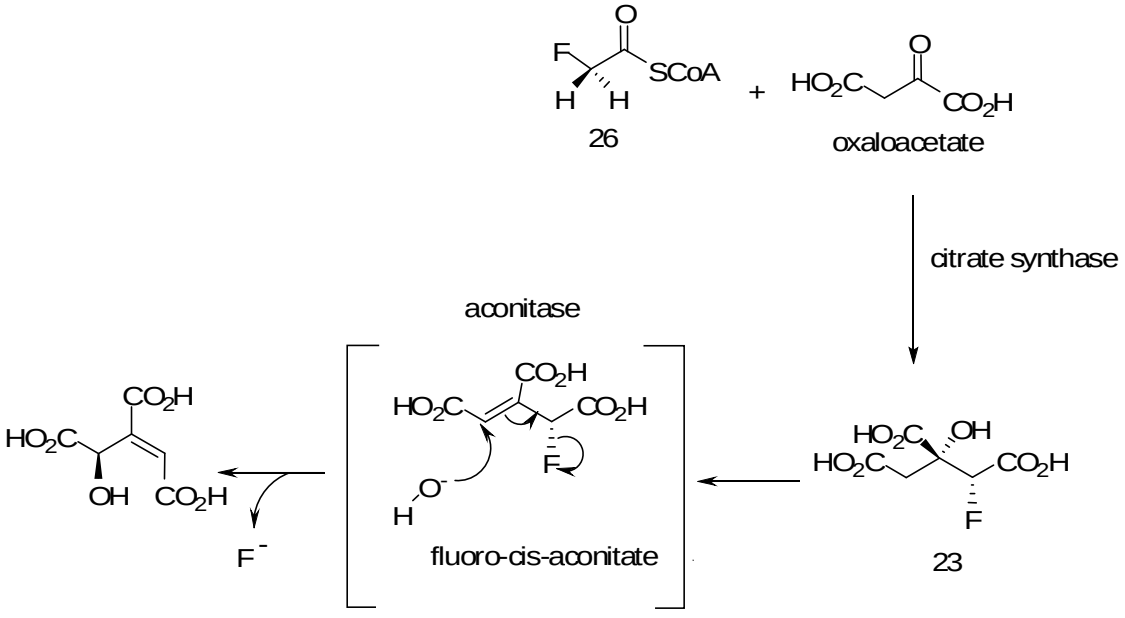
1.4.1.1.2 Fluorocitrate 23

The toxicity of fluoroacetate **8** is due to its *in vivo* activation to fluoroacetyl CoA **26**, which is then combined with oxaloacetate by citrate synthase.



23

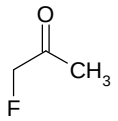
This highly stereospecific reaction generates the toxic stereoisomer, *(2R, 3R)*-fluorocitrate **23**, a competitive inhibitor of aconitase; the subsequent enzyme in the citric acid cycle.[60] This has a toxic effect on cells, because the pathway through which cellular energy is generated is blocked by this "lethal synthesis"[61] (Scheme 1.9).



Scheme 1.9. The 'lethal synthesis' of fluorocitrate **23** from fluoroacetyl-CoA **26** and oxaloacetate.[61]

1.4.1.1.3 Fluoroacetone 24

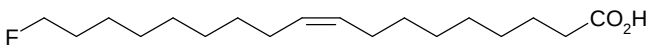
Fluoroacetone **24** was first identified in the Australian plant *Acacia georginae* as a "volatile" organo-fluorine compound.[62, 63] However due to problems with the derivatisation of these compounds, it is possible that fluoroacetaldehyde **40** (FAld) may be the metabolite as suggested in the initial report itself.



24

1.4.1.1.4 Fluorinated fatty acids

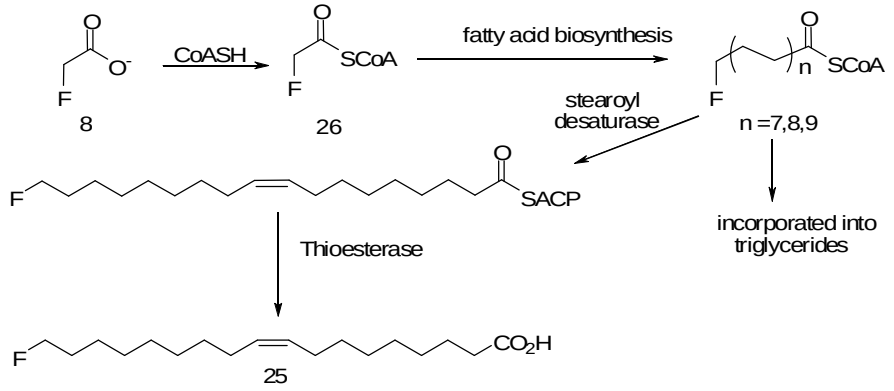
ω-Fluorooleic acid **25** was discovered as a constituent (~3%) of the seed oil of the West African plant *Dichapetalum toxicarium*.[64]



25

More recent evaluation has revealed up to six more fluorinated fatty acids in *D. toxicarium* with varying chain lengths and all of which possess a fluorine atom at the terminal carbon (ω).[65] This may occur as a direct incorporation of fluoroacetyl-CoA **26**

during fatty acid biosynthesis rather than direct synthesis of an analogue by the plant itself (Scheme 1.10).

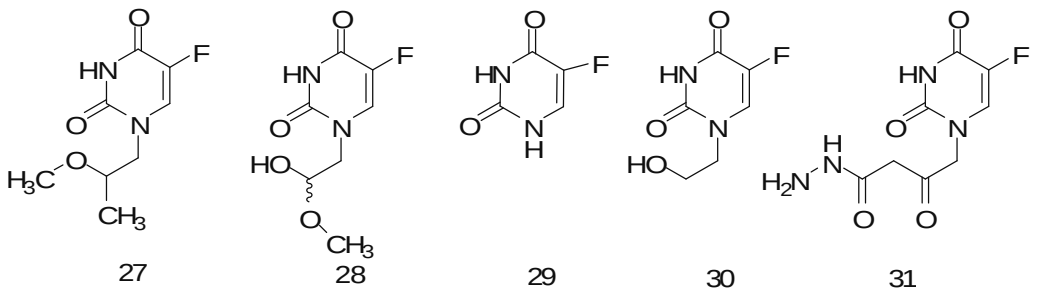


Scheme 1.10. The putative incorporation of fluoroacetate in ω-fluorofatty acid biosynthesis in *D. toxicarium*.

1.4.1.2 Organo-fluorine metabolites from marine sources

1.4.1.2.1 5'-Fluorouracil derivatives from the sponge *Phakellia fusca*

The only example of marine natural products containing a fluorine are the 5-fluorouracil alkaloids, **27-31** which were isolated from the sponge *Phakellia fusca* from the South China Sea.[66] Five compounds were isolated, including compounds **29** and **31** which are known to possess anti-tumour activity[67]. The remaining three were novel compounds.

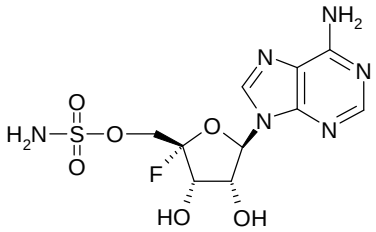


There appears to be some doubt over the biosynthesis of these compounds. They may have been accumulated as a result of industrial contamination in the ocean, and then uptake by the sponge, rather than a *de novo* biosynthesis.[66]

1.4.1.3 Organo-fluorine metabolites from bacteria

1.4.1.3.1 Nucleocidin 32

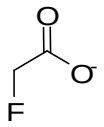
Nucleocidin **10** is an anti-trypanosomal antibiotic isolated from the actinomycete bacterium *Streptomyces calvus.*[68] It possesses a fluorine atom at the 4' position of the ribosyl ring system[69] and was the first organo-fluorine compound to be isolated from a bacterial source. Further attempts to re-isolate nucleocidin **32** have failed, possibly due to high levels of sub-culturing, which appears to have lead to a loss of biosynthetic capacity.



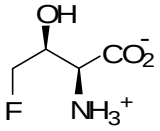
32

1.4.1.3.2 Fluoroacetate 8 and 4-fluorothreonine 33 from *Streptomyces cattleya*

The actinomycete *Streptomyces cattleya* was first recognised for its production of the β-lactam antibiotic thienamycin,[70] subsequently its ability to generate organo-fluorine metabolites was discovered.[71] Extracts containing fluoroacetate (FAc) **8** and 4-fluorothreonine (4-FT) **33** could be obtained during optimisation of thienamycin production.[71] It was discovered that growth media containing soy-bean casein was responsible for supplying fluoride for biosynthesis (0.7% inorganic fluoride). Fluorometabolite production occurs after a lag of up to five days in *S. cattleya* resting cell suspensions,[72] indicating that the fluorometabolites are secondary metabolites.



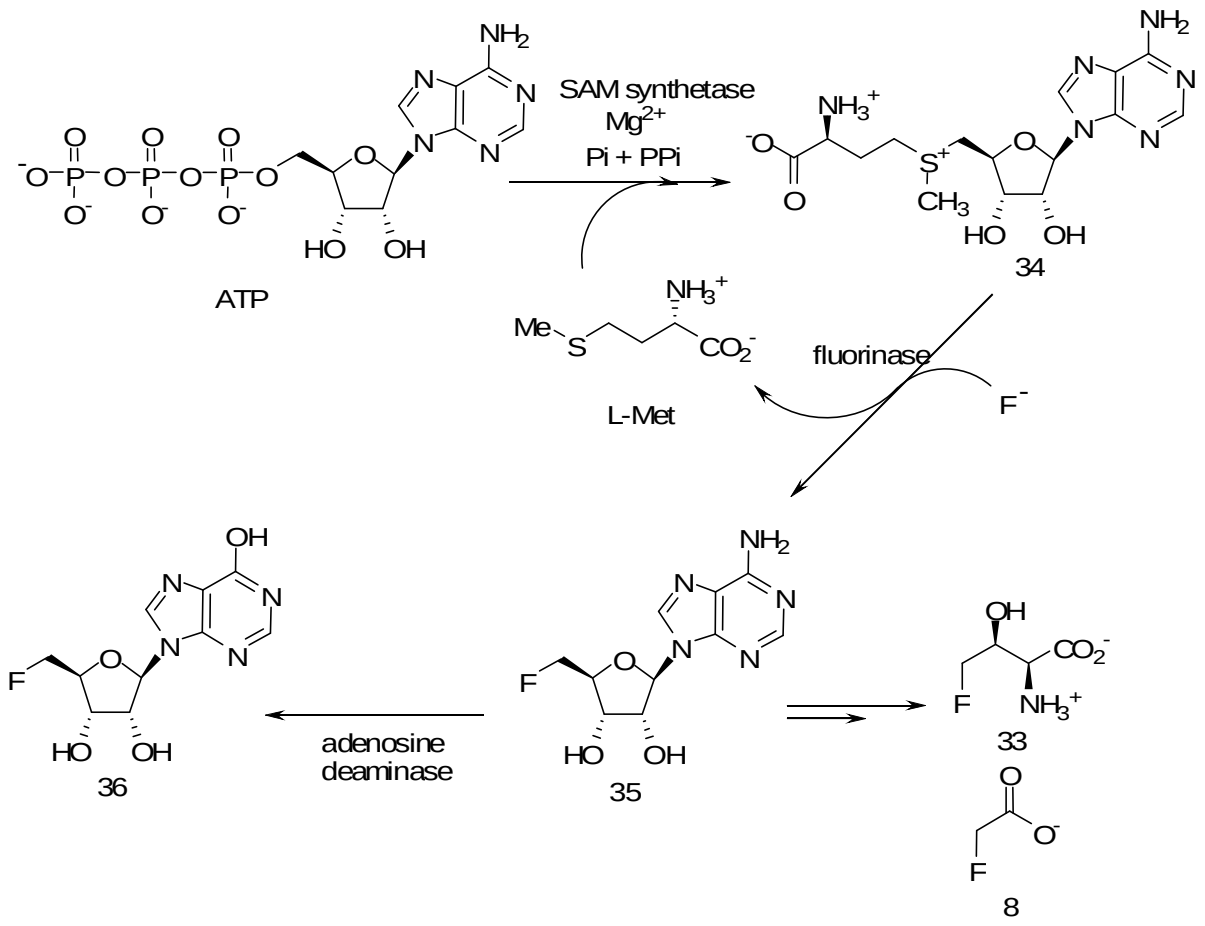
8



33

It has since been established that these fluorometabolites are synthesised by an enzyme capable of catalysing formation of the C-F bond. The conversion of ATP and inorganic fluoride to three fluorinated metabolites by cell free extracts of *S. cattleya* was shown by ^{19}F NMR.[73] Further experiments established that S-adenosyl methionine (SAM) **34** was also capable of being fluorinated under similar conditions. SAM **34** is metabolically related to ATP; SAM synthetase promotes a reaction between ATP and L-methionine to produce SAM **34**.[73] ^{19}F-NMR showed that SAM **34** was converted to 5'-fluoro-5'-

deoxyadenosine (5'-FDA) **35** by a "fluorination" enzyme contained in the cell free extract (Scheme 1.11).



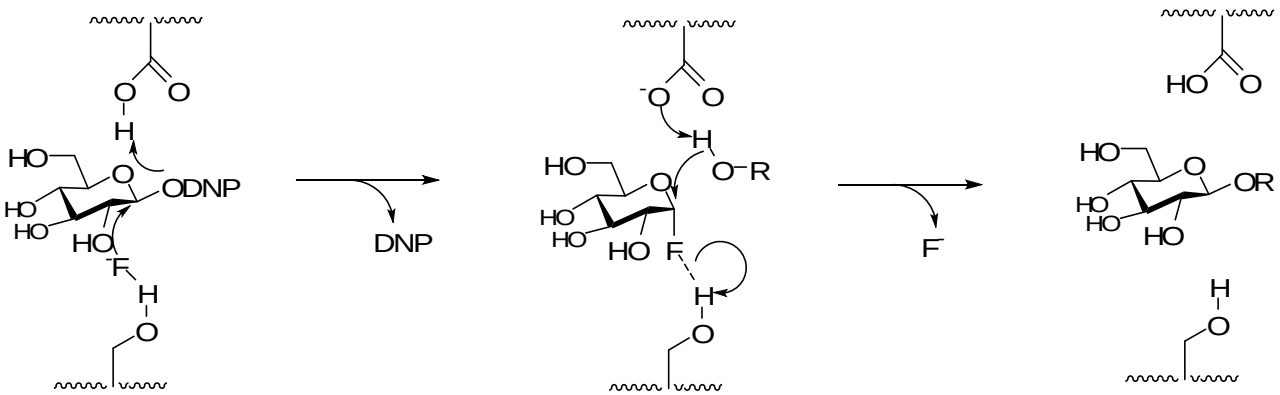
Scheme 1.11. The generation of 5'-FDA **35**, 5'-FDI **34**, 4-FT **33** and FAc **8** from ATP in *S. cattleya*.

Further inspection of the fluorinated products by ES-MS analysis also revealed the generation of the shunt product 5'-fluoro-5'-deoxyinosine **36** (5'-FDI), produced by the action of an endogenous deaminase on 5'-FDA **35** in the cell free extract.[74] Time course ^{19}F-NMR experiments revealed that 5'-FDA **35** synthesis by the fluorination enzyme is the first committed step on the biosynthetic pathway to FAc **8** and 4-FT **33** in *S. cattleya*[75, 76, 77] (Scheme 1.11).

1.5 Enzymatic C-F bond formation

1.5.1 Mutant glycosidases

Generation of the C-F bond has significance in the synthesis of commercial organo-fluorine compounds for use in the agrochemical, pharmaceutical and fine chemicals industries.[78, 79] Enzymatic formation of the C-F bond i.e. converting inorganic fluoride to organic fluorine, was first reported in mutant glycosidases that were capable of generating α-fluoroglycosides as transient intermediates from dinitrophenyl (DNP) activated sugars [80, 81] (see Scheme 1.12).

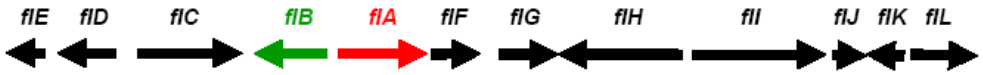


Scheme 1.12. The proposed mechanism of enzymatic C-F bond formation by mutant glycosidases.[80, 81]

1.5.2 The fluorinase from *S. cattleya*

The enzyme responsible for C-F bond formation in *S. cattleya*, 5'-fluorodeoxyadenosine synthase (fluorinase), was purified from wild type cell free extracts. SDS-PAGE showed that the fluorinase has a subunit mass of 32 kDa.[76] The fluorinase enzyme found in *S. cattleya* is the only native enzyme identified so far able to form the C-F bond.

Subsequent size exclusion chromatography revealed that this protein had a native mass of 180 kDa, revealing that the active protein exists as a hexamer.[75] N-Terminus amino-acid analysis and trypsin digest enabled the design of PCR primers which were used to amplify the fluorinase gene (*fl*A) from the genomic DNA of *S. cattleya* by J. Spencer and co-workers at Cambridge University.[82] Gene walking using these primers identified the location of the fluorinase gene within the genomic DNA and more recently a 10 kb gene cluster containing other genes involved in the fluorometabolite pathway of *S. cattleya* was identified (Figure 1.5).[83] Sequencing of *fl*A showed it to be 897 base pairs in length, coding for a protein of 299 amino acids corresponding to a monomer of 32 kDa,[82, 83] and confirming the initial purification results.



ORF	Start/Stop (bp)	Length (aa)	Function/Homology
E	130-795c	222	DNA binding regulatory protein
D	857-1504c	216	Dehalogenase/Phosphatase
C	1845-3036	397	MFS permease
B	3057-3953c	299	5'-FDA phosphorylase
A	4173-5069	299	5'-FDA syntase
F	5197-5751	185	DNA binding regulatory protein
G	5951-6652	234	DNA binding regulatory protein
H	6652-8052c	467	Na^+/H^+ antiporter
I	8314-9780	489	Homocysteine lyase
J	9803-10195	131	DNA binding protein
K	10592-10176c	139	Thioesterase/acyltransferase
L	10700-11374	225	DNA binding regulatory protein

Figure 1.5. Organisation of the 10kb gene cluster from *S. cattleya*, highlighting the fluorinase (*flA*, red) and the PNP (*flB*, green) genes which mediate the first two enzymes (steps **a** and **b**) of fluorometabolite biosynthesis. The annotations for the remaining genes are deduced from sequence homologies.[83]

The cloned *fl*A was then inserted into a pET28(a) plasmid and was over expressed in *E. coli* in the presence of IPTG. Fluorinase can be purified to around 9 mg/ml using nickel affinity and size exclusion chromatography. Kinetic data showed a catalytic rate constant (k_{cat}) of 0.07 min^{-1} and a Michaelis constant (K_m) for F^- of 2 mM and 74 µM for SAM **34**.[74] The low affinity for fluoride is thought to be linked to the difficulty with which the enzyme secures the desolvated fluoride ion due to its high heat of hydration. [54, 55]

1.5.2.1 Crystal structure of the fluorinase

Crystallisation of both wild-type and the over expressed fluorinase was carried out in order to determine its structure. Structures were solved with SAM **34** bound (PDB 1RQP) and also with the products of the fluorination reaction, 5'-FDA **35** and L-methionine bound (PDB 1RQR). These studies showed that the fluorinase is a hexamer, consisting of a dimer of trimers (Figure 1.7) constructed from 32 kDa monomers (Figure 1.6).[84]

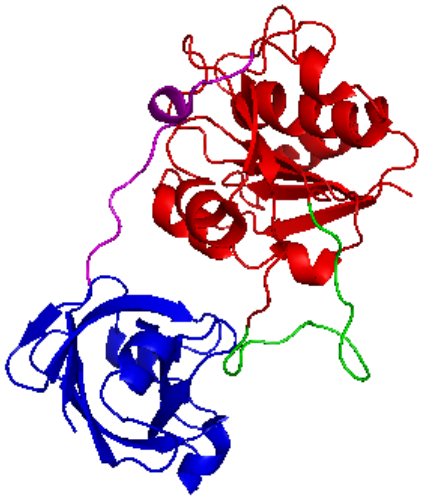


Figure 1.6. Monomeric structure of the fluorinase (PDB 1RQP). The N-terminal domain is coloured red, C-terminal domain in blue, 20 amino acid 'loop' in green.[84]

The X-ray structure of the fluorinase revealed that the 299 amino acid monomer was organized into two main domains, the amino- and carboxy- terminals. The N-terminal domain consists of residues 8-180 forming a seven stranded β-sheet which is contained between α-helices.[82] Within the N-terminus, an extended loop, consisting of residues 98-114 is apparent. This loop is putatively involved in the formation of the trimer and catalytically active hexamer structures, although its true role is not very clear.[82] The smaller C-terminal domain (residues 195-298) is made up of a 5- and a 4-stranded antiparallel β-sheet.[82]

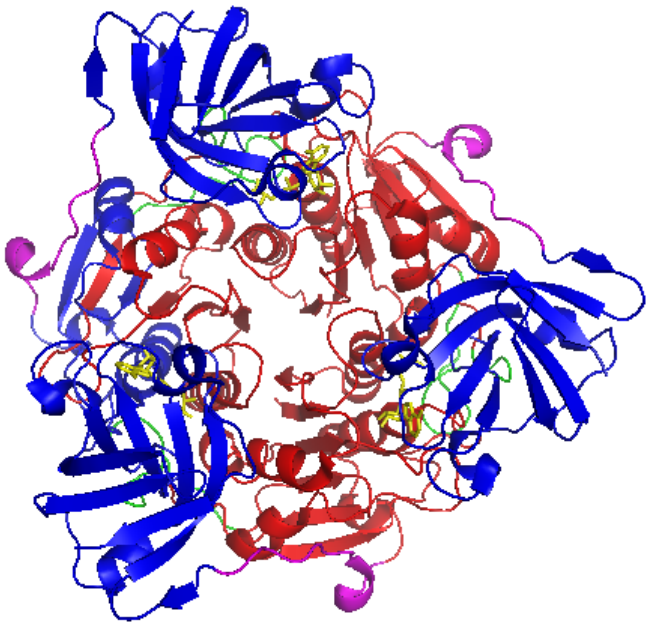


Figure 1.7. SAM-bound trimeric structure of the fluorinase (PDB 1RQP). The N-terminal domain is coloured red, C-terminal domains in blue, 20 amino acid loops in green, linker regions in magenta and the substrate SAM in yellow.[84]

The trimeric X-ray structure of fluorinase revealed that the three N-terminal domains are arranged in a 3-fold axis, making intimate contacts with each other.[82] The three C-terminal domains make contacts with N-terminal domains from neighbouring monomers.

X-ray structures with the substrates SAM **34** and the product, 5'-FDA **35**, revealed that the active site of the fluorinase is located at the interface between neighbouring N- and C-terminal domains in the trimeric structure. The active form of the fluorinase was identified as a dimer of trimers. [82]

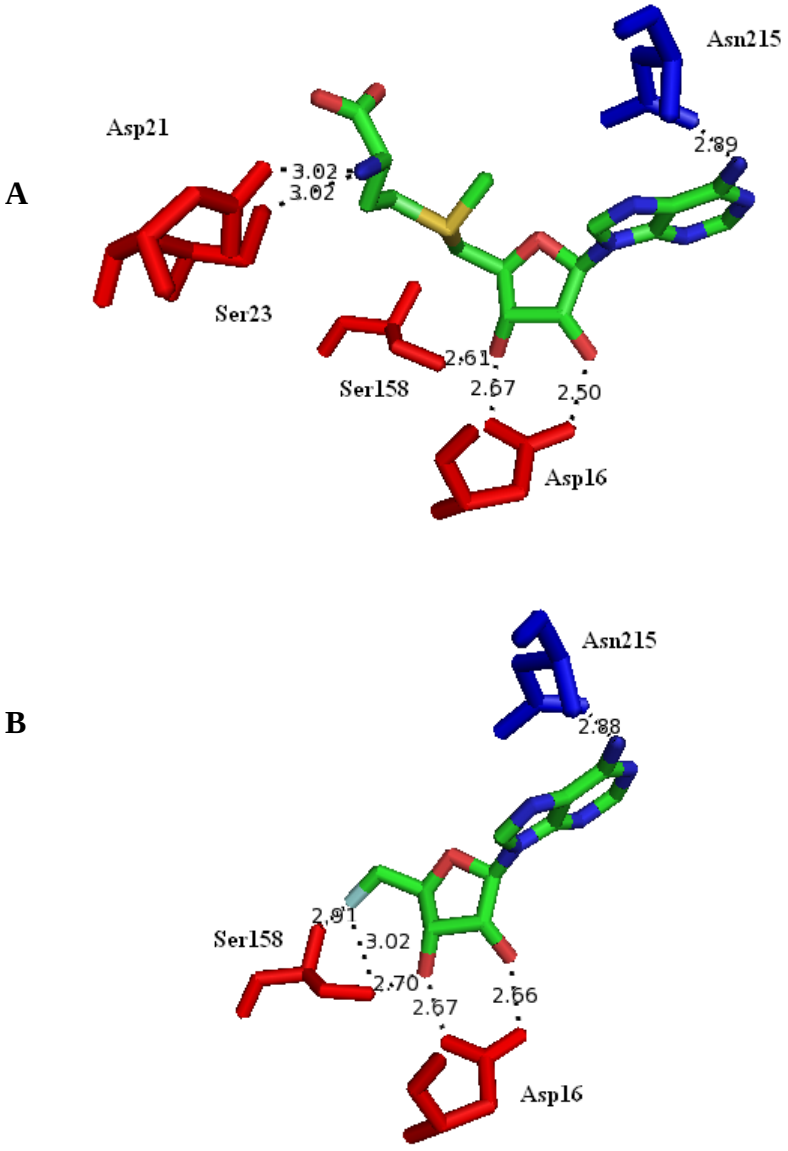


Figure 1.8. Selected residues at the active site of fluorinase involved in H-bonding with **A)** SAM **34** (PDB 1RQP) and **B)** 5'-FDA **35 B** (PDB 1RQR). N-terminal residues are coloured red, C-terminal residues blue.[82]

1.5.2.2 Mechanism of the fluorinase

The crystal structure of the fluorinase with SAM **34** or 5'-FDA **35** bound do not have water molecules near the fluorine pocket.[82] This suggests that fluoride ion is desolvated at the reaction centre. This desolvation is compensated for by two hydrogen bonds to Ser158. A third hydrogen bond is formed with Thr80, which is predicted to break its ground state hydrogen bond to the backbone carbonyl of Pro154 and form a new hydrogen bond with fluoride, as fluoride becomes fully desolvated.[78, 83] The fluoride ion is further stabilised at the active site by the positively charged sulphur of SAM **34** (Figure 1.9A)**.**[85] The full desolvation of fluoride ion is driven by the binding of SAM **34**. Dehydrated fluoride ions are potent nucleophiles, and SAM **34** then gets attacked to generate 5'-FDA **35** and L-methionine (Figure 1.9B). Stereospecific deuterium labelling studies, at the 5'-*pro*-S site of SAM **34**, was used to show that the newly formed C-F bond occurred with an inversion of configuration, indicative of an S_N2 reaction.[86] QM/MM calculations have suggested that the fluorinase lowers the barrier for C-F bond formation by 39 kJ mol^{-1} and a 10^6 fold increase in the rate of reaction compared with the (non-existant) reaction in solution.[77]

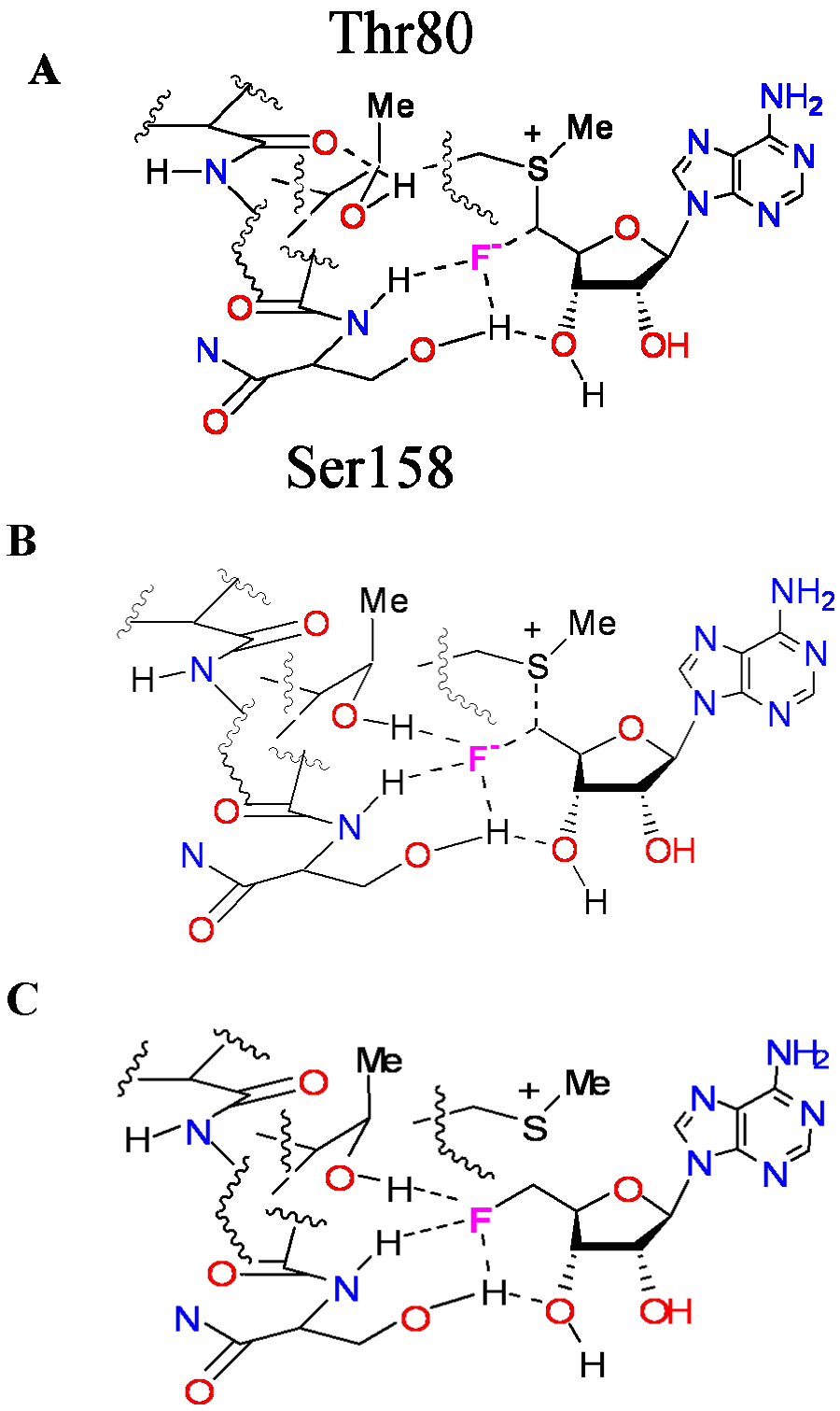


Figure 1.9. Proposed mechanism of the fluorinase determined by X-ray crystallography[82], QM/MM calculations[77] and labelling studies.[85] **A)** Ground state interactions with SAM **34** and fluoride ion. **B)** Transition state. **C)** Completed reaction.

Recent mechanistic studies using isothermal titration calorimetry (ITC) have revealed more information about substrate/product binding during fluoride turnover.[87] Structural studies using the apo (without adenosine bound at the active site) enzyme have revealed the presence of 4 water molecules in the active site in the place of adenosine. [87] During 5'-FDA **35** generation by the fluorinase, fluoride ion binds before SAM **34** in the catalytic cycle. The K_m of fluoride is high (10 mM), and increases in the presence of high SAM **34** concentrations (47 mM in the presence of 300 µM SAM **34**[87]) which indicates competitive binding at the active site. It is suggested that fluoride ion will passively diffuse into the active site of fluorinase, and upon binding of SAM **34 (**K_m= 6.5 µM[87]) it becomes trapped. SAM **34** binding also squeezes out any remaining water in the active site, leading to the full desolvation of the fluoride ion. In the reverse direction, it was discovered that 5'-FDA **35** binds first, and upon this event the binding site for L-methionine is formed through reorganization of the protein, specifically residues Thr75 to Arg85 and consequently Ala95 to Glu102, located on the 20 amino acid loop determined in the X-Ray structure. [87]

1.5.2.3 Site directed mutagenesis of the fluorinase

Site directed mutagenesis of the fluorinase has been useful for exploring the roles of individual residues of the active site during the mechanism of catalysis. These methods, based upon analysis of the crystal structure and QM/MM calculations have established the putative hydrogen bonding networks important for catalysis and the integrity of the

active site pocket. Dr Xiaofeng Zhu (University of St Andrews) generated several site specific mutants, with an interest in determining the mechanistic roles of specific residues at the active site.[87] Four residues were identified from the crystal structure of the fluorinase that were thought to be critical to its catalytic activity, Ser158, Thr80, Phe156 and Asp16.

1.5.2.4 Serine 158 fluorinase mutant

The role of Ser158 has been discussed previously and a mutant possessing a glycine residue at this position exhibited only 8% activity compared to the native enzyme.[87] The crystal structure of this mutant revealed that a water molecule had take the place of the OH side chain of S158, suggesting that this residue is critical in the desolvation of fluoride ion. Ser158 was also mutated to an alanine, to generate the mutant S158A. Alanine is a non-polar amino acid, with a lipophilic methyl group. This mutant exhibited 38% activity.[87] In both of these mutants, disruption to the hydrogen bond network has a significant effect upon the catalytic activity of the fluorinase, but the more lipophilic alanine presumably promotes desolvation over the less lipophilic glycine, and hence was a more efficient catalyst.

1.5.2.5 Threonine 80 fluorinase mutant

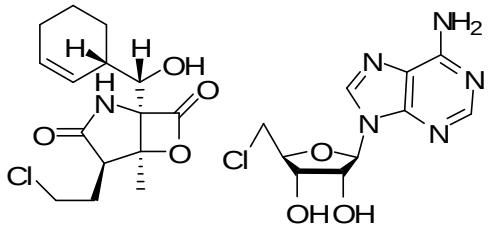
The threonine residue at position 80 (Thr80) in the fluorinase lines the fluoride binding pocket. Theoretical studies carried out by W. Thiel and H. Senn in Mülheim suggest that H bonding between a tyrosine at position 157 and the side chain OH of Thr80 (1.83Å)

stabilizes the pocket (Figure 1.9).[77] Substituting this residue with alanine (Thr80A mutant), a non-polar residue with a CH_3 side group, eliminates this H bond interaction and diminishes the enzyme's activity (15% activity). When Thr80 was however exchanged for a serine then the mutant retained almost full activity (95 %). Notably the integrity of the important hydrogen bond has been maintained in converting Thr80 to serine.

1.5.2.7 Structural homologs and the origins of the fluorinase

1.5.2.7.1 The chlorinase

In 2007, the gene for an enzyme was discovered from the marine actinomycete *Salinispora tropica* in a gene cluster responsible for the biosynthesis of the chlorinated natural product, salinisporamide A **36**.[88]



36 37

The first committed step in the biosynthesis of **36** is through the action of a SAM-dependant chlorinase, *Sal*L, generating 5'-chlorodeoxyadenosine **37**.[89] Not only is the halogenation mechanism analogous to the fluorinase, the chlorinase also shows 35% amino acid identity to the fluorinase enzyme and exhibits identical characteristics in its

tertiary structure. Closer analysis of the active site reveals that similar residues are required for SAM **34** coordination. The way in which the halide is bound differs somewhat. Notably a glycine residue (Gly131) replaces Ser158 and a tyrosine residue (Tyr70) replaces Thr80 in the fluorinase. The chlorinase is incapable of fluorination, unlike the fluorinase, which also accepts chloride ion as a substrate.[90] Crystallography has revealed that chloride ion sits in the active site at approximately 180º from the C-5' carbon of SAM **34**, consistent with an S_N2 mechanism, as exhibited by the fluorinase. Currently it is presumed that the reaction mechanism of halide ion binding and desolvation is similar to the fluorinase.

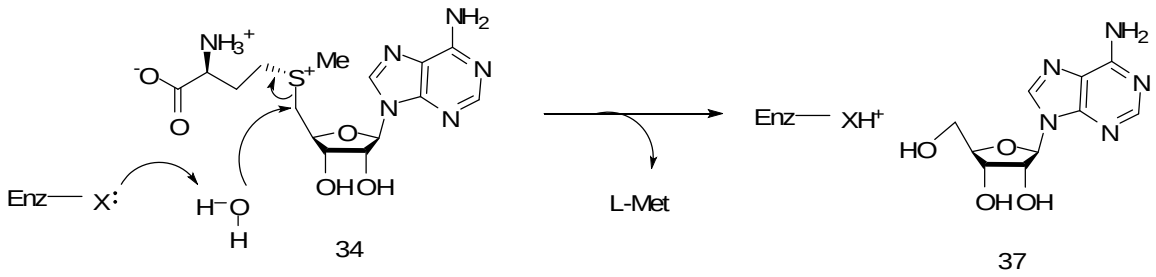
1.5.2.7.2 The Duf62 Superfamily

A BLAST search of the fluorinase amino acid sequence reveals homology (25-38%) with a range of proteins named the domains of unknown function-62 (duf-62), identified in a series of genome sequencing projects but as the name suggests their function was yet to be revealed. These proteins are generally localized in extremophile and pathogen-related microorganisms. Four of these enzymes have been subjected to X-ray crystallography studies and recently the duf-62 from the deep sea vent-dwelling bacterium *Pyrococcus horikoshii* OT3 was identified as a SAM-dependant hydroxide adenosyltransferase.[91] Duf-62 from *P. horikoshii* is thermostable at 80 ºC and has a K_m for SAM **34** of 39.2 µM and a k_{cat} of 0.14 s^{-1} (similar to the fluorinase of 0.07 s^{-1}). The X-ray structure of this enzyme (PDB 1WU8) reveals remarkable similarities to both the chlorinase and fluorinase enzymes (Figure 1.10), also consisting of a trimer made up of three identical

monomers. The X-Ray structures also reveal an adenosine bound to the protein at the subunit interfaces, clearly identifying the active sites of these proteins.

Figure 1.10. X-ray crystal structures of the fluorinase (blue)[82], chlorinase (pink) [88] and duf-62 from *P. horikoshii* OT3.

Despite the structural similarities with these halogenating enzymes, duf-62 is incapable of performing fluorination or chlorination reactions in the presence of high halide ion concentrations and SAM **34**. However it is able to catalyse the conversion of SAM **34** to produce adenosine and H^+. Labelling studies using $^{18}OH_2$ and GC-MS analysis revealed that activated water was used as a nucleophile to attack the electrophilic C-5' carbon.[91] (Scheme 1.13).



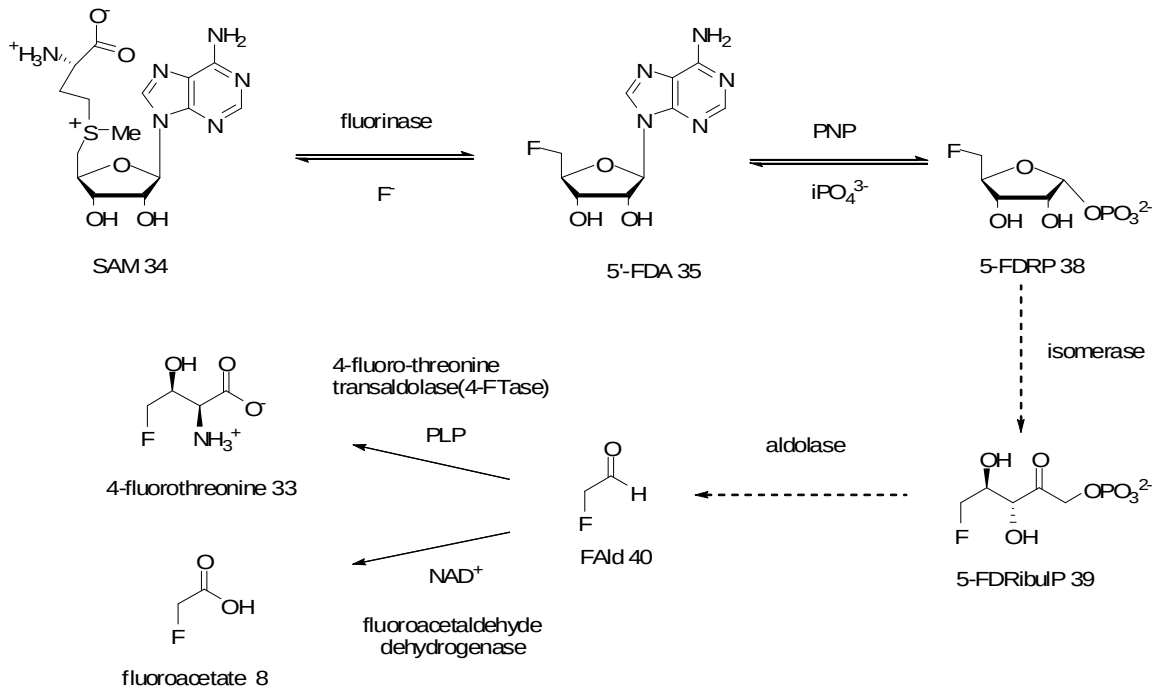
Scheme 1.13. Most likely mechanism for the duf-62 enzyme from *P. horikoshii*, deduced by $^{18}OH_2$ labelling and GC-MS analysis.[91]

The substitution by hydroxide ion, generated from activated water is probably an S_N2 process in a similar manner to fluorinase. There are three conserved amino acids in the duf-62 proteins that are not present in the chlorinase or the fluorinase (Asp68, Arg75 and His127). These residues are H-bonded together in a triad and may be involved in the catalytic cycle.[91]

The function of the duf-62 proteins is still unclear, however it may be involved in a sensitive regulation of pH as it produces one H^+ for every reaction cycle. It also has an optimal pH of 8.5 and is completely inactivated at pH 5, suggesting a regulatory role for this protein. It appears from amino acid sequence homologies that these duf-62 proteins are relatives of the halogenating fluorinase and chlorinase enzymes, although the active sites evolved in different directions from their ancestors.

1.5.2.8 The metabolic fate of 5'-FDA in *S. cattleya*

The fluorinase enzyme from *S. cattleya* is the first committed step in the production of the fluorometabolites 4-FT **33** and FAc **8**.[75-77] Previous work in our research group has revealed some of the biosynthetic steps and intermediates in the pathway to fluorometabolite production. Scheme 1.14 below reveals the status of the pathway when this project started.



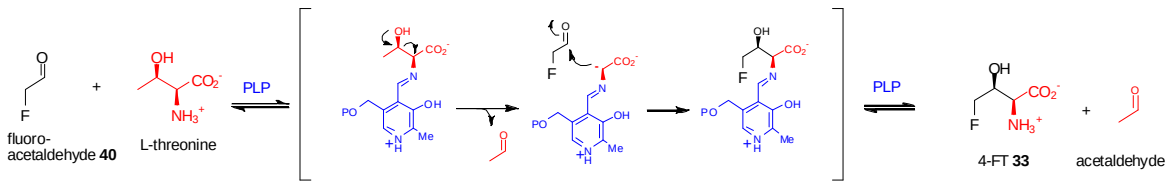
Scheme 1.14. The fluorometabolite pathway in *S. cattleya*, enzymes and intermediates, at the beginning of this project. Broken lines indicated steps which had yet to be characterized.

Following the generation of 5'-FDA **35** from SAM **34** and fluoride ion by the fluorinase, the next transformation in the fluorometabolite pathway is carried out by a purine nucleoside phosphorylase (PNP)[92], which converts 5'-FDA **35** to 5-fluoro-5-deoxyribose-1-phosphate **38** (FDRP).[93] The gene, *FlB*, responsible for this enzyme is located directly alongside the fluorinase in the gene cluster identified by J. Spencer and co-workers. (Figure 1.5).[83] PNPs catalyse the reversible phosphorolysis by inorganic phosphate and the glycosidic bond of purine ribo- and deoxyribonucleosides to generate the free purine and a (deoxy)ribose sugar.[94] The *FlB* gene encodes a protein of 299 amino acids with a molecular weight of ~36 KDa. A BLAST search reveals that this PNP belongs to a family of 5'-methylthioadenosine phosphorylases (MTAPs), a key component of the L-methionine salvage pathway, discussed in Chapter 2. Over expression of this enzyme in *E. coli* was achieved, although the protein formed was largely insoluble.[95]

The sugar-phosphate, 5-FDRP **38** has been isolated and characterized from partially purified cell free extracts of *S. cattleya*.[96] A further intermediate, 5-fluorodeoxyribulose phosphate **39** (5-FDRulP) has also been identified from cell free extracts of *S. cattleya*.[97] It is proposed that this is the next intermediate in the biosynthetic pathway, generated from **38** by an aldose-keto isomerase (Scheme 2.10, Chapter 2), also involved in the L-methionine salvage pathway. Sugars such as 5-FDRulP **39** are well known as products of aldolases, particularly fuculose aldolases. These enzymes are also capable of utilizing these sugars as substrates to generate dihydroxyacetone phosphate (DHAP) and an associated aldehyde. In this case FAld **40** is formed as the last common intermediate in 4-FT **33** and FAc **8** biosynthesis.

1.5.2.9 The 4-fluorothreonine transaldolase (4-FTase) gene from *S. cattleya*

The final step in 4-FT biosynthesis in *S. cattleya* involves a pyridoxal phosphate (PLP) dependant transaldolase enzyme that mediates a cross-over reaction between L-threonine and fluoroacetaldehyde **40** to yield 4-FT **33** and acetaldehyde. The few bacterial PLP threonine aldolases that have been identified to date utilize acetaldehyde and glycine in a direct condensation reaction.[98] The *S. cattleya* enzyme does not utilize glycine but instead carries out a mechanistically more elaborate reaction (Scheme 1.15).[96]



Scheme 1.15. Minimal mechanism of the PLP-transaldolase involved in the final step of 4-FT **33** biosynthesis in *S. cattleya*.[99]

The generation of 4-FT **33** by the PLP-transaldolase requires the presence of the amino acid L-threonine, the co-factor PLP and the fluorinated intermediate FAld **40**.[99] The gene for this 4-FT transaldolase enzyme is not in the 10kb *flA* gene cluster and this gene was identified through a reverse genetic approach after purification and N-terminal sequencing of the wild type 4-FT transaldolase by Dr Hai Deng (University of St Andrews).[99] Trial and error PCR and subsequent chromosomal gene walking identified a 2.2 kbp DNA sequence, which contained a complete open reading frame (*FTase*) of ~1.9 kbp (1905 bp).[99]

The *FTase* coded for a 634 amino acid protein composed of two domains (Figure 1.11). The larger domain (440 amino acids) is homologous to the PLP binding domain of serine hydroxylmethyl transferase (SHMT) enzymes in micro-organisms such as archaea and thermophilic bacteria (~35 % amino acid identity). The smaller domain (145 amino acids) has homology with the phosphate binding domain of bacterial ribulose-1-phosphate-4 epimerases (*araD*) or L-fuculose aldolases (>28%). A region of 35 amino acids between the SHMT-like and araD-like domains appears to act as a linker, perhaps bearing no catalytic function. The PLP transaldolase appears to have a hybrid construction with key binding motifs from these enzymes. Enzymes from the *araD*

superfamily catalyse reversible aldol/retro-aldol carbon-carbon bond cleavage, often resulting in epimerization,[100, 101] which is similar in nature to the PLP-transaldolase reaction.

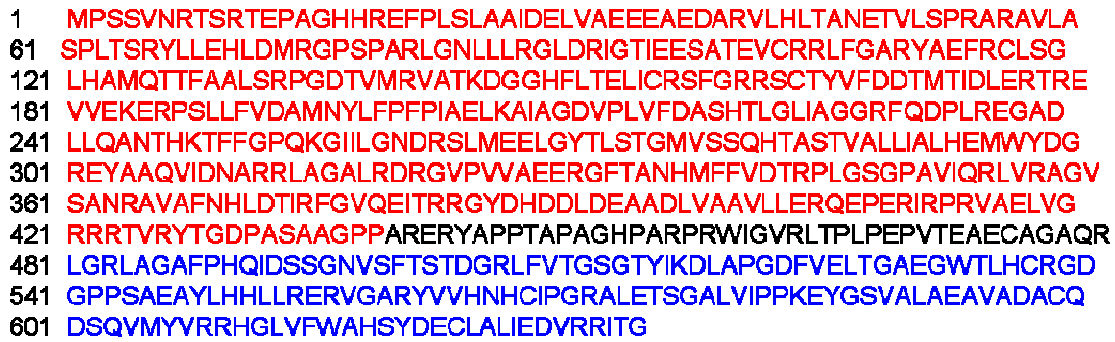


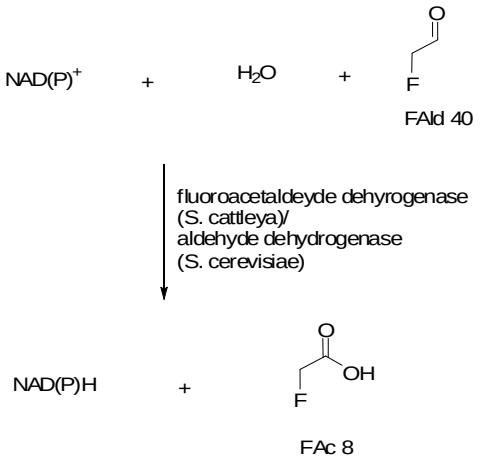
Figure 1.11. The full amino acid sequence of the PLP-transaldolase from *S. cattleya* with three main domains. Red= putative SHMT-PLP binding domain. Black= putative linker domain and Blue= putative phosphate binding domain of bacterial epimerase/aldolase.[99]

Expression of the 4-FT transaldolase in *E. coli* lead to the generation of insoluble inclusion bodies. Therefore in order to generate an active protein, the gene was inserted into *Streptomyces lividans* on an *E. coli*:*Streptomyces* shuttle vector, enabling over expression and purification of this enzyme by affinity chromatography.[99]

1.5.2.10 A role for aldehyde dehydrogenase in the fluorometabolite pathway of *S. cattleya*

The enzyme responsible for the oxidation of FAld **40** to FAc **8** in the fluorometabolite pathway of *S. cattleya* has been attributed to an NAD^+-dependant aldehyde dehydrogenase (E.C.1.2.1.69). The activity of this aldehyde dehydrogenase observed in the prepared CFE of *S. cattleya* after 4 days growth.[102] This suggests that this enzyme is a part of secondary metabolism, committed to the fluorometabolite pathway and not

performing some other function. The gene for this enzyme has yet to be characterized and it was not found in the 10 kb gene cluster. However it was discovered that a homologous enzyme from yeast was capable of carrying out the same reaction, albeit with a K_m, 6 fold higher.[102] The NAD^+-dependant aldehyde dehydrogenase from *Saccharomyces cerevisiae* (E.C. 1.2.1.5) is commercially available, and a minimal scheme for the reversible oxidation of FAld **40** to FAc **8** in the presence of $NAD(P)^+$ and water is shown in Scheme 1.16.



Scheme 1.16. A minimal scheme for the irreversible oxidation of FAld **40** to FAc **8** by aldehyde dehydrogenase from *S. cerevisiae*, in the presence of $NAD(P)^+$ and water.[102, 103]

1.5.2.11 Application of the fluorinase: Positron emission tomography

Positron emission tomography (PET) is a non-invasive imaging technique used for medical imaging and diagnostics.[104] The technique uses radiotracers labelled with positron emitting radionuclides with various *in vivo* properties that permit imaging of the distributions of binding ligands that have been taken up into metabolising tissues. The most common PET radionuclides are ^{11}C, ^{18}F, ^{15}O and ^{13}N which have half lives of 20,

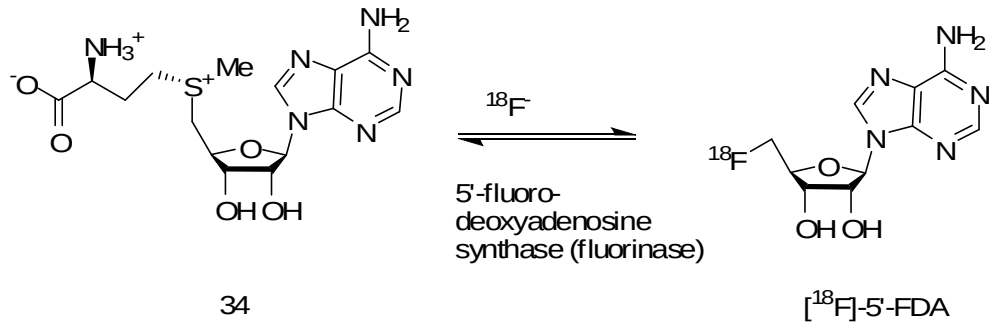
110, 2 and 10 minutes respectively. Fluorine-18, with a relatively long half life ($t^{1/2}$= 110 m), permits more time for radiochemical synthesis and purification for use in *in vivo* experiments and therefore is an attractive radionucleotide for PET. [^{18}F] Fluoride is generated in a cyclotron in very high specific activity (~ G Bq's), without the need for a cold carrier ([^{19}F] fluoride) to be added.[105]

The most common radiotracer in this arena is [^{18}F]-labelled 2-fluorodeoxyglucose (FDG) which has routinely been used in brain and tumour imaging.[106] The discovery of specific adenosine[107] and uridine[108] receptors in the brain also increases the significance of these compounds with regard to neurological imaging. This has led to extensive studies using radiolabelled nucleosides such as the adenosine analogue, 2'-fluoro-2'-deoxyadenosine and preliminary work with this compound has shown promising results in the evaluation of tumour cell proliferation.[109]

1.5.2.11.1 PET-labelled production of fluorinated metabolites

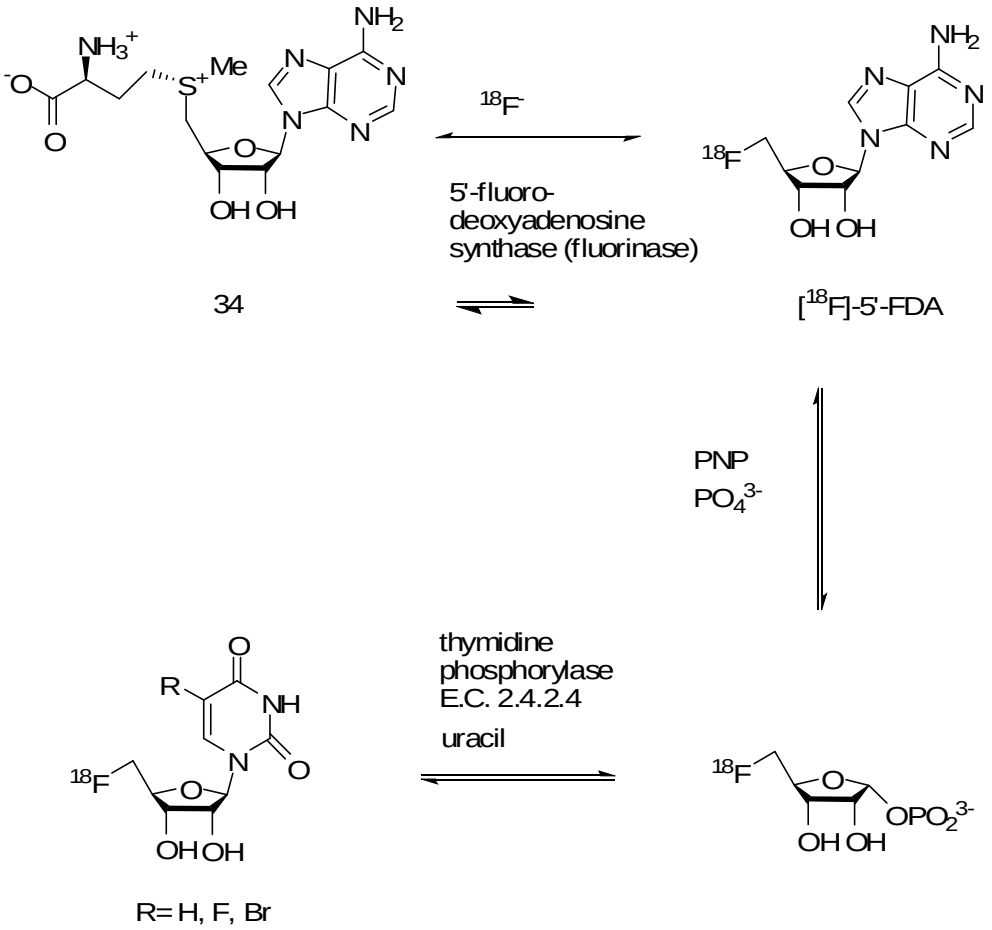
The radioactive isotope [^{18}F] can be incorporated into 5'-FDA **35** by the fluorinase enzyme (Scheme 1.17). Two [^{18}F] compounds have been synthesised that are also intermediates in the fluorometabolite pathway of *S. cattleya*, [^{18}F]-FAld **40**[110] and [^{18}F] 5'-FDA **35**.[111] [^{18}F]-FAld **40** was generated for use as a [^{18}F]-fluoroethylating agent and synthesis of the adenosine analogue [^{18}F]-5'-FDA **35** was achieved with only a ~1% radiochemical yield (RCY).[111] With over expressed fluorinase however, the enzymatic method towards radiolabelled [^{18}F]-5'-FDA **35** was achieved with RCYs of up to 95%.[112]

This was accomplished by first coupling the fluorinase to an L-amino acid oxidase, which removed the co-substrate L-methionine and also prevented the reverse reactions.



Scheme 1.17. Production of [^{18}F]-5'-FDA by fluorinase.[112]

Secondly a deaminase was added to produce a second labelled purine nucleoside, [^{18}F]-5'-fluorodeoxyinosine **36**. Most recently, fluorinase and various nucleotide phosphorylase-coupled base swap experiments have been carried out. This involved the biocatalytic removal of the adenosine base of [^{18}F]-5'-FDA **35** and then utilising the reversible nature of nucleotide phosphorylases to generate nucleoside analogues with uracil bases (Figure 1.18).[113]



Scheme 1.18. [^{18}F]-5'-Fluorodeoxyuracil derivatives generated by fluorinase-coupled base swapping experiments.[113]

There appears to be potential for using the fluorinase to produce [^{18}F]-radiolabelled nucleosides for PET imaging of certain cancers, and other fluorinated compounds. Aggressive cancers can take up these fluorometabolites through the purine nucleotide salvage pathway, due to the high demand for energy to fuel growth and replication. [^{18}F] through PET imaging offers a sensitive visual technique for cancer diagnosis and treatment by deploying suitable tracers e.g. [^{18}F]-FDG.

1.6 Analytical Methods

1.6.1 ^{19}F NMR spectroscopy

^{19}F NMR is used in this thesis as an analytical tool for identifying the production of fluorometabolites and intermediates generated by enzymes identified from *Streptomyces cattleya*. This technique allows the identification of fluorinated products without the need for isolating the metabolite. Coupling of fluorine (^{19}F, I= 1/2) with hydrogen (^{1}H, I= 1/2) allows the chemical environment of the fluorine to be determined (Figure 1.12). Figure 1.13 shows the ^{19}F-NMR spectras of two fluorinated secondary metabolites, fluoroacetate **8** and 4-fluorothreonine **33** from *S. cattleya*.

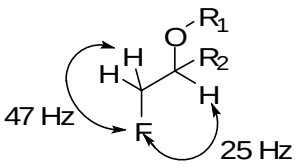


Figure 1.12. Typical *J* coupling of an organo-fluorine compound by ^{19}F NMR.

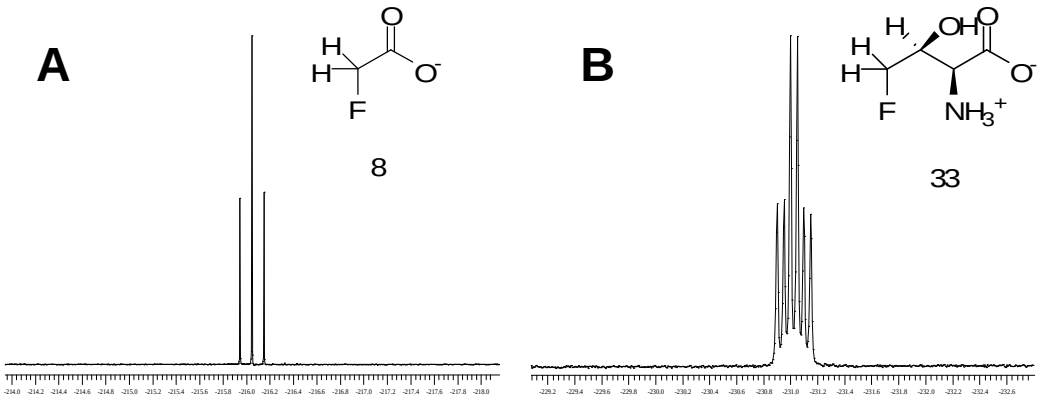


Figure 1.13. ^{19}F NMR spectra of (**A**), fluoroacetate **8** and (**B**), 4-fluorothreonine **33**.

The ^{19}F NMR spectrum (Spectrum A above) of FAc **8**, reveals that the fluorine of the fluoromethyl group is a triplet. The corresponding spectrum **B** for 4-FT **33** is a doublet of doublets of doublets arising from coupling to the non-equivalent methylene protons and then the vicinal methine-proton.

Intermediate	^{19}F NMR Chemical Shift (ppm)
5'-FDA **35**	-231.5
5'-FDRP **38**	-231.3
5'-FDRulP **39**	-231.8
FAld **40**	-231.45
FAc **8**	-217.4
4-FT **33**	-231.2
4-FDI **36**	-231.35

Table 1.2. ^{19}F NMR chemical shifts of the identified intermediates of the fluorometabolite pathway in *S. cattleya*.

The chemical shifts of FAc **8** (-217.4 ppm) and 4-FT **33** (-231.2 ppm) are easily distinguishable, as are the other fluorinated metabolites on the biosynthetic pathway (Table 1.2).

1.6.2 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a sensitive technique used to determine intermolecular recognition and binding of ligands to macromolecules such as proteins. It can be used to determine the thermodynamic relationships of ligand binding, giving quantitative values of free energy change (ΔG), enthalpy (ΔH) and entropy (ΔS) which allows the accurate characterization of binding events. The K_A, and stoichiometry of binding can also be effectively measured by this technique.

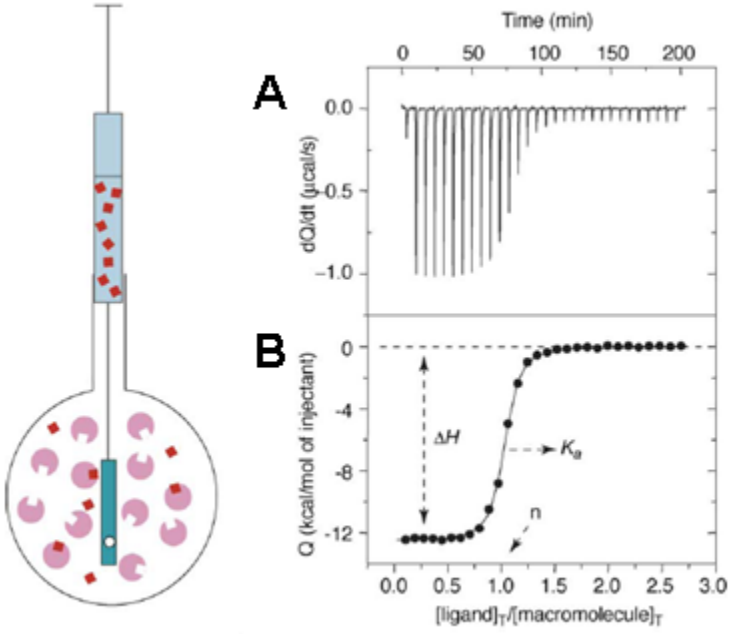


Figure 1.14. Schematic diagram of an ITC sample cell, and typical results of an ITC experiment. A= Raw data power difference upon ligand titration. B= Raw data converted according to molecular concentration of the titrant and titrand.[115]

ITC is a direct method used to measure the heat change on formation of a binding complex at a constant temperature.[114] Figure 1.14 shows a schematic of an ITC experiment.[115] At a constant temperature, either the ligand or the macromolecule of

interest (the 'titrant') is titrated into a solution containing its putative binding partner (the 'titrand'). The titrand is contained within a cell consisting of a highly efficient thermal conducting material (e.g. gold). There is also a reference cell which contains sample buffer. Both of these cells are surrounded by an adiabatic jacket and are maintained at a constant temperature set by an operating computer terminal. The temperature of the cells is measured as a value of power ($J\ s^{-1}$) required to maintain the cells at a constant temperature. Typically a small volume of the titrant is injected into the sample cell in stepwise manner using a fixed volume. The injections are separated in time (~180 s) to allow for the heat of binding to be measured for that particular injection event. Upon binding of the titrant to the titrand, heat is released or absorbed depending on the binding relationship. This heat change effects the amount of power ($J\ s^{-1}$) required to maintain a constant temperature in the sample cell. The difference in power required between the sample and reference cells is measured, and converted to produce quantitative values for the binding properties listed above.

ITC has been used extensively in drug discovery, allowing for the screening of compounds against protein targets.[115] Recently this technique has also been successfully used to determine the binding parameters and mechanisms of the fluorinase, providing interesting insights into the binding order of substrates and products.[87] ITC is used in this thesis to determine putative binding of several molecules to protein targets.

1.7 Conclusions and project aims

The fluorometabolite pathway of *S. cattleya* has been well characterized with several intermediates and enzymes identified. Two putative enzymatic steps remained to be characterized at the enzymatic level when this project began; that for the conversion of 5-FDRP **38** to 5-FDRulP **39** by an isomerase and that of the conversion of 5-FDRulP **39** to FAld **40** by the action of an aldolase. The identification of these enzymes became the initial research focus of this project.

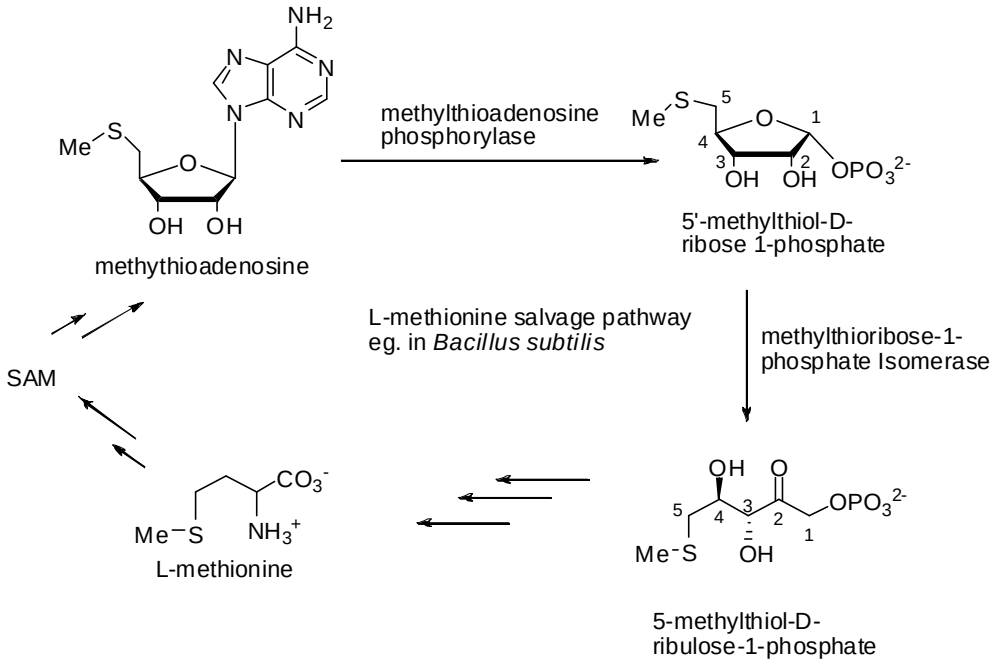
2 The identification of an isomerase from *S. cattleya*

The identification of 5-FDRP **38**[93]and 5-FDRulP **39**[97] as intermediates in the fluorometabollite biosynthesis pathway of *S. cattleya* revealed some interesting similarities with the SAM-derived L-methionine salvage pathway. It was postulated that 5-FDRulP **39** is generated by an isomerisation directly from 5-FDRP **38**. The SAM-methionine pathways in bacteria[116,117,118,119,120,121,122] and yeast[123] are very similar to that observed in fluorometabolite biosynthesis in *S. cattleya*. In this chapter the identification of two enzymes from *Streptomyces*, capable of performing the isomerisation of 5-FDRP **38** to 5-FDRulP **39** is described.

2.1 Methionine salvage pathway

The methionine salvage pathway was first identified in the bacteria *Klebsiella pneumoniae*[116-120] and then subsequently in *Bacillus subtilis*.[121, 122] L-Methionine is an essential amino acid, and is involved in many cellular functions such as the initiation of protein synthesis, methylation of DNA and rRNA as well as the biosynthesis of cysteine, phospholipids and polyamines. Energetically, the *de novo* synthesis of L-methionine is expensive, and consequently the L-methionine salvage pathway has developed as a key

process for the recycling of both L-methionine and sulphur for redistribution within the cell.[123] Identified intermediates in the L-methionine salvage pathway of both *K. pnuemoniae*[118] and *B. subtilis*[121] include 5'-S-methylthioadenosine (MTA), methylthioribose 1-phosphate (MTR-1-P), with methylthioribulose-1-phosphate (MTRul-1-P) generated directly downstream (see Scheme 2.1).[116-122] These intermediates are analogous to intermediates of fluorometabolite biosynthesis.



Scheme 2.1. Illustration of the intermediates and enzymes from the L-methionine salvage pathway identified from *B. subtilis* and *K. pnuemoniae* with particular focus on the compounds analogous to those involved in the fluorometabolite pathway from *S. cattleya*.[116-123]

2.2 Methylthioribose isomerases (MTRIs)

The enzymatic step responsible for the generation of MTRul-1-P from MTR-1-P in the methionine salvage pathway is an aldose-ketose isomerase.[116,122] These enzymes were

identified from *B. subtilis* (E.C. 5.3.1.23) and the YPR118W gene from the yeast, *Saccharomyces cerevisiae*.[121] A BLAST search using the sequence of the methylthioribose-1-phosphate isomerase (MTRI) of *B. subtilis* reveals homology with a number of genes belonging to the PFAM family PF01008[124] and the TIGR 00512 and 00524 families (See Figure 2.1).

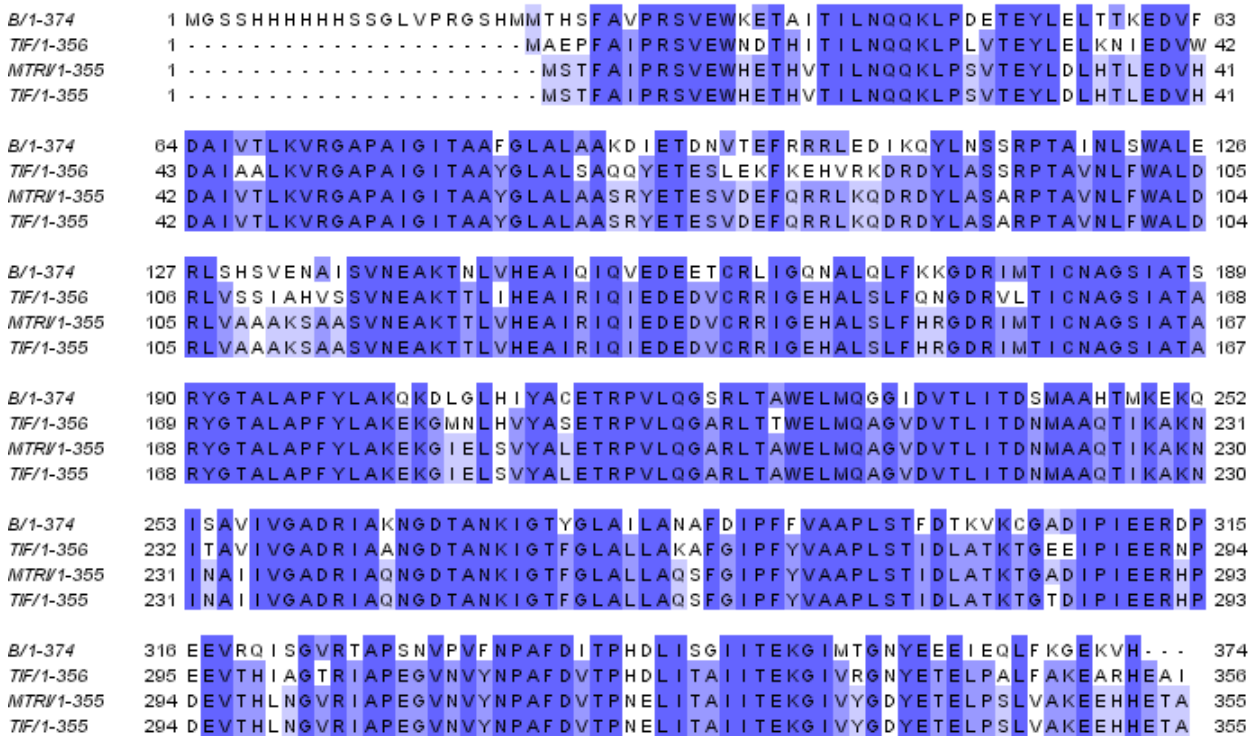


Figure 2.1. Selected sequences from a BLAST search of the MTRI from *B. subtilis*. Sequences shown are from *B. subtilis MTRI*, Geobacillus sp WCH70 Translation initiation factor (TIF), *Geobacillus thermodenitrificans* MTRI and Geobacillus sp G11MC16 TIF respectively. Regions of homology are highlighted in blue.

Genes belonging to PF01008 contain the α-, β-, and δ- subunits of eukaryotic initiation factor 2B (eIF2B) found in yeast and mammals.[124] However they lack the ε-subunit, which is responsible for the catalytic activity of these eIF2Bs. In eukaryotic translation initiation, heterotrimeric eIF2 acts in the presence of GTP and Met-tRNA to interact with

the 40S ribosomal subunit and other initiation factors to form the 43S preinitiation complex. This complex is responsible for the binding of the 5'-end of mRNA, before scanning in the 3'-direction for the AUG initiation codon. AUG is identified by codon-anticodon recognition through GTP hydrolysis which is triggered by another initiation factor, eIF5. The subsequent complex is necessary for the formation of the 80s initiation complex through the binding of the large ribosomal subunit and the release of translation initiation factors from the 40S subunit, resulting in translation initiation. eIF2-GDP is released, eIF2B then binds and catalyzes the exchange of GDP for GTP so that the cycle can start again.

As well as encompassing the well characterised eIF2B proteins, the PF01008 family also includes a subfamily of proteins known as the eIF2B-related proteins. This subfamily has homology with the eIF2B proteins and is known to exist in eukaryotes, archaeae and eubacteria however their roles in these organisms are largely unknown. The existence of these eIF2B-like proteins in non-eukaryotes is particularly intriguing, as there is no evidence that implicates IF2 proteins in prokaryotic translation initiation. Many of these proteins have been annotated as putative translation initiation factors based on homology with proteins of this function. However recent studies have identified enzymes from *B. subtilis*[121] and *Saccharomyces cerevisiae*[123] as MTR-1-P isomerases (MTRIs), an enzyme with 37% sequence identity to that involved in the methionine salvage pathway. Enzymes of this type are capable of converting MTR-1-P to MTRul-1-P.[121-123]

2.2.1 MTRI crystal structures

Two MTRIs have recently had their crystal structure elucidated, YPR118W from *S. cerevisae*[123] and the MTRI from *B. subitilis*.[125] Only the MTRI from *B. subtilis* has been crystallized with a substrate bound to the active site, in the form of MTRul-1-P (PDB 2YVK).[125] This has revealed several candidates for the key catalytic residues at the active site, and also gives some clues on the mechanism of these enzymes. The residues that are putatively involved in substrate binding can be separated into three groups: Those that hydrogen bond with **1)** the phosphate group, **2)** the backbone oxygens and hydrogens of ribulose and **3)** those which have hydrophobic interactions with the methylthio group.

The residues within reasonable H-bonding distance of the phosphate moiety are the side chains of Arg51 (3.37 Å), Arg94 (2.72 Å), Gln199 (3.27 Å), Lys251 (2.96 and 3.11 Å) and the backbone amide of Gly52 (3.16 Å) (Figure 2.2B). The Asn250 (3.43 Å) and Asp240 (3.99 Å) residues probably interact with the C2 carbonyl of ribulose. The C3 hydroxyl is coordinated by C160 (3.61 Å) and to water a molecule located in the active site, and the OH of C4 with Asp240 (2.42 Å) and the backbone amide of Ala53 (3.01 Å) (Figure 2.2C).

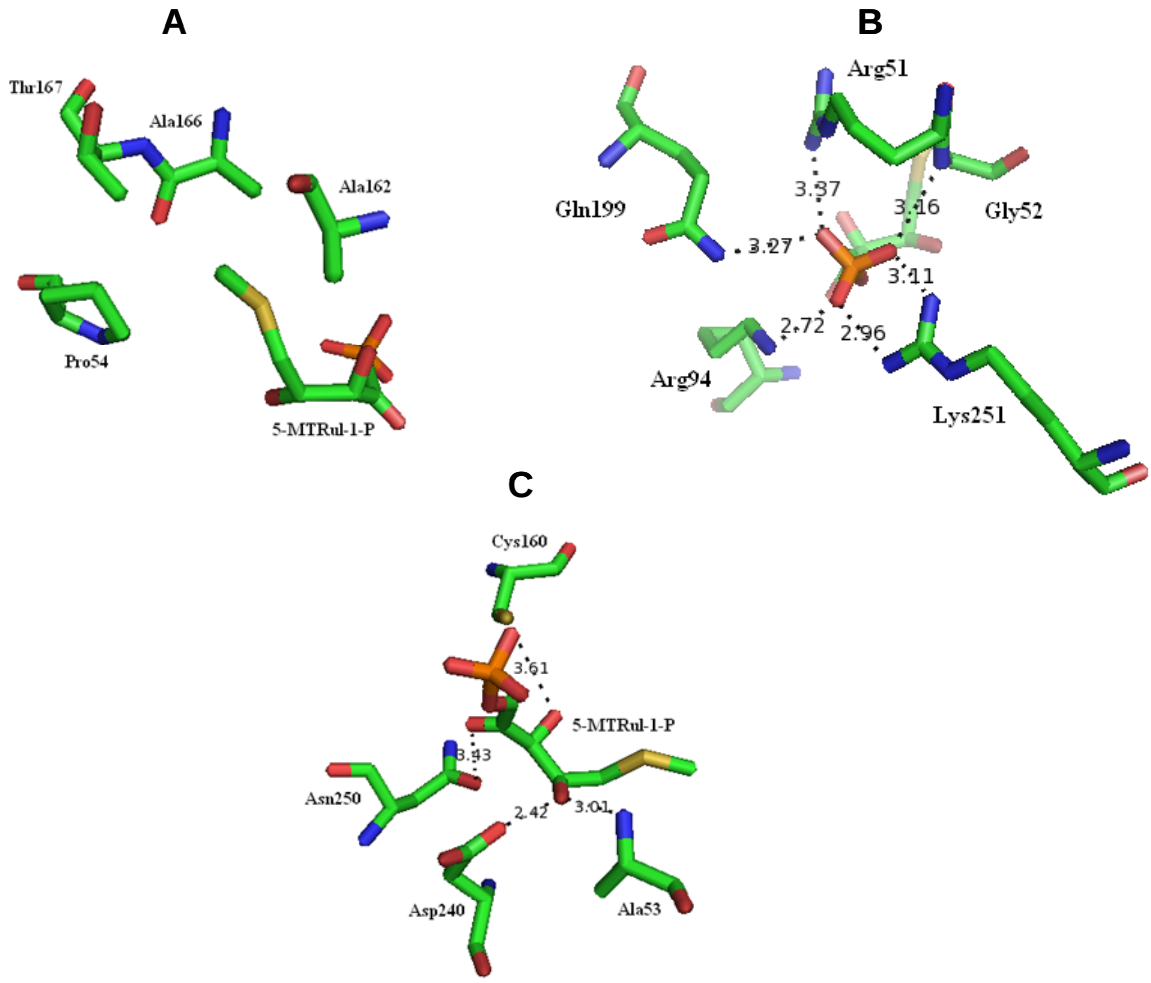
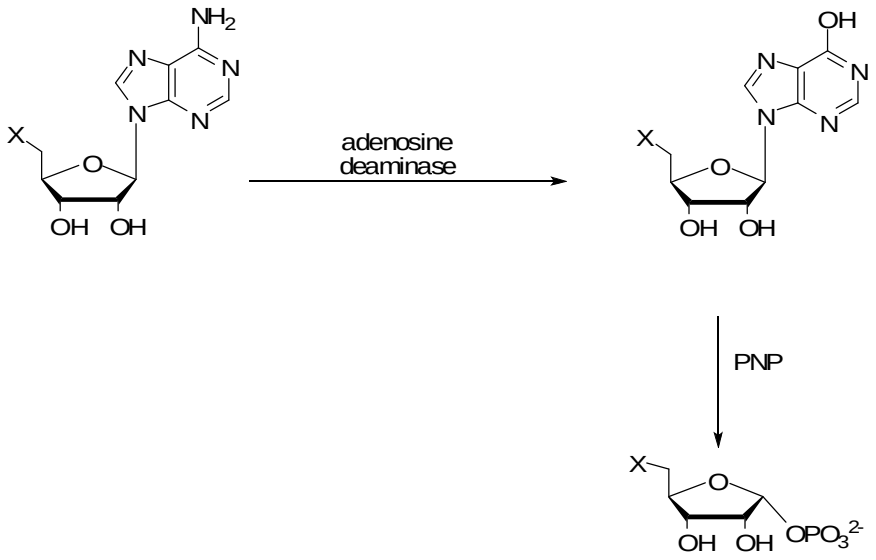


Figure 2.2. The crystal structure of the putative active site from the MTRI of *B. subtilis* with 5-MTRul-1-P bound.[125] **A**= Hydrophobic interactions with the methylthio group. **B**= Putative hydrogen bonding to the phosphate group. **C**= Putative hydrogen bonding to the backbone of ribulose. All distances measured are in Å.

The side chains of Pro54, Ala162, Ala166 and Thr167 are thought to exhibit hydrophobic interactions with the methylthio group (Figure 2.2A). All of these residues are highly conserved amongst MTRIs from various species, apart from three residues. Ala162 is commonly substituted for threonine, Ala166 for valine and Thr167 for serine. Although they are not absolutely conserved, the residues with which they are substituted exhibit similar structural properties and therefore interact similarly at the active site.

The crystal structure of an MTRI from *S. cerevisiae* revealed the monomeric structure possesses two main domains, the N- and C- terminals.[123] The N-terminal domain, consists of 138 residues and shows structural homology to the pollen allergen phl 6 (PDB entry 1NLX) and ATP Synthase subunit C (PDB entry 1C17).[126] The C-terminal domain has structural homology to ribose-5-phosphate isomerase from *E. coli* (PDB entry 1LKZ).[127] This evidence suggests that proteins that possess similar structures are capable of the isomerisation of sugar-phosphate moieties. The active MTRI protein from both *B. subtilis* and *S. cerevisiae* are homodimers, which has been determined by gel filtration and X-ray structure studies. These eIF2B proteins are monomeric which is a further difference between MTRI proteins and eIF2Bs.

Despite the identification of a number of MTRI's, kinetic data is rare. Quantitative kinetic data is difficult to produce due to problems with the synthesis of MTR-1-P. Currently the only route to MTR-1-P is through the conversion of commercially available 5'-MTA by a deaminase to 5'-methylthioriboseinosine (5'-MTI) and then phospholytic cleavage of the base by a purine nucleotide phosphorylase (PNP) to generate 5-MTR-1-P (Scheme 2.2). The same approach can be used to generate 5-FDRP **38** using synthetic 5'FDA **35** as a starting material.[97]

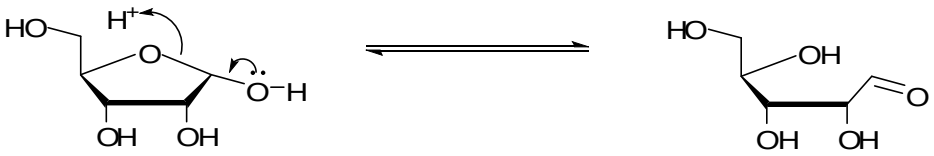


Scheme 2.2. The generation of the substrate of MTRIs. X= F[97], SMe.

5'-MTA, 5-MTR-1-P and 5-MTRul-1-P have all been characterized by ^{1}H NMR[116], and a method for the colorimetric assay of reducing sugars[128] can be modified to quantitatively measure the isomerization reaction.[129] So it is possible to analyze the isomerization with a view to attaining kinetic data, providing that you can generate enough substrate.

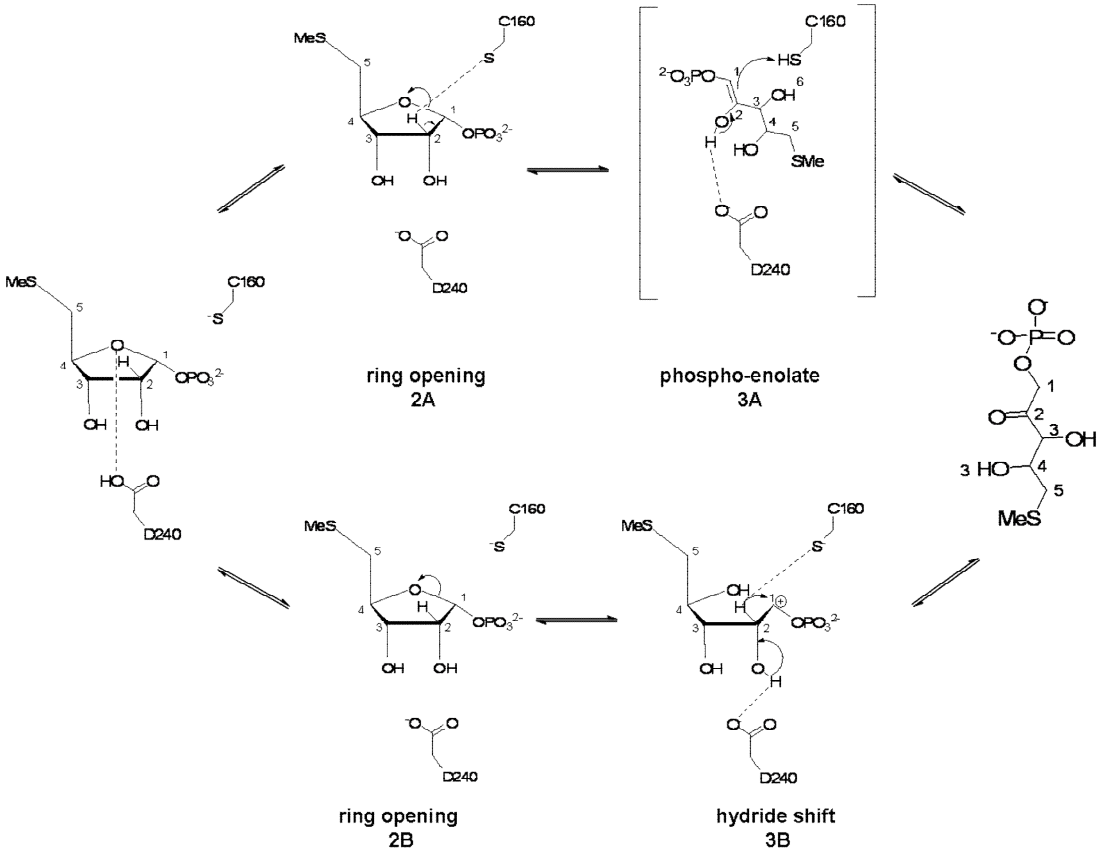
2.2.2 Putative mechanisms for MTRIs

The catalytic mechanism of MTRI's has yet to be elucidated. It is well known that in sugars with a free anomeric OH (hemiacetals), the sugar equilibriates between the ring-opened aldehyde and the ring closed hemiacetal (Scheme 2.3). However the presence of the phosphate group at the 1-position stops such ring-opening.



Scheme 2.3. Classical ring opening of a 5-membered sugar.

MTRIs belong to the aldose-ketose isomerases, which catalyse the isomerisation of its cyclic substrate.[122] Crystallization of the MTRI from *B. subtilis* with substrate bound at the active site, has triggered discussion of possible mechanisms for this reaction. Two putative mechanisms for aldose-ketose isomerases have been put forward.[130] These are the cis-enediol and hydride transfer mechanisms. Analysis of the active site structure of MTRI from *B.* subtilis does not discriminate against either of these putative mechanisms (Scheme 2.4).[125]



Scheme 2.4. Proposed reaction mechanisms of MTRIs based on the active site residues from the crystal structure of an MTRI from *B. subtilis*. Reaction steps 2A and 3A represent the proposed cis-enediol mechanism. Steps 2B and 3B represent the proposed 1,2-hydride shift mechanism.[125]

2.2.2.1 Cis-enediol mechanism

The putative mechanisms in Scheme 2.4 begin with binding of the phosphate group to the positively charged region generated by the side chains of Arg51, Arg94 and Lys251. This triggers a conformational change in the MTRI structure, isolating the active site and substrate from solvent. The side chain of Asp240 is expected to play a role as either a proton donor or acceptor. Protonation to the ring oxygen of MTR-1-P would trigger the ring opening. In the cis-enediol mechanism (steps 1 to 2A in Scheme 2.3) Cys160 removes the proton from C2 and the resulting flow of electrons forms a double bond between C1 and C2 to generate the cis-enediol intermediate (2A and 3A, Scheme 2.3).

Asp240 mediates proton transfer between O2 and O4 simultaneously, and the proton abstracted by Cys160 is donated back to form the MTRul-1-P product.[125] This mechanism is potentially facilitated by the isolation of the active site from solvent in the process of substrate binding.[125]

2.2.2.2 Hydride transfer mechanism

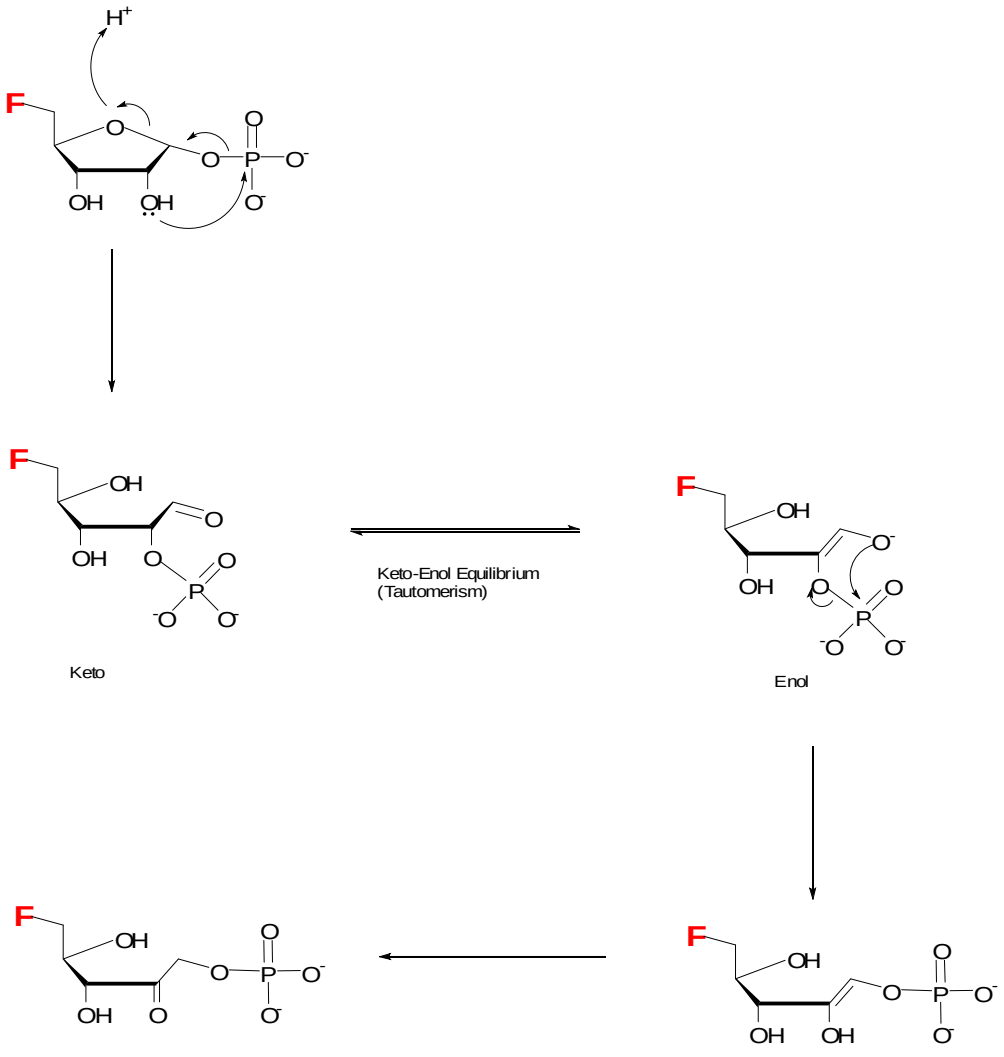
The hydride transfer mechanism (2B and 3B, Scheme 2.4) would be initiated by protonation by the side chain of Asp240 to O2 of MTR-1-P.[125] This would initiate ring opening and Cys160 may stabilize the positive charge generated at C1. The hydride on C2 then transfers to C1, following the flow of electrons from O2 to C2, forming a carbonyl and generating the product, MTRul-1-P.[125]

Xylose isomerase (XI), also a aldose-ketose isomerase, is thought to proceed *via* a hydride shift mechanism.[131, 132] Previous work involving the MTRI from *B. subtilis* in D_2O by NMR and Mass spectrometry are similar to results gained from XI, however they are inconclusive. NMR and GC-MS studies could not detect deuterium incorporation into the product, MTRul-1-P at C1.[133] This effectively rules out the cis-enediol mechanism. Although only implied by this negative result, the hydride shift in this case remains to be proven.

2.2.2.3 Phosphate transfer mechanism

The presence of a phosphate group and its putative effects upon the stability of the 5-MTR-1-P molecule have prompted the proposal of a third mechanism for the catalytic

activity of MTRIs. This mechanism involves phosphate transfer to the C2 OH causing a keto-enol tautomerism. In the presence of the enol, the oxygen at C1 would attack the phosphate causing it to return to its initial position and the molecule then isomerizes to, ribulose-1-phosphate (Scheme 2.5).



Scheme 2.5. A putative phosphate transfer mechanism for isomerisation of 5-FDRP **38** to 5-FDRulP **39**.

The proposed phosphate transfer mechanism implicates the role of a basic amino acid side chain in the active site, to deprotonate C2 and trigger phosphate transfer. Crystal

structures of the MTRI from *B. subtilis* have identified Asp240 as a potential key catalytic residue, and this residue is conserved amongst all known MTRIs. This residue sits in a hydrophobic pocket in the active site. Hydrophobicity has been reported to increase the pK_a of Asp26 in thioredoxin of *E. coli* from 4.4 to 7.5.[134] This increase in pK_a may also occur in MTRIs and consequently Asp240 may act as a base to trigger ring opening in this mechanism. It may also explain why enzymes of this type are found to be most active under basic conditions (~ pH 8). There may also be a need for stabilisation of the enol transition state, in order for the second phosphate transfer to occur. The residues Arg51, Gly52, Arg94, Gln199 and Lys251 have been implicated in the hydrogen bonding of the phosphate moiety of MTRul-1-P. It appears that significant changes in the interactions at the active site would be required for phosphate transfer to occur.

2.3 Identification of an MTRI from *S. coelicolor*

The peptide sequences of the known MTRIs from *S. cerevisiae* and *B. subtilis* was used in a BLAST search against the full *Streptomyces coelicolor* genome database.[135] This search highlighted a putative translation initiation factor within the *S. coelicolor* genome (*SCO3014*) with 38% and 33% identity to the *S. cerevisiae* and *B. subtilis* MTRIs respectively. SCO3014 possesses an open reading frame (ORF) of 1124 bp, and encodes a peptide of 39.1 kDa (39135 Da). With a knowledge that MTRIs have been consistently mis-annotated because of their close homology with eIFs, the peptide sequence for SCO3014 was aligned alongside the known MTRIs in order to assess if this may also have been a case of mis-annotation (Figure 2.3).

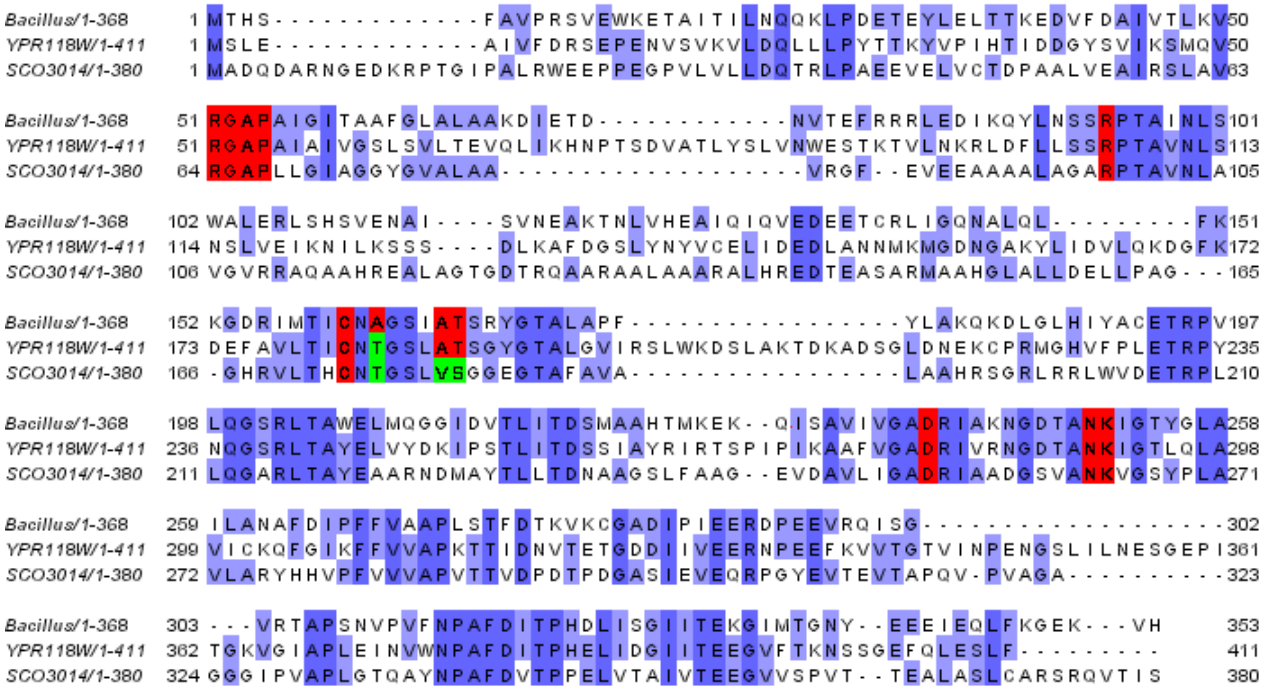


Figure 2.3. Alignment of the sequences of the known MTRIs from *B. subtilis* and *S. cerevisiae* with the sequence of the putative translation initiation factor *SCO3014* from *S. coelicolor*. Blue= conserved residues. Red= conserved catalytic residues determined from *B. subtilis*. Green= catalytic residues substituted by residues common in other MTRIs.

The alignment of the amino acid sequence of the putative initiation factor *SCO3014* from *S. coelicolor* reveals many regions of homology with known MTRIs (Figure 2.3). The putative catalytic residues located in the active site of the MTRI from *B. subtilis* (highlighted in red in Figure 2.3) are also conserved in SCO3014, suggesting that they may serve the same function. Particularly, Asp240 and Cys160 are conserved across all three peptide sequences, and their putative role in the catalysis of isomerisation has been discussed previously (Chapter 2.2.1). There are some small differences (highlighted in green) in the active site residues, however these residues are not absolutely conserved amongst MTRIs. In these cases, Ala162 is often exchanged to a threonine residue, Ala166 to a valine and Thr167 is often substituted for a serine. This variation is also shown to occur in the YPR118W sequence. These observations suggest that *SCO3014*

may well have been mis-annotated, and that its function is that of an MTRI. It became an objective to amplify the *SCO3014* gene and overexpress the protein product in order to assay it for this activity.

2.3.1 *SCO3014* amplification

The genetic sequence of *SCO3014* is complete and specific DNA primers were designed in the 5'-3' complement and 3'-5' reverse complement directions, with EcoRI and XhoI restriction sites respectively (Table 2.1).

Primer/ Restriction Site	Sequence 5'-3'
Forward/ EcoRI	**gcaggaggaattcatatggctgatcaggacgcgc**
Reverse/ XhoI	**ccctcacgccgctcgagttagctaatcgttacctgg**

Table 2.1. Specific DNA primers for the amplification of SCO3014 from *S. coelicolor* DNA. Red= DNA complementary to SCO3014. Blue= Restriction enzyme sites.

Genomic DNA was prepared from *S. coelicolor* and SCO3014 was amplified using the primers described in Table 2.1 and using the *pFu* DNA polymerase. In the event, a PCR product of about 1.1 kb was identified by DNA gel analysis (Figure 2.4). The PCR product highlighted in Figure 4 was then excised from the gel and purified in water to give ~65 ng/µl of DNA.

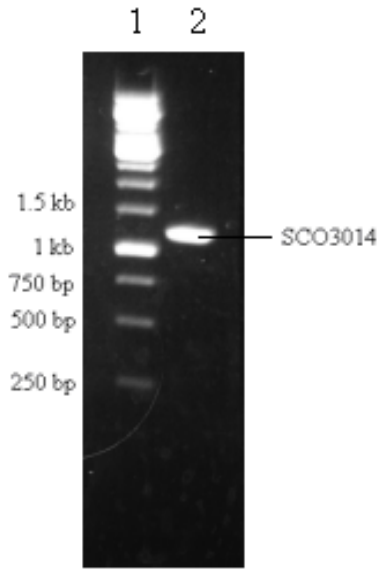


Figure 2.4. UV photograph of the 1% agarose gel containing the product of the PCR of *SCO3014* from *S. coelicolor* genomic DNA using the primers from Table 2.1.

2.3.2 Expression of the SCO3014 protein in *E. coli*

The prepared PCR DNA, and the *E. coli* expression plasmid pHISTEV were digested by *EcoRI* and *XhoI* restriction enzymes for 4 h at 37°C. The DNA preparations were then re-purified and introduced to each other in a 3:1 (PCR : pHISTEV) ratio, in the presence of T4 DNA ligase for 16 h at 4 °C. The ligation product was then transfected into competent *E. coli* BL21 (DE3) Gold cells and colonies were selected for resistance to kanamycin. Resistant clones were then picked from the petri dish and subjected to colony PCR using the primers in Table 2.1. Those colonies that exhibited a PCR product of ~1.1 kb were then picked and inoculated in LB medium (5 ml) containing kanamycin in a 15 ml falcon tube (agitated for 16 h at 37ºC). The cells were then pelleted by centrifugation and the recombinant plasmid extracted and prepared for DNA sequencing using the QIAPrep Spin Miniprep Kit.

tggatggaaaaatttcccctctagaaataattttgtttaactttaagaaggagatatacatatgtcgtactaccttcaccatcaccatcacgatt
acgatatcccaacgaccgaaaacctgtattttcagggcgccATGGCTGATCAGGACGCGCGAAACGGCGAGGA
CAAGCGGCCGACCGGGATACCGGCCCTTCGCTGGGAGGAACCCCCCGAGGGCCCGGTGC
TGGTGCTGCTGGACCAGACCAGGTTGCCGGCCGAGGAGGTCGAGCTGGTCTGCACGGAC
CCGGCCGCGCTGGTGGAGGCGATCCGCTCGCTCGCCGTGCGCGGGGCACCGCTGCTGG
GCATCGCGGGCGGCTACGGCGTCGCGCTCGCCGCCGTACGGGGCTTCTAGGTCGAGGAG
GCCGCGGCGGCGCTGGCGGGGGCGCGCCCGACCGCGGTGAACCTCGCCGTCGGGGTGC
GCCGGGCGCAGGCCGCGCACCGGGAGGCGCTCGCCGGGACCGGTGACACCCGGCAGGC
CGCCCGGGCGGCGCTGGCCGCGGCAAGGGCGCTGCACCGGGAGGACACCGAGGCCAGC
GCCCTGATGGCCGCGCACGGACTCGCGCTGCTCGACGAGCTGCTGCCCGCCGGAGGACA
CCGCGTCCTCACGCACTGCAACACCGGTTCGCTGGTGTCGGAGGGGGATGTGACCGCCTT
TCGCGGTTGCTCTCTCGGCGCACCTATCGGGACGGCTGCTACGGCTGTGGGTGGACATAT
GCGTCCCTTGTCTGTAGGGTGCTCGCAGAACGTATACGAGAGGTCCTCCACGACTTGCGTA
CACCTTGCTCCCCAACAAAGCGGAATTTTCATGTTTCCTGAGGGGAGAAGCTGACCCTTAAT
TGATT

Figure 2.5. Confirmed DNA sequence of SCO3014-pET28 recombinant plasmid. Black= pHISTEV, Red= SCO3014.

The presence of *SCO3014* in pHISTEV was confirmed by DNA sequencing (Figure 2.5). The *SCO3014*-pHISTEV recombinant plasmid was then transfected into *E. coli* BL21 Gold competent cells and the SCO3014 protein was expressed after incubation with IPTG (1mM) for 16 h at 16ºC. Following Ni^{2+} affinity chromatography, samples from the cell free extract, cell debris, supernatant, column flow through, column wash, and column elution were mixed with SDS dye at 95 ºC for 5 min. Samples were then loaded onto a 1 mm 4-12% Bis-Tris gel submerged in MES SDS running buffer for SDS PAGE analysis. The resulting SDS-PAGE revealed a band in the eluent fraction of ~40 kDa (Figure 2.6).

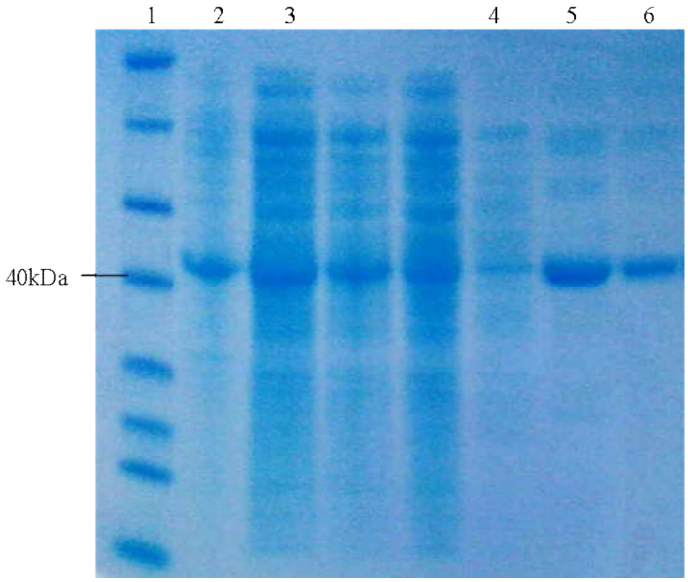


Figure 2.6. SDS PAGE of the Nickel Column Purification of MTRI from *S. coelicolor*. 1) SDS PAGE protein ladder 2) Cell Free Extract. 3) Supernatant. 4) Column Flow Through. 5) Wash. 6) Elution.

The expression of SCO3014 was confirmed by in-gel tryptic digest and analysis of the resultant peptides by nanoLC-ESI MSMS (UltiMate (Dionex) and Q-Star Pulsar XL (Applied Biosystems)). The MS/MS data file generated was analysed using the Mascot 2.1 search engine (Matrix Science, London, UK) against an internal database consisting of a bacterial genome background to which the SCO3014 sequence (amongst others) had been added. The data was searched with tolerances of 0.2 Da for the precursor and fragment ions. Trypsin was used as the cleavage enzyme and up to one missed cleavage was assumed. Carbamidomethyl modification of cysteines was selected as a fixed modification and L-methionine oxidation was selected as a variable modification.

The elution fraction following Ni^{2+} affinity chromatography contained the His_6 tagged protein product of *SCO3014* (~2 mg/ml). This fraction was concentrated (2ml) and

subjected to FPLC size exclusion (Figure 2.7). Size exclusion chromatography using phosphate buffer (10mM, pH 7.8) revealed that the SCO3014 protein eluted between 75 and 80 ml (the highlighted region of Figure 2.7) with a protein concentration of 1.3 mg/ml. This data indicates that the soluble MTRI exists as a dimer and supports the hypothesis that this protein is not an initiation factor, as active eIF proteins exist as monomers.

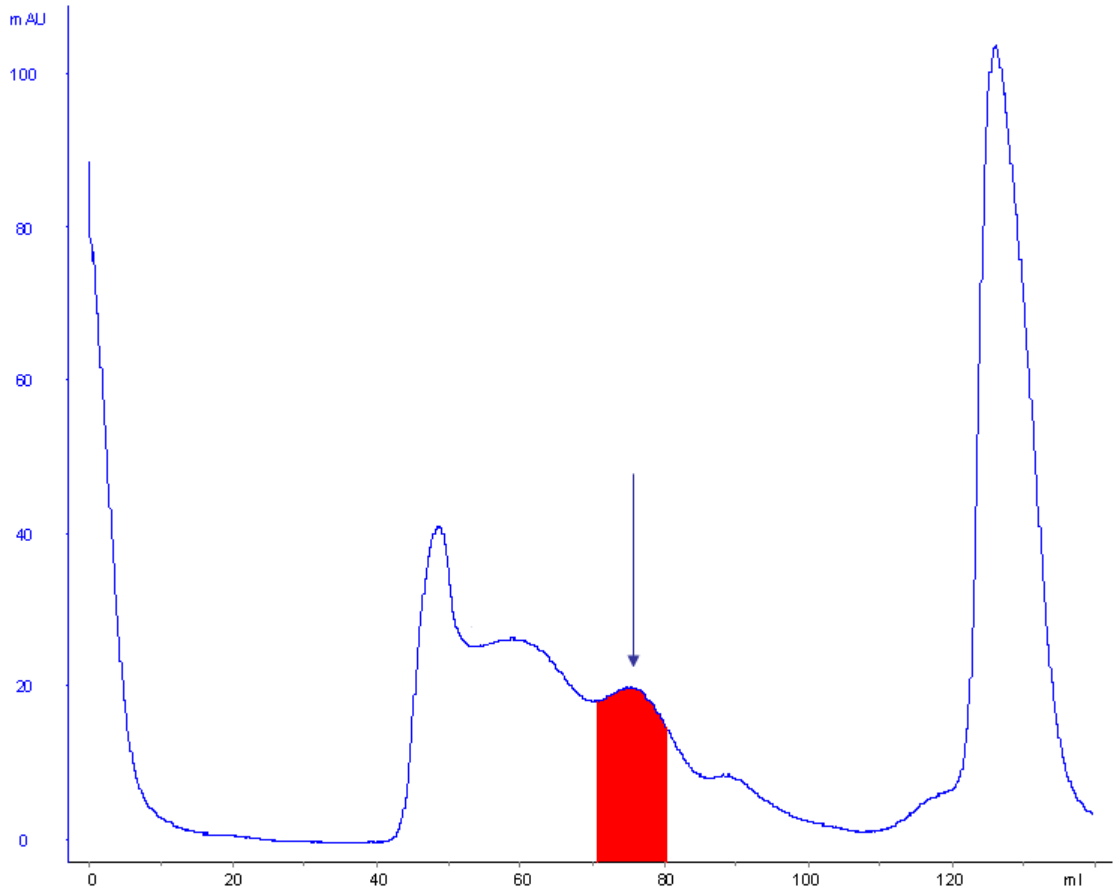


Figure 2.7. Chromatogram obtained by size exclusion using a Superdex 200 column (120 ml) after the injection of a sample (2 ml) from the eluent fraction of the Ni^{2+} affinity chromatography for the SCO3014 protein product. The highlighted region exhibits the elution fraction containing SCO3014.

After size exclusion chromatography, the fractions containing the SCO3014 protein product were identified by SDS PAGE analysis. Another level of purification was required due to the lack of baseline separation as illustrated in Figure 2.7. This was achieved by concentrating the pooled fractions containing the SCO3014 protein (2 ml), and applying the sample to HiTrap Q HP anion exchange column (Amersham Biosciences, UK) equilibriated with phosphate buffer (10 mM, pH 7.8). The proteins contained in the sample were seperated by increasing the concentration of NaCl (from a 1M stock) and the resulting chromatogram is shown in Figure 2.8.

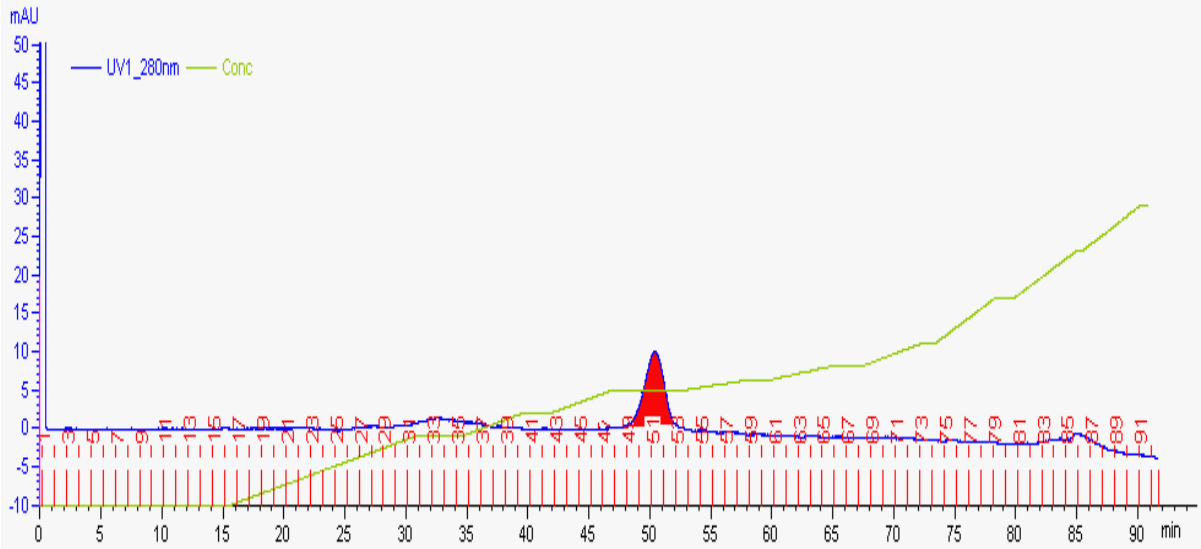


Figure 2.8. Chromatogram obtained by anion exchange, using a HiTrap Q HP column, of the concentrated SCO3014 protein fraction after size exclusion chromatography. SCO3014 eluted between fractions 49 and 53 at a NaCl concentration of 300 mM (highlighted red). Blue line= UV monitoring at 280 nm. Green line= NaCl concentration, ranging from 0 to 500 mM.

During anion exchange chromatography, the SCO3014 protein eluted from the column in the presence of 300 mM NaCl, giving relatively pure protein. The fractions containing SCO3014 were pooled, and concentrated (2 ml). The resultant sample was then injected

to a 5 ml HiTrap desalt column (Amersham Biosciences, UK) equilibriated with phosphate buffer (10 mM, pH 7.8) to remove NaCl from the sample (Figure 2.9).

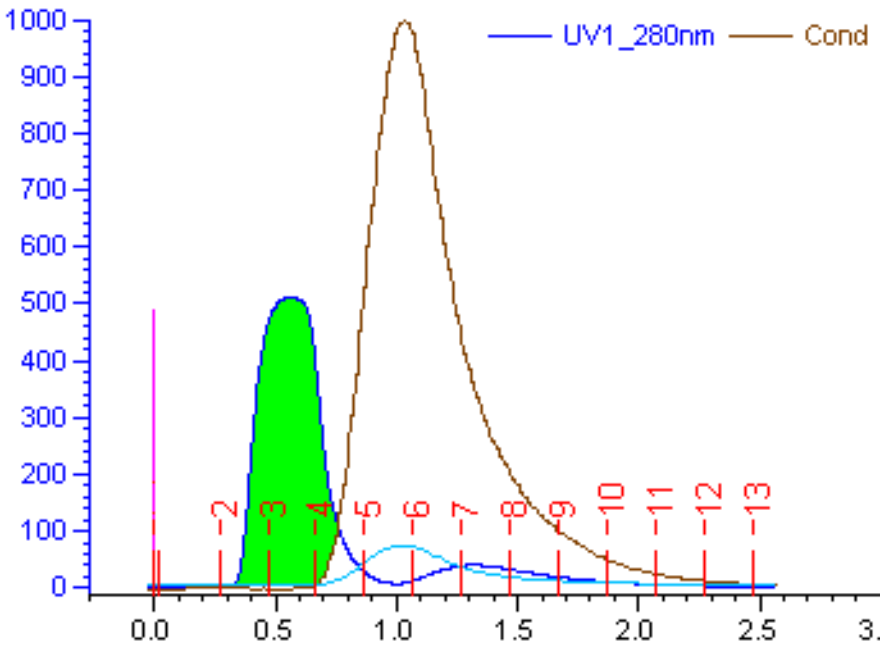


Figure 2.9. Chromatogram obtained by desalting (HiTrap desalting column, 5 ml) of a sample (2 ml) containing the SCO3014 protein after anion exchange chromatography. Sample vial numbers are shown in red, UV at a wavelength of 280 nm in blue and the conductivity (i.e salt) in brown.

The resultant chromatogram (Figure 2.9) indicates that the desalting process was successful. Sample vials 2 and 3 were pooled resulting in a final product containing ~ 1.3 mg/ml of relatively pure SCO3014 protein.

2.3.3 Assay of the SCO3014 protein

The putative MTRI protein was then incubated with 5-MTRP generated from synthetically produced 5'-FDI **36**, itself generated by treatment of 5'-FDA **35** with a commercially available adenosine deaminase. Treatment of 5'-FDI **36** with a PNP then

generated 5-FDRP **38**. This was achieved from synthetic 5'-FDA **35** (supplied by M. Onega), suspended in phosphate buffer (20 mM, pH 7.5) and incubated with adenosine deaminase (5 mg/ml, Sigma UK) for 1 h at 37 ºC. The reaction was stopped by heat deactivation at 95 ºC for 5 min, and then centrifuged at 12,000 rpm for 2 min. The resulting supernatant was removed and incubated with commercially available PNP suspended in phosphate buffer (5 mg/ml, 20 mM, pH 7.5) for 16 h at 37 ºC. The reaction was stopped by heat deactivation at 95 ºC for 5 min, and then centrifuged at 12,000 rpm for 2 min. A sample of the supernatant was then examined by ^{19}F NMR, and a second sample was incubated with the SCO3014 protein product for 6 h at 37 ºC. A control experiment was set up identically without the SCO3014 protein preparation. The reactions were stopped by heat deactivation and centrifugation as before and the supernatants were subject to ^{19}F NMR analysis (Figure 2.10).

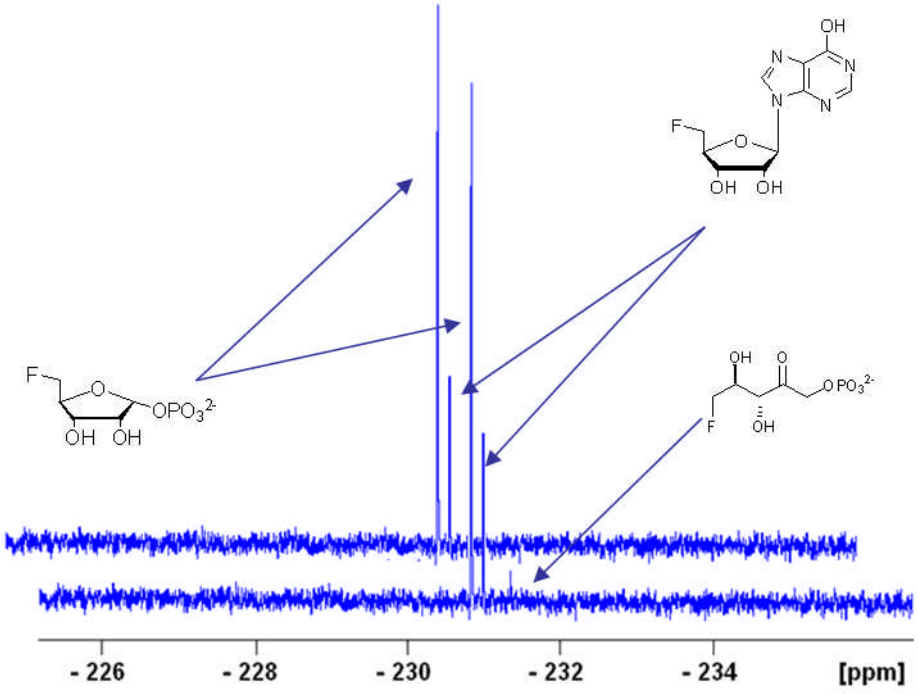


Figure 2.10. ^{19}F {^{1}H}NMR spectra of 5-FDRP **38** incubated with MTRI from *S. coelicolor* for 6 hours at 37ºC and a control without the MTRI protein added.

Previous work in the research group had identified the ^{19}F NMR chemical shift associated with 5-FDRulP **39** as -231.8 ppm.[93,97] Following incubation of SCO3014 with a preparation of 5-FDRP **39**, ^{19}F NMR analysis revealed a new signal at -231.8 ppm consistent with isomerisation. After 6 hours incubation at 37 ºC a peak corresponding to 5-FDRulP **39** was identified by ^{19}F NMR (Figure 2.10). The results clearly show that the over expressed protein was indeed an MTRI from *S. coelicolor*. Perhaps even more interestingly, this enzyme is capable of isomerising a substrate with fluorine at the 5-position as opposed to the thiomethyl group, its natural substrate in the L-methionine salvage pathway.

2.4 Identification of an MTRI from *S. cattleya*

The goal now was to identify an analogous MTRI gene in *S. cattleya*. The isolation of the enzyme from *S. coelicolor* gave us a framework to scan the genome of another close relative to *S. cattleya*, that of *Streptomyces avermilitis*. A BLAST search against the full *S. avermilitis* genome, using the protein sequence of SCO3014, enabled the indentification of a protein of 346 amino acids (SAV6658) with 45% homology to a putative eIF2 initiation factor. The identification of this sequence of high homology to a known MTRI afforded confidence to align the two sequences from *S. coelicolor* and *S. avermilitis* in order to establish highly conserved regions within the amino acid sequences. These are shown in Figure 2.11. From these highly conserved sequences, it was possible, with the use of a *S. cattleya*-specific codon usage table (Figure 2.12), to design degenerate DNA primers in order to attempt to isolate and amplify a gene from genomic DNA from *S. cattleya*.

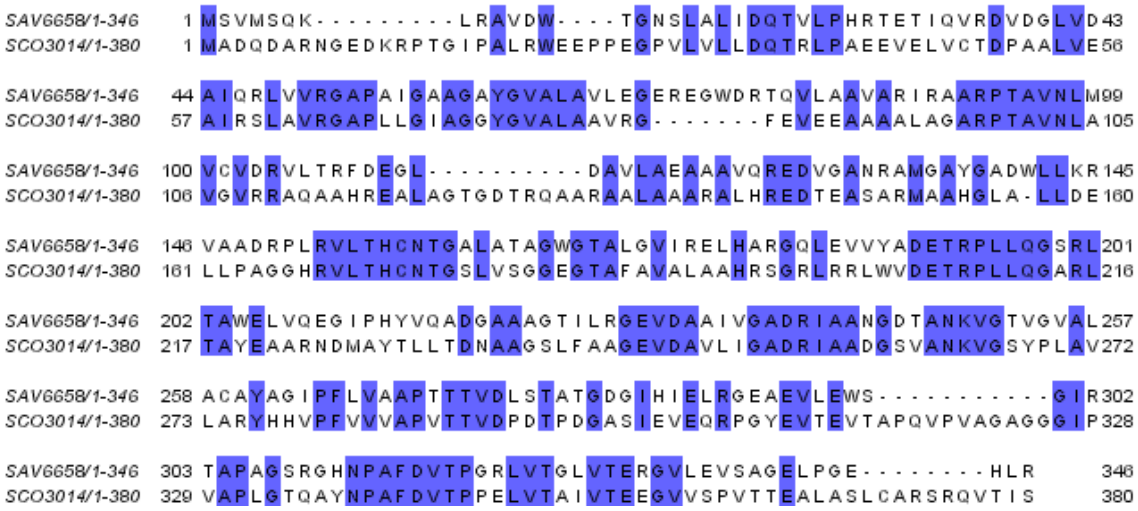


Figure 2.11. Alignment of putative MTRI from *S.coelicolor* SCO3014 and a homologous protein from *S. avermilitis* SAV6658. Conserved residues highlighted in blue, putative catalytic residues highlighted in red.

Following the alignment of the peptide sequences of SCO3014 and SAV6658, degenerate PCR primers were designed for attempted PCR amplification using *S. cattleya* genomic DNA as the template. This amplification was carried out by Dr Hai Deng (University of St Andrews). In the event a 288 bp PCR product was amplified and sequence comparison indicated a deduced amino acid sequence with 90% homology to SCO3014 and SAV6658. Chromosomal walking (Dr H. Deng) from this DNA fragment resulted in the sequencing of a 1161 bp ORF (*MTRI-Sca*) shown in Figures 2.12 and 2.13. The deduced amino acid sequence gives a putative protein consisting of 386 amino acids with a molecular weight of 38.22 kDa.

gaacacgcgtcgtttacctccggggtgagagtcctggtccggtcagctcggtcacctcctgccgccgcgcgaagcccgccccgcgcccggctccgcgcc
cggcccggaacggtcggtacggcggcggcaccgggccaccacgaccggccggcctcggacgaattgcgccgcgcccgttcaacgcctccgctccgc
ggggcagactgacgac**atgggtgatcagtccgtacagcctttggccaagggcacggggtccgggaccccggagccgaaacccgctctccgctgg**
gaagagcctcccgaagggcccgtgctggtcctcctcgaccagacccggctccccgtcgaggaggtcgaactgttctgtacggacgtgcccgcgct
cgtccaggccatccgtaccctcgccgtccgcggcgcgccgctgctcgggctcgccggagcgtacggcgtcgccctggccgccgcccgtggctacg
acgtcgggcaggccgccgacgaactcgccggcgcccggcccaccgccgtcaacctctcctacggggtgcgccgcgcgctggccgcgtaccgtac
cgcggtcaccggcggcgccgacgacacgggcgcggcggcggccaccctcgccgaggcccgcgcgctgcacgccgaggacgccagggccagc
gaacgcatggcccgcaacggcctggcgctgctggacgaactcgtccccggcggcggctaccgggtgctgacccactgcaacaccggcgccctgg
tctccggcggcgagggcaccgccctggccgtcgtcctcgccgcccaccgcggcggactcctgcgccggctgtgggtggacgagacccggccgct
gctccagggcgcccggctgaccgcctacgaggccgcccgggccggcgtcgcccacacgttgctgccggacggcgcggccgggtcgctcttcgcg
gccggcgaggtcgacgcggtgctgatcggcgccgaccggatcgccgcggacggctcgaccgccaacaaggtcggcagctacccgctggccgtc
ctcgcccggtaccacaacgtccccttcgtcgtggtcgcccccaccaccacgatcgacctggccacccccgacggcaccgcgatcgaggtggagca
gcgcccgcgcaggaggtgaccgagctgaccggaccgcgccccggcccggaccgcgagggcgccaccggcatccccgtcgcgcccctgggca
cgccggcgtacaacccggcgttcgacgtcaccccgcccgaactgatcaccgccgtggtcaccgagaccggcgtggcctccccggtcaccggctcc
tccatagccgccctggccgcccgccccggccccgtccgcgcccagccgtgacggcctcgacggtcgtcatccgccgccgatgacgaccgtcaccc
gtcaccgatgacgaccgtcccttcc

Figure 2.12. The *MTRI-Sca* ORF identified from *S. cattleya* genomic DNA after degenerate PCR and gene walking targeted towards MTRI identification. Red= ORF of 1161 bp.

```
SCa/1-386      1 MGDQSVQPLAKGTGSGTPEPKPALRWEEPPEGPVLVLLDQTRLPVEEVELFCTDVP 56
SCa/1-386     57 ALVQAIRTLAVRGAPLLGLAGAYGVALAAARGYDVGQAADELAGARPTAVNLSYGV 112
SCa/1-386    113 RRALAAYRTAVTGGADDTGAAAATLAEARALHAEDARASERMARNGLALLDELVPG 168
SCa/1-386    169 GGYRVLTHCNTGALVSGGEGTALAVVLAAHRGGLLRRLWVDETRPLLQGARLTAYE 224
SCa/1-386    225 AARAGVAHTLLPDGAAGSLFAAGEVDAVLIGADRIAADGSTANKVGSYPLAVLARY 280
SCa/1-386    281 HNVPFVVVAPTTTIDLATPDGTAIEVEQRPAQEVTELTGPRPGPDREGATGIPVAP 336
SCa/1-386    337 LGTPAYNPAFNVTPPELITAVVTETGVASPVTGSSIAALAARPGPVRAQP       386
```

Figure 2.13. The 386 amino acid sequence derived from the 1161 bp ORF which was identified by degenerate PCR using genomic DNA from *S. cattleya*.

2.4.1 Amplification of *MTRI-Sca*

Alignment of the amino acid sequence from the identified ORF of *S. cattleya* with those from the known MTRIs from *S. coelicolor* (SCO3014) and *B. subtilis* is shown in Figure 2.14.

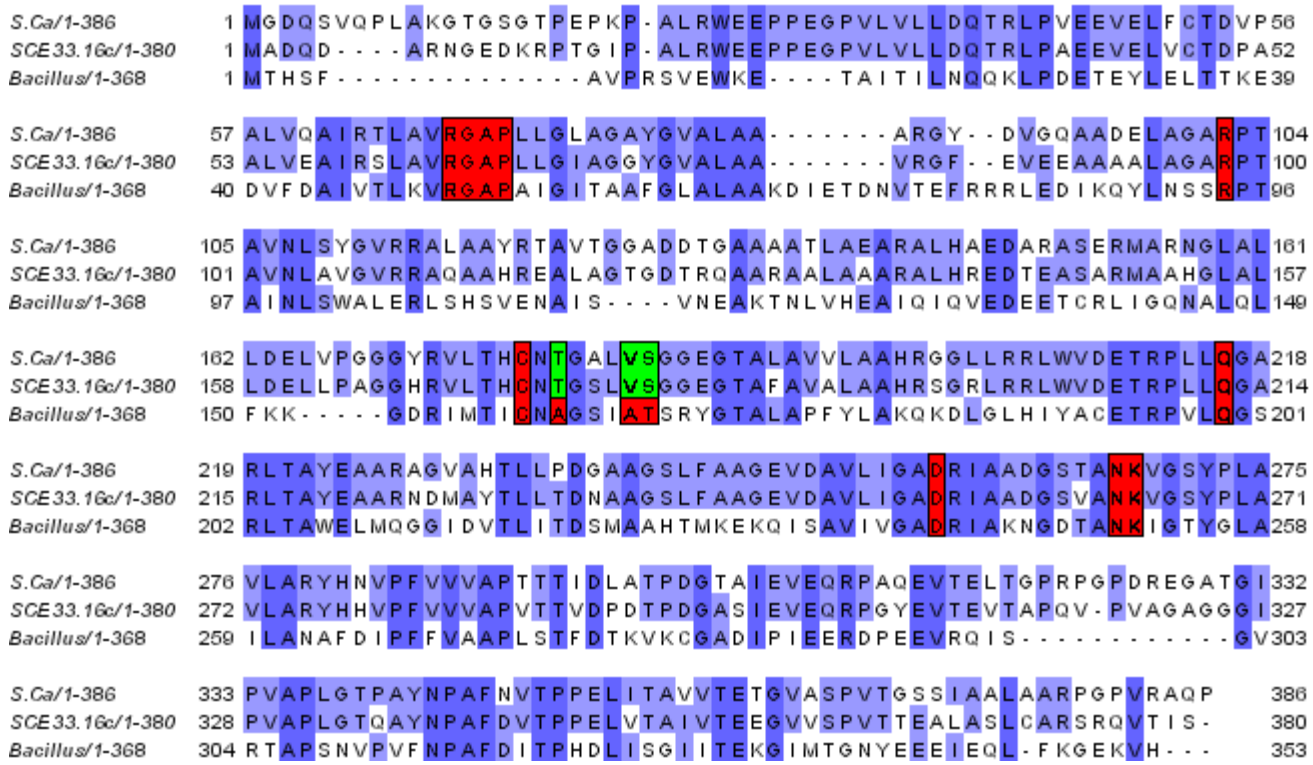


Figure 2.14. Alignment of amino acid sequence derived from the 1161 bp ORF obtained by degenerate PCR using genomic DNA from *S. cattleya* with the known MTRIs from *S. coelicolor* (SCO3014) and *B. subtilis*. Conserved amino acids are shown in blue. Red= Conserved putative catalytic residues. Green= Common variations in putative catalytic residues within MTRIs.

The subsequent alignment reveals that the putative protein sequence from *S. cattleya* contains many of the conserved putative catalytic residues that have been identified in other MTRIs.[12] The amino acid sequence from this ORF had ~75% identity to the putative isomerase from *S. coelicolor*. This suggests that the ORF identified from the genomic DNA of *S. cattleya* is likely to an MTRI. As a result, the specific DNA primers shown in Table 2.3 were designed.

Primer/ Restriction Site	Sequence 5'-3'
Forward/ EcoRI	GCAGGAGGAATTCCATGGGTGATCAGTCCGTACAGCCTTTGGC
Reverse/ XhoI	CGCCGCTCGAGCGGAAGGGACGGTCGTCATCGGTGAC

Table 2.3. Specific DNA primers designed for the PCR of the putative MTRI from *S. cattleya* from genomic DNA

Genomic DNA was prepared from *S. cattleya* and the putative MTRI gene was amplified using the primers described in Table 2.3 and using the *pFu* DNA polymerase. In the event, a PCR product of about 1.1 kb was identified by DNA gel analysis (Figure 2.15). The PCR product highlighted in Figure 2.15 was then excised from the gel and purified in water to give ~80 ng/µl of DNA.

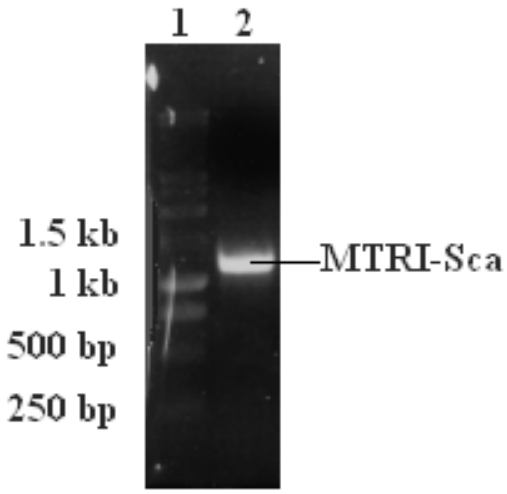


Figure 2.15. UV photograph of the 1% agarose gel containing the product of the PCR MTRI-Sca from *S. cattleya* genomic DNA using the primers from Table 2.3.

2.4.2 Expression and purification of *MTRI-Sca* in *E. coli*

The prepared PCR DNA, and the *E. coli* expression plasmid pHISTEV were digested by *EcoRI* and *XhoI* restriction enzymes for 4 h at 37°C. The restricted DNA preparations were then repurified into water and introduced in a 3:1 (PCR : pHISTEV) ratio, in the presence of T4 DNA ligase for 16 h at 4°C. The ligation product was then transfected into competent *E. coli* BL21 (DE3) Gold cells and colonies were selected for resistance to kanamycin. Resistant clones were then picked from the petri dish and subjected to colony PCR using the primers in Table 2.3. Those colonies that exhibited a PCR product of ~1.1 kb were then picked and inoculated in 5 ml LB medium containing kanamycin in a 15 ml falcon tube (agitated for 16 h at 37ºC). The cells were then pelleted by centrifugation and the recombinant plasmid extracted and prepared for DNA sequencing using the QIAPrep Spin Miniprep Kit (Figure 2.16).

tggatggaaaaatttcccctctagaaataattttgtttaactttaagaaggagatatacatatgtcgtactaccttcaccatcaccatcacgatt
acgatatcccaacgaccgaaaacctgtattttcagggcgccATGGGTGATCAGTCCGTACAGCCTTTGGCCAAG
GGCACGGGGTCCGGGACCCCGGAGCCGAAACCCGCTCTCCGCTGGGAAGAGCCTCCCGA
AGGGCCCGTGCTGGTCCTCCTCGACCAGACCCGGCTCCCCGTCGAGGAGGTCGAACTGTT
CTGTACGGACGTGCCCGCGCTCGTCCAGGCCATCCGTACCCTCGCCGTCCGCGGCGCGC
CGCTGCTCGGGCTCGCCGGAGCGTACGGCGTCGCCCTGGCCGCCGCCCGTGGCTACGAC
GTCGGGCAGGCCGCCGACGAACTCGCCGGCGCCCGGCCCACCGCCGTCAACCTCTCCTA
CGGGGTGCGCCGCGCGCTGGCCGCGTACCGTACCGCGGTCACCGGCGGCGCCGACGAC
ACGGGCGCGGCGGCGGCCACCCTCGCCGAGGCCCGCGCGCTGCACGCCGAGGACGCCA
GGGCCAGCGAACGCATGGCCCGCAACGGCCTGGCGCTGCTGGACGAACTCGTCCCCGGC
GGCGGCTACCGGGTGCTGACCCACTGCAACACCGGCGCCCTGGTCTCCGGCGGCGAGGG
CACCGCCCTGGCCGTCGTCCTCGCCGCCCACCGCGGCGGACTCCTGCGCCGGCTGTGGG
TGGACGAGACCCGGCCGCTGCTCCAGGGCGCCCGGCTGACCGCCTACGAGGCCGCCCG
GGCCGGCGTCGCCCACACGTTGCTGCCGGACGGCGCGGCCGGGTCGCTCTTCGCGGCC
GGCGAGGTCGACGCGGTGCTGATCGGCGCCGACCGGATCGCCGCGGACGGCTCGACCG
CCAACAAGGTCGGCAGCTACCCGCTGGCCGTCCTCGCCCGGTACCACAACGTCCCCTTCG
TCGTGGTCGCCCCCACCACCACGATCGACCTGGCCACCCCCGACGGCACCGCGATCGAG
GT

Figure 2.16. Confirmed DNA sequence of the PCR product using the primers from Table 2.3 with *S. cattleya* genomic DNA. Black = pHISTEV vector. Red= Putative MTRI.

The presence of the putative MTRI gene in pHISTEV was confirmed by DNA sequencing (Figure 2.16). The MTRI-pHISTEV recombinant plasmid was then transfected into *E. coli* BL21 Gold competent cells and the SCO1844 protein was expressed after incubation with IPTG (1mM) for 16 h at 16ºC. Following Ni^{2+} affinity chromatography, samples from the cell free extract, cell debris, supernatant, column flow through, column wash, and column elution were mixed with SDS dye at 95 ºC for 5 min. Samples were then loaded onto a 1 mm 4-12% Bis-Tris gel submerged in MES SDS running buffer for SDS PAGE analysis. The resulting SDS-PAGE revealed a band in the eluent fraction of ~40 kDa (Figure 2.17).

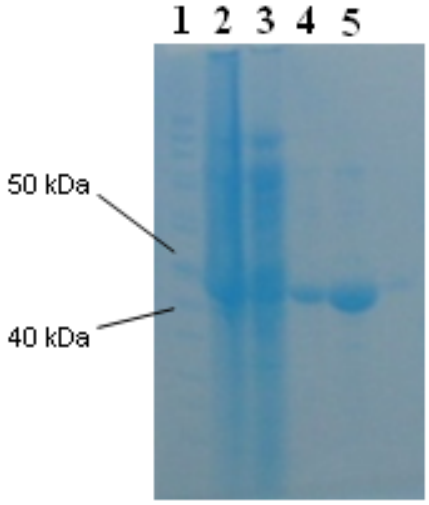


Figure 2.17. SDS-Page of the progressive purification of *S. cattleya* isomerase over-expressed in *E. coli*. 1) Protein molecular weight markers (Fermentas); 2) cell-free extract; 3) cell-free extract supernatant; 4) column wash from Ni-affinity column; 5) Eluent from Ni-affinity column; The identity of the protein was confirmed by MS-MS mass spectrometry.

The elution fraction following Ni^{2+} affinity chromatography contained the protein product of MTRI-Sca (~2 mg/ml). This fraction was concentrated (2ml) and subjected to FPLC gel filtration (Figure 2.18). Gel filtration using phosphate buffer (10mM, pH 7.8)

revealed that the SCO3014 protein eluted between 56 and 60 ml (the highlighted region of Figure 2.7) with a protein concentration of 1.3 mg/ml. This data indicates that the soluble MTRI exists as a dimer and is unlikely therefore to be an initiation factor eIF.

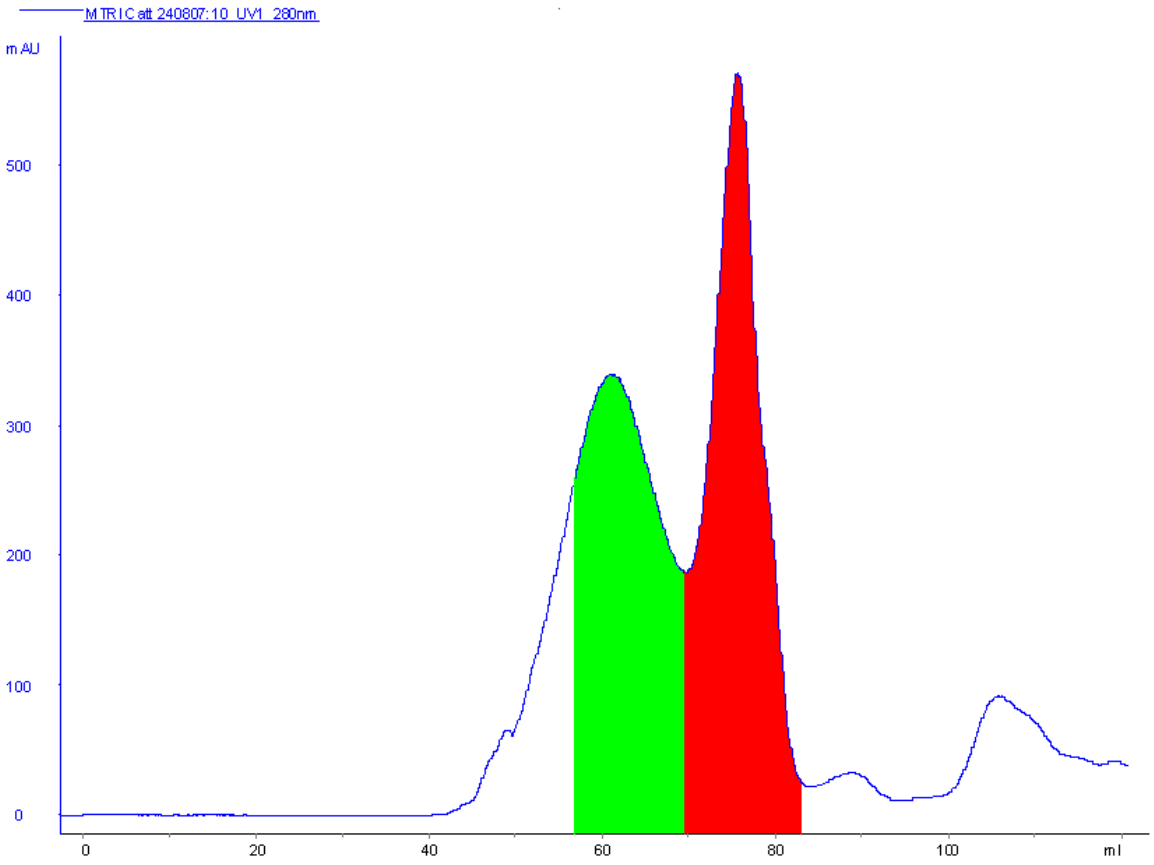


Figure 2.18. Chromatogram obtained by size exclusion using a Superdex 200 column (120 ml) after the injection of a sample (2 ml) from the eluent fraction of the Ni^{2+} affinity chromatography for the MTRI-Sca protein product. The red region exhibits the elution fraction containing dimeric MTRI-Sca protein, the green represents tetrameric MTRI-Sca.

Analysis of the MTRI-Sca protein product by size exclusion chromatography revealed that the expressed protein was present in solution in two forms, as a homodimer and a homotetramer. The homotetramer eluted from the gel filtration column after ~65 ml (green area, Figure 2.18), and the homodimer after ~75 ml (red fraction, Figure 2.18). Samples were taken from the column eluant that corresponded to each of these fractions,

and the eluted protein was identified as MTRI-Sca by SDS PAGE (Figure 2.19) and also by MS-MS analysis.

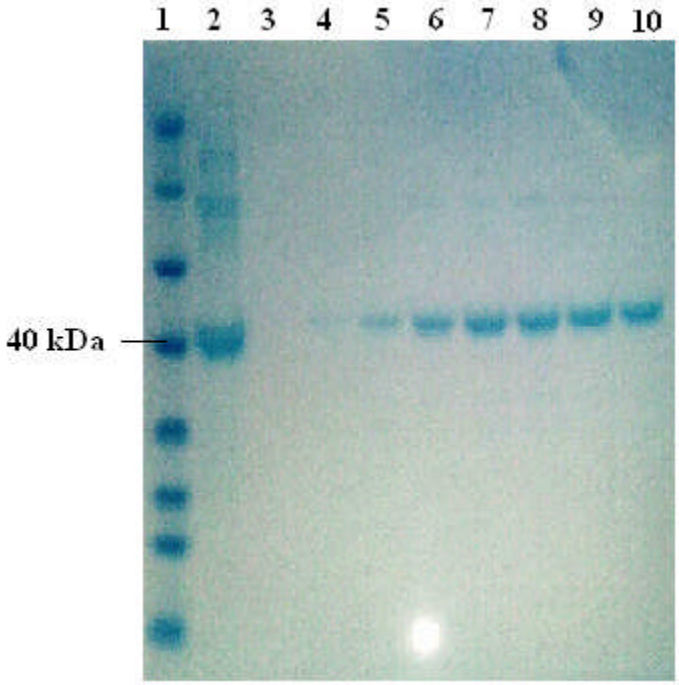


Figure 2.19. SDS-PAGE of the size exclusion fractions from *S. cattleya* isomerase over-expressed in *E. coli*. 1) Protein molecular weight markers (Fermentas); 2) Ni^{2+} affinity eluent 3) Fractions 40-45 ml 4) Fractions 45-50 ml; 5) Fractions 50-55 ml. 6) Fractions 65-67. 7) Fractions 67-70 ml. 8) Fractions 70-72 ml. 9) Fractions 72-75 ml. 10) Fractions 75-78 ml. The identity of the protein was confirmed by MS-MS mass spectrometry.

The size exclusion analysis of over expressed MTRI-Sca shows that either the protein product has been expressed in two forms, or there is a transient movement between the two. It was particularly important to determine this information in order to enter X-ray crystal trials with this protein. To establish whether each, or both of the soluble forms of MTRI-Sca are transient or fixed. The eluant containing the dimeric form of the protein (the major product) was collected and pooled and left for 6 hours at room temperature.

Following this period, the pooled fractions were concentrated to 2 ml and subject to a second round of size exclusion chromatography, using the same method as before (Figure 2.20).

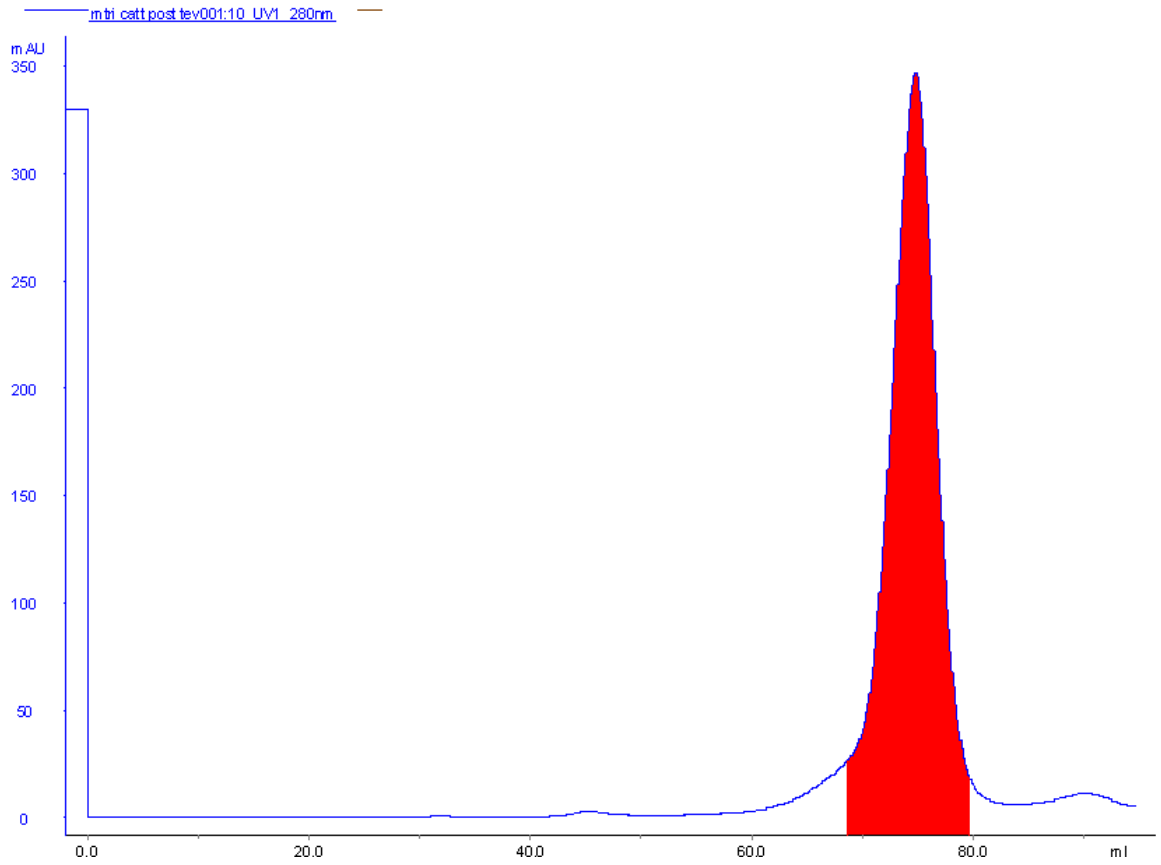


Figure 2.20. Chromatogram obtained by size exclusion using a Superdex 200 column (120 ml) after the injection of a 2 ml sample from the eluent fraction of the Ni^{2+} affinity chromatography for the MTRI-Sca protein product. The red region exhibits the elution fraction containing dimeric MTRI-Sca protein.

The results of a second size exclusion chromatography on the protein product of MTRI-Sca reveals that only the homodimer is present and that the homodimeric form of this protein is stable. The homodimer-containing fractions were then collected, to a final concentration of ~2 mg/ml in phosphate buffer (10 mM. pH 7.8). The data in Figure 2.20 suggest that formation of the homotetrameric protein occurs as an artefact of

overexpression in an *E. coli* host, and not as a transient form in the presence of concentrated protein. As the major protein product of MTRI-Sca exists as a homodimer, combined with the fact that known MTRIs are active as dimers, this form of the protein is assumed to be the native, active structure.

2.4.3 Assay of *MTRI-Sca*

The MTRI-Sca protein product was then incubated with 5-FDRP **38** generated from synthetic produced 5'-FDA **35** *via* 5'-FDI **36** by commercially available adenosine deaminase and PNP respectively. Synthetic 5'-FDA **35** was suspended in phosphate buffer (100mM, pH 7.5) and incubated with adenosine deaminase (Sigma, UK) for 1 h at 37 ºC. The reaction was stopped by heat deactivation at 95 ºC for 5 min, and then centrifuged at 12,000 rpm for 2 min. The resulting supernatant was removed and incubated for 16 h at 37 ºC. The reaction was stopped by heat deactivation at 95 ºC for 5 min, and then centrifuged at 12,000 rpm for 2 min. A sample of the supernatant was then analysed by ^{19}F NMR, and a second sample was incubated with the MTRI-Sca protein product (~0.1 mg) for 6 h at 37 ºC. The reaction was stopped by heat deactivation and centrifugation as before and the supernatant was subject to ^{19}F NMR analysis (Figure 2.21).

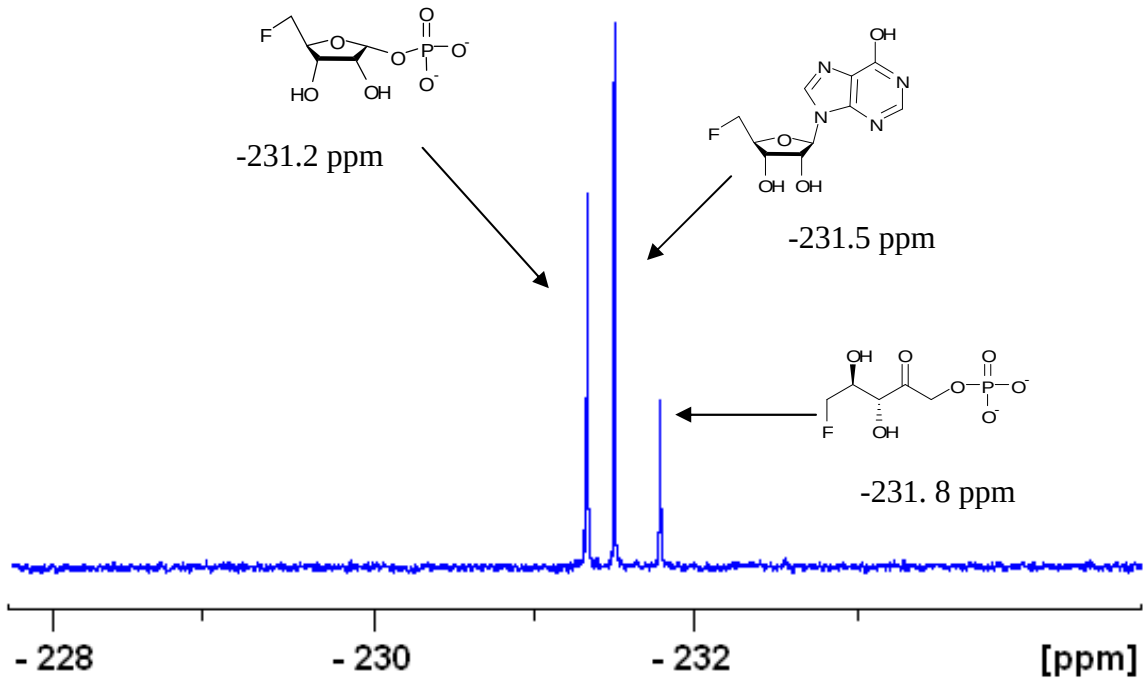


Figure 2.21. $^{19}F\{^{1}H\}$ NMR 5-FDRP incubated with the putative MTRI from *S. cattleya* for 6 h at 37ºC.

Upon incubation of the MTRI-Sca protein with a preparation of 5-FDRP **38**, analysis by ^{19}F NMR revealed a new signal at -231.8 ppm. This peak is consistent with the generation of the open sugar 5-FDRulP **39** (Figure 2.21). The results clearly show that the MTRI-Sca protein was an MTRI from *S. cattleya*. Similarly to its homolog from *S. coelicolor* (SCO3014), MTRI-Sca is capable of performing the isomerisation of the fluorinated substrates, 5-FDRP **38** to 5-FDRulP **39**. The MTRI-Sca protein may well be the enzyme responsible for isomerisation in the elaboration of fluorometabolites in *S. cattleya* as well as playing a dual role in the methionine salvage pathway.

2.4.3.1 Assay of MTRI-Sca in the presence of EDTA

It is known that some aldose-ketose isomerases require two divalent cations (e.g. Zn^{2+}) in order to be catalytically active. This is the case for xylose isomerase.[132] The MTRIs from *B. subtilis*, *S. cerevisiae* and *S. coelicolor* do not require such metals for activity. Studies on the cell free extracts of *S. cattleya* revealed that incubation with EDTA does not inhibit the isomerisation of 5-FDRP **38** to 5-FDRulP **39.** [97]

It was important to establish at this stage if the isomerase identified from *S. cattleya* was inhibited by the presence of EDTA. Accordingly MTRI-Sca (0.1 mg) in phosphate buffer (10 mM, pH 7.8) was incubated with synthetic 5-FDRP in the presence of EDTA (1 mM) (6 hours, 37 °C). The reaction was stopped by heat deactivation (95 °C, 5 min), followed by centrifugation (12,000 rpm, 2 min). The resulting supernatant was removed and the volume made up to 700 µl using ultrapure water. The mixture was then added to D_2O (100 µl) and subject to ^{19}F NMR. The resulting ^{19}F NMR spectrum is shown in Figure 2.22.

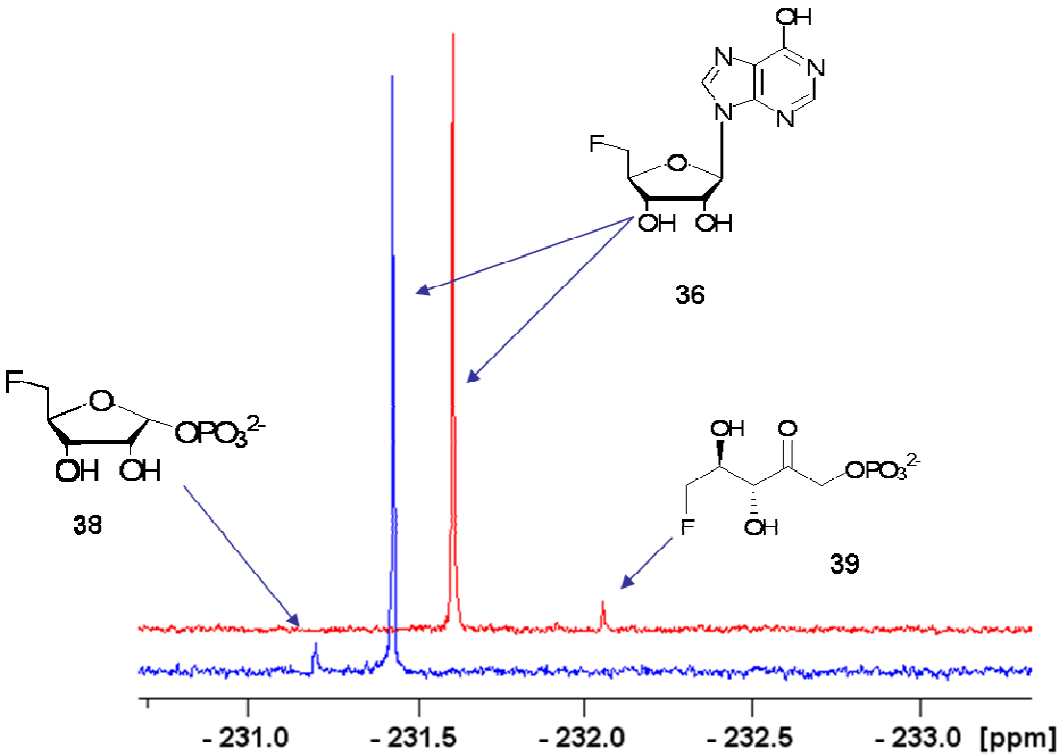


Figure 2.22. $^{19}F\{^{1}H\}$ NMR spectra of MTRI-Sca incubated with 5-FDRP **38** in the presence of EDTA (1 mM) and a control without the MTRI-Sca protein (blue)

Incubation of MTRI-Sca in the presence of 1 mM EDTA did not appear to reduce catalytic activity. A new product signal peak at -231.8 ppm was clear, which corresponds to the production of 5-FDRulP **38** (Figure 2.22). Thus the MTRI does not appear to be dependant upon divalent ion co-factors. It is also known from previous *in vitro* studies that the fluorometabolite pathway in *S. cattleya* is not inhibited by the presence of EDTA.[97]

2.4.3.2 *In vitro* generation of 5-FDRulP from fluoride ion

MTRI-Sca was now assayed for its ability to generate 5-FDRulP **38** in the reconstitution of the first three steps of the fluorometabolite pathway *in vitro*. This could be achieved by

over expressing the fluorinase, PNP and isomerase in *E. coli* and to generate 5-FDRulP **39** from SAM **34** and fluoride ion in a one-pot reaction. The fluorinase and PNP (*FlB*) genes from *S. cattleya* have been reported, and successful over expression of these proteins in *E. coli* has been achieved through previous work within the research group. The identification now of an isomerase from *S. cattleya*, that is capable of converting 5-FDRP **38** to 5-FDRulP **39**, opens up the possibility to reconstitute the pathway *in vitro* from SAM **34** and fluoride ion to generate 5-FDRulP **39**. Establishing this route to 5-FDRulP **39** will have connotations to the viability of reconstituting the complete pathways to 4-FT **33** and FAc **8** *in vitro*.

The assay for the MTRI-Sca in this case is incubation of the protein with the fluorinase, and PNP enzymes in the presence of SAM **34** and fluoride ion. The reaction was followed by ^{19}F NMR. The fluorinase and PNP genes were inserted into the *E. coli* expression vectors pET28(b) and pLou respectively and transfected into *E coli* (BL21 Gold) competent cells. Expression and purification of these proteins was identical to that of MTRI-Sca. They were purified to final concentrations of ~1 mg/ml in phosphate buffer (10 mM, pH 7.6). Equimolar amounts (0.1 mg) of these proteins were incubated together in the presence of SAM **34** (2 mM) and fluoride ion (50 mM) for 16 h at 37ºC. A control experiment was set up with the MTRI-Sca protein excluded from the reaction, and the experiments were run simultaneously. The reactions were stopped by heat deactivation (95ºC, 5 min) and centrifuged (12,000 rpm, 2 min). The resulting supernatant was made up to a volume of 700 µl, and D_2O (100 µl) was added. The resulting mixture was then subject to ^{19}F NMR, and typical spectra are shown in Figure 2.23.

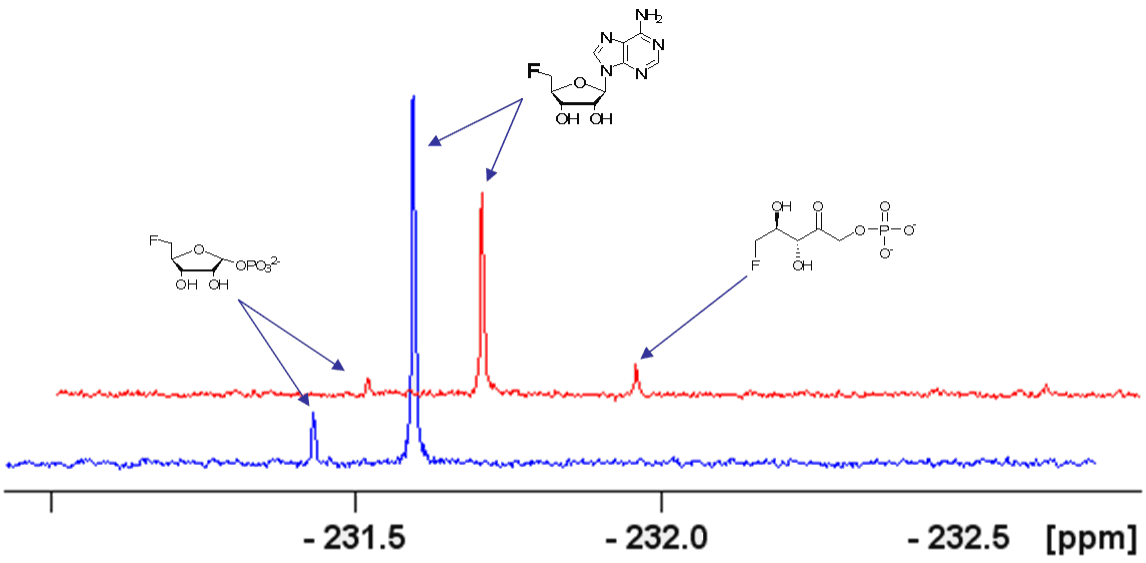


Figure 2.23. $^{19}F\{^{1}H\}$ NMR spectra of the MTRI-Sca protein in an *in vitro* pathway experiment starting from SAM **34** and fluoride ion. The blue spectrum represents a control experiment, without the MTRI-Sca protein. The red spectrum represents an experiment with the MTRI-Sca protein incubated alongside the fluorinase and PNP enzymes from *S. cattleya* overexpressed in *E. coli.*

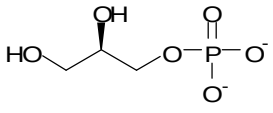
Figure 2.23 clearly shows the conversion of 5-FDRP **38**, generated by an *in vitro* reconstitution of the fluorometabolite pathway from SAM **34** and fluoride ion, to the product 5-FDRulP **39**. In the presence of MTRI-Sca, the ^{19}F NMR spectra revealed three organofluorine signals at -231.41 ppm, -231.60 ppm and -231.85 ppm respectively. In these experiments 5'-FDA **35** (-231.60) and 5-FDRP **39** (-231.41 ppm) remain, with 5'-FDA the major product of the reaction. This indicates that the PNP enzyme is acting as a bottleneck in both of these experiments. In the experiment containing MTRI-Sca, 5-FDRP is diminished, but is still present suggesting that the isomerisation reaction has at some point reached an equilibrium. Control experiments where the MTRI-Sca protein was excluded from the reaction show the accumulation of 5-FDRP **38** and 5'-FDA **35** only. Both were confirmed by the add-mixing of synthetic samples.[74, 97] These results

show that the MTRI-Sca is indeed capable of converting 5-FDRP **38**, generated from fluoride ion, to 5-FDRulP **39**.

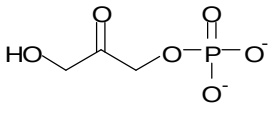
2.4.4 Isothermal titration calorimetry of *MTRI-Sca* with putative substrates

With the isomerase from *S. cattleya* in hand, it became a focus to identify the nature of binding at the active site of the enzyme. Two putative compounds were identified as structural homologs to the ribulose-1-phosphate moiety. They were dihydroxyacetone phosphate (DHAP) and glycerol-3-phosphate (LG3P).

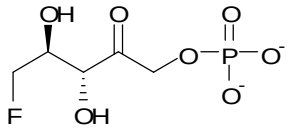
DHAP is a key intermediate in metabolism and is involved the Calvin cycle, ether lipid biosynthesis and L-methionine salvage pathways. Its major biochemical role is in the metabolic pathway of glycolysis, where it is a breakdown product of fructose-1-phosphate. L3GP is also an intermediate in glycolysis, and is responsible for the entry of glycerol into the glycolytic pathway. The glycerol-3-phosphate shuttle is used to rapidly generate NAD^+ in the brain and skeletal muscle through the activity of glycerol-3-phosphate dehydrogenase. The oxidation of L3GP at C2 position yields DHAP. Both L3GP and DHAP possess structural similarities to the ribulose-phosphate substrate for MTRIs (Scheme 2.5).



L3GP



DHAP



5-FDRulP 39

Scheme 2.5. Structural similarities between L3GP, DHAP and 5-FDRulP **39**.

Due to this structural similarity it was thought that LG3P and DHAP may display some affinity for the active site of MTRI-Sca. Isothermal calorimetry (ITC) was used to try and explore this affinity.

Accordingly, the MTRI-Sca protein was over expressed and dialysed into Hepes buffer (10 mM, pH 7.8) to a final concentration of 20 µM (1.6 mg/ml). Both DHAP and L3GP (Sigma Ltd, UK) were also suspended in the same buffer to a concentration of 600 µM. Solutions were degassed with a Thermovac degasser (Microcal Inc, USA). The protein solution was applied to the sample cell of a VP-ITC Microcalorimeter (Microcal Inc, USA) and the DHAP solution was loaded to the injection syringe and the cell jacket was equilibriated at 25 ºC. An initial injection of 2 µl over 10 s followed by 180 s equilibriation, was followed by 29 subsequent injections of 5 µl over a 10 s time period with 180 s between injections. The subsequent results are shown in Figure 2.24.

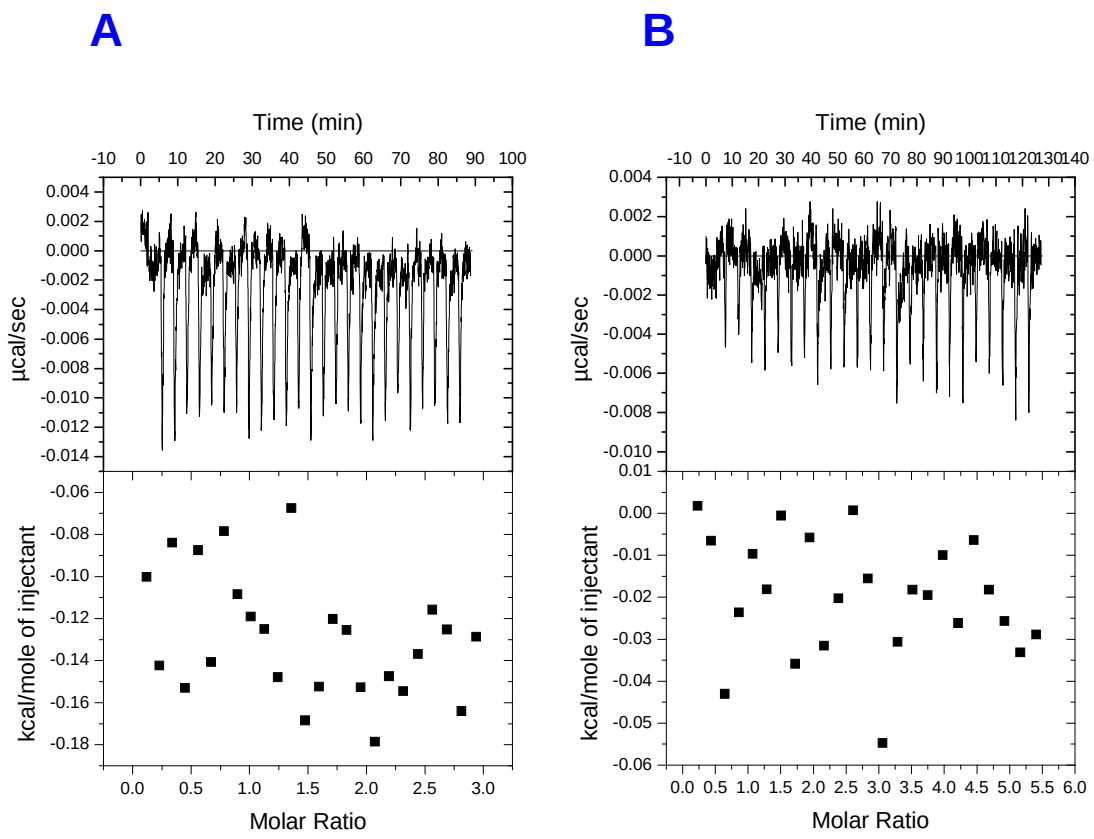


Figure 2.24. The ITC results following incubation of the MTRI-Sca in Hepes buffer (10 mM, pH 7.8) with **A.** DHAP and **B.** L3GP.

Results for the ITC of MTRI-Sca with DHAP and L3GP did not reveal any heat of binding to the active site of the enzyme for either putative ligand. Both of these molecules have fewer hydrogen bonding sites available than the confirmed isomerisation product 5-FDRulP **39**, and as a result may not possess the necessary characteristics to initiate efficient binding.

2.4.5 A Role for *MTRI-Sca* in *S. cattleya*

The L-methionine salvage is a critical process in many organisms, and the enzymes involved are clearly part of primary metabolism. In *S. cattleya*, the fluorinase and the PNP are located in a 10 kb gene cluster.[83] Fluorometabolite production occurs after 5 days of incubation in optimal medium for growth, suggesting that fluorometabolite production occurs as a consequence of secondary metabolism.[72] This raises the question of whether this enzyme is a primary metabolism enzyme, that exists as part of the L-methionine salvage pathway but that it is also capable of elaborating fluorometabolites i.e. it may have a dual role. However it may only be a secondary enzyme, expressed later in the growth cycle of *S. cattleya* specifically to perform this step in fluorometabolite biosynthesis. This is unresolved.

Some clues to the origin of MTRI-Sca can be found by using information from the genomic DNA of its homolog from *S. coelicolor*, SCO3014. Figure 2.25 shows the location of SCO3014, and the genes that are located next to it in the genomic DNA. Unlike *B. subtilis* and *S. cerevisiae*, the MTRI gene is not located in a gene cluster consisting of the component parts of the methionine salvage pathway.

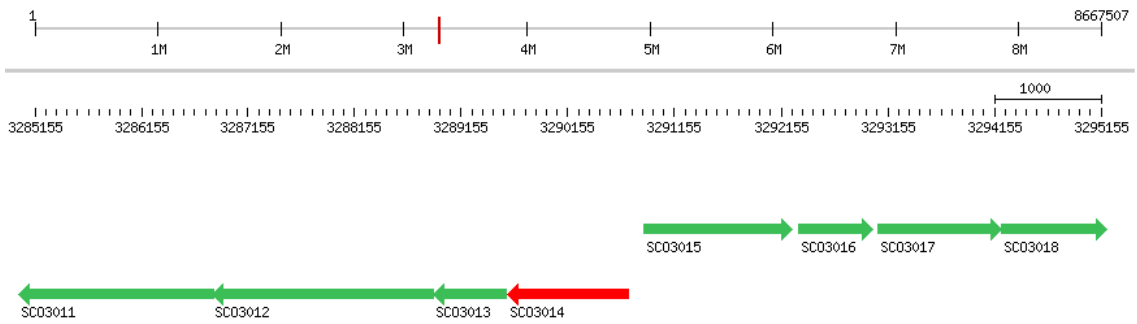


Figure 2.25. *SCO3014* and the organisation of the surrounding genes in the genomic DNA of *S. coelicolor*. SCO3011= putative lipoprotein. SCO3012= putative two-component system histidine kinase. SCO3013= putative two-component system response regulator. SCO3014= MTRI (highlighted in red). SCO3015= putative integral membrane protein. SCO3016= putative integral membrane protein. SCO3017= putative secreted protein. SCO3018= putative regulatory protein.[180]

The genes surrounding SCO3014 are documented in Figure 2.25. In the identification of MRTI-Sca, some information concerning the surrounding genomic DNA was uncovered during gene walking. Immediately downstream from the *MTRI-Sca*, 70 bp of genomic DNA were also revealed. Translation of this DNA into its subsequent peptide sequence and BLAST search against the genomic database of *S. coelicolor* identified that this DNA corresponds to a homolog *SCO3013*, the gene immediately downstream of the MTRI in *S. coelicolor* (Figure 2.26). Although this information is very limited, on the face of it, it implies that the *MTRI-Sca* is organised within the genomic DNA of *S. cattleya* in the same way.

atgggtgatcagtccgtacagccttggccaagggcacggggtccgggaccccggagccgaaacccgctctccgctggg
aagagcctcccgaagggcccgtgctggtcctcctcgaccagacccggctccccgtcgaggaggtcgaactgttctgtacg
gacgtgcccgcgctcgtccaggccatccgtaccctcgccgtccgcggcgcgccgctgctcgggctccgccggagcgtacg
gcgtcgccctggccgccgcccgtggctacgacgtcgggcaggccgccgacgaactcgccggcgcccggcccaccgcc
gtcaacctctcctacggggtgcgccgcgcgctggccgcgtaccgtaccgcggtcaccggcggcgccgacgacacgggc
gcggcggcggccaccctcgccgaggcccgcgcgctgcacgccgaggacgccagggccagcgaacgcatggcccgc
aacggcctggcgctgctggacgaactcgtccccggcggcggctaccgggtgctgacccactgcaacaccggcgcccctg
gtctccggcggcgagggcaccgccctggccgtcgtcctcgccgcccaccgcggcggactcctgcgccggctgtgggtg
gacgagacccggccgctgctccagggcgcccggctgaccgcctacgaggccgcccgggccggcgtcgcccacacgtt
gctgccggacggcgcggccgggtcgctcttcgcggccggcgaggtcgacgcggtgctgatcggcgccgaccggatcg
ccgcggacggctcgaccgccaacaaggtcggcagctacccgctggccgtcctcgcccggtaccacaacgtcccttcgt
cgtggtcgcccccaccaccacgatcgacctggccaccccccgacggcaccgcgatcgaggtggagcagcgccccgcgc
aggaggtgaccgagctgaccggaccgcgcccggcccggaccgcgagggcgccaccggcatccccgtcgcgccccct
gggcacgccggcgtacaacccggcgttcgacgtcaccccgcccgaactgatcaccgccgtggtcaccgagaccggcgt
ggcctcccggtcaccggctcctccatagccgccctggccgcccgccccggccccgtccgcgcccagccgtgacggcc
tcgacggtcgtcatccgccgccgatgacgaccgtcacccgtcaccgatgacgaccgtcccttccg

Figure 2.26. *MTRI-Sca* ORF and followed by a 70 bp sequence, homologous to *SCO3013* from *S. coelicolor* (highlighted in red).

The generation of a cosmid library of the genomic DNA of *S. cattleya* was carried out by Dr Hai Deng (University of St Andrews). Interestingly it was discovered that both *MTRI-Sca* and the fluorometabolite cluster were present in the same 40 kb fragment, i.e they were present on the same cosmid. The exact location of *MTRI-Sca*, and its position relative to the 10 kb gene cluster remains to be identified by the genome mapping project of *S. cattleya* that is currently being undertaken in a collaboration with the University of Edinburgh.

2.5 Conclusions

Two enzymes have been identified from *Streptomyces* that are capable of performing the isomerisation of 5-FDRP **38** to the subsequent intermediate in the fluorometabolite pathway of *S. cattleya*, 5-FDRulP **39**. The MTRI identified from *S. coelicolor* was initially annotated as a putative translation initiation factor, a common error associated with MTRIs. As a result of these studies, this has been corrected. Alignments of this

putative peptide sequence reveal that conserved residues associated with the catalytic activity of MTRIs were present. Amplification of the gene and its insertion into an *E. coli* expression vector allowed the over-expression of the SCO3014 protein in good yield (~5 mg/L). The incubation of the over expressed SCO3014 protein with 5-FDRP **38** revealed that 5-FDRulP **39** had been generated.

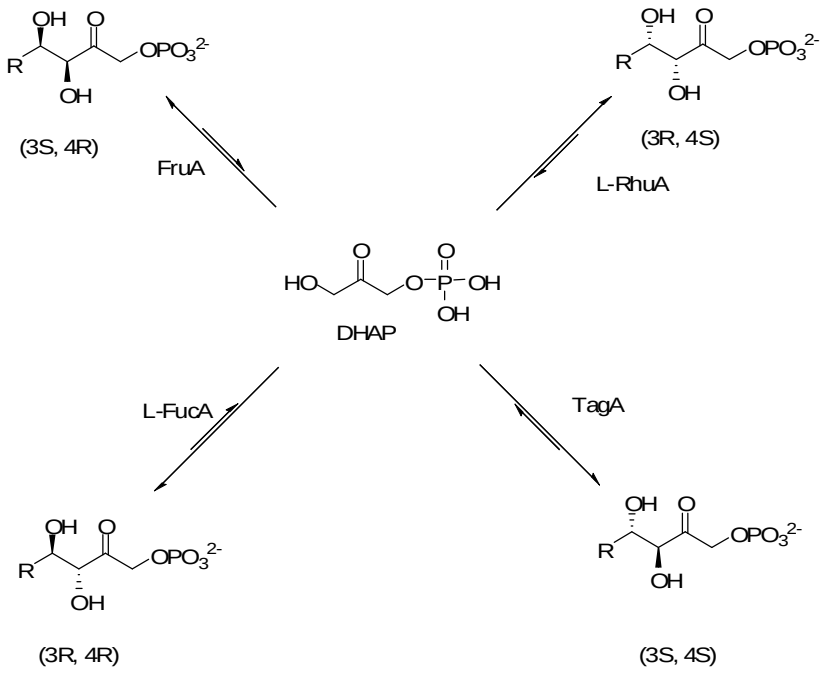
This demonstrated that enzymes of this class are capable of isomerisation of the fluorinated substrate analogue and gave a target for the identification of its homolog which may be involved in the fluorometabolite pathway in *S. cattleya*. The SCO3014 and SAV6421 (a putative MTRI from *S. avermitilis*) were aligned to identify conserved regions for the design of degenerate PCR primers. Degenerate PCR and gene walking revealed a 1161 bp ORF, which were translated and through BLAST search analysis against the *S. coelicolor* genome revealed high homology with the SCO3014 peptide sequence. The subsequent amplification of this ORF and insertion into an *E. coli* expression vector allowed the over-expression of the putative MTRI in good yield (~5 mg/L). Incubation of this protein with 5-FDRP **38** generated from synthetic 5'-FDA **35**, and generated from fluoride ion in the presence of the fluorinase and PNP enzymes from *S. cattleya* both produced the 5-FDRulP **39** intermediate. Consequently, an enzyme capable of the isomerisation reaction of the fluorometabolite pathway was identified. It remains to be confirmed whether *MTRI-Sca* is an enzyme of primary or secondary metabolism, a question that may be answered by genome mapping of *S. cattleya*.

3 DHAP aldolases from *Streptomyces*

This chapter describes the over expression of a dihydroxyacetone phosphate (DHAP) dependant aldolase from *S. coelicolor*. The enzyme was isolated with a view to identifying the gene of its homolog from *S. cattleya*, an enzyme which is involved in fluorometabolite biosynthesis, and generates fluoroacetaldehyde **40** from 5-fluoro-5-deoxyribulose phosphate (5-FDRulP) **39** in a retro-aldol reaction.

3.1 Dihydroxyacetone phosphate (DHAP) dependant aldolases

DHAP dependant aldolases are of interest synthetically as they form a C-C bond and generate two stereogenic centres by the aldol addition of DHAP and an aldehyde acceptor. DHAP aldolases can generate four different stereochemical outcomes and each is catalysed by an individual subset of enzymes within the DHAP aldolase family (Scheme 3.1). Because these aldolases construct two stereogenic centres in one reaction they are valuable biotransformation catalysts.[136, 137, 138, 139, 140] All of these aldolases show absolute substrate specificity for DHAP, however they can accept a variety of different aldehyde electrophiles giving product diversity as biotransformation catalysts.

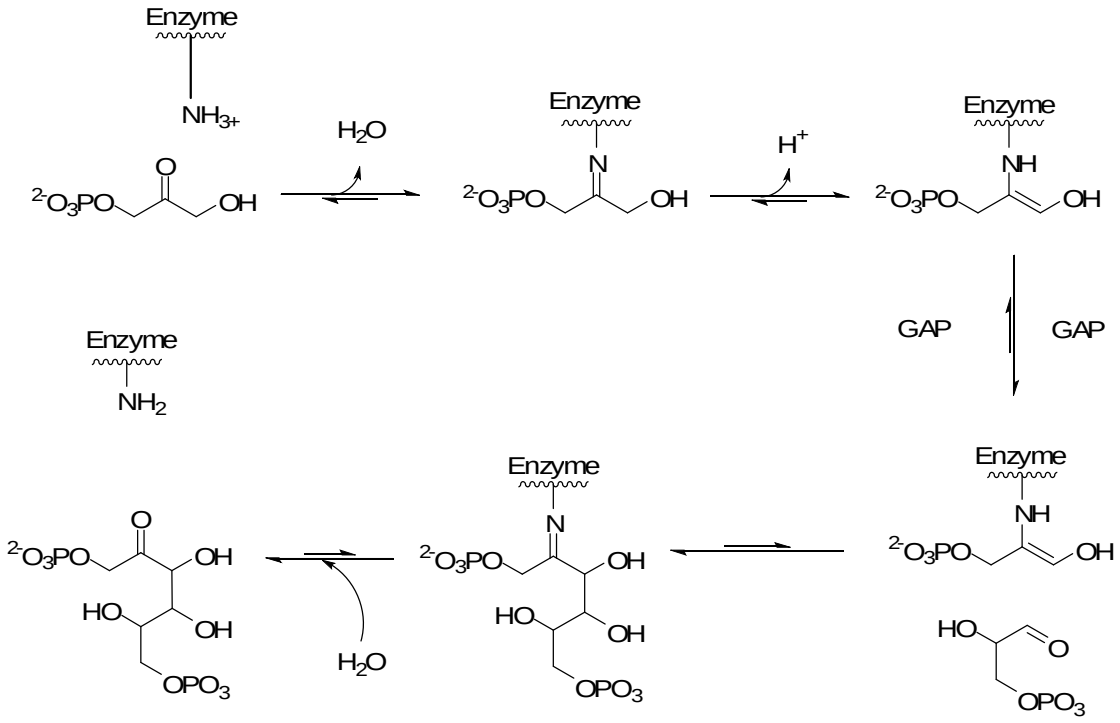


Scheme 3.1 L-fuculose-1-phosphate aldolase (L-FucA) (E.C. 4.1.2.17), L-rhamnulose-1-phosphate aldolase (L-RhuA) (E.C. 4.1.2.19), L-tagatose-1,6-bisphosphate aldolase (L-TagA) (E.C. 4.1.2.40) and L-fructose- 1,6-bisphosphate aldolase (L-FruA) (E.C. 4.1.2.13)[141, 142] R= OH.

3.1.1 Mechanism of class I aldolases

DHAP dependant aldolases fall into Class I or Class II. The different classes catalyse identical reactions however the mechanisms by which they operate are different. [141, 142] Class I aldolases are generally homotetrameric enzymes that form a Schiff-base intermediate during the catalytic cycle.[143] These enzymes are usually found in eukaryotes or higher organisms, although Class I aldolases have been reported in prokaryotes.[144] The generally accepted mechanism for Class I aldolases proceed *via* the formation of a Schiff-base intermediate between a γ-amino lysyl group at the active site, and the carbonyl of

DHAP.[145, 146] A general mechanism for the Class I fructose 1,6-bisphosphate aldolase is shown in Scheme 3.2.

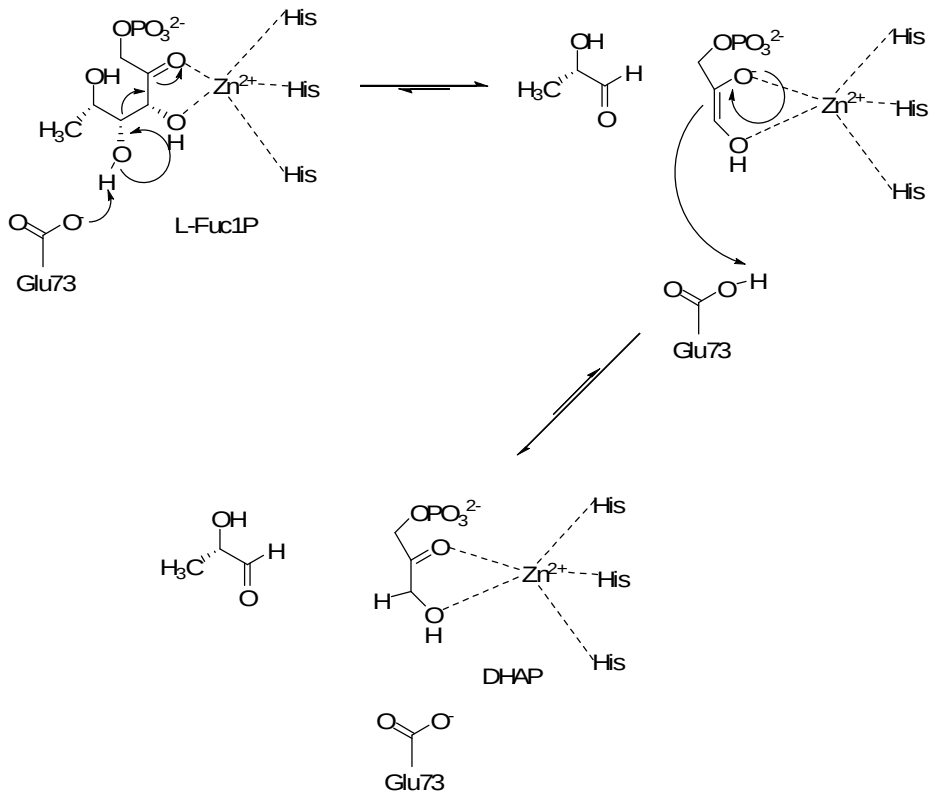


Scheme 3.2. General mechanism for Class I L-fructose 1,6-bisphosphate aldolase (L-FruA).[8]

3.1.2 Mechanism of Class II DHAP aldolases

Class II DHAP dependant aldolases are usually found in prokaryotes and lower eukaryotic organisms such as yeast, fungi and algae.[147] Class II enzymes are homodimeric and require a divalent metal ion, usually Zn^{2+} as an essential Lewis acid co-factor. As a result, they can be inhibited by chelating compounds such as EDTA, which sequester the Zn^{2+}, inactivating the enzyme. The mechanism of Class II aldolases is highlighted in Scheme 3.3. The mechanism for this reaction in the aldol direction is initiated by deprotonation at C3 of DHAP by the glycolic acid side chain of the glutamic

acid 73 residue of L-FucA from *E. coli* (Glu73). The resultant ene-diolate is stabilized by Zn^{2+} complexation. A nucleophilic attack at the carbon *si* face of the aldehyde results in subsequent C-C bond formation.[147, 148]



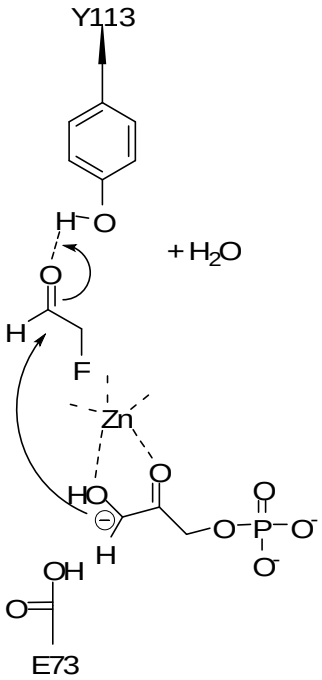
Scheme 3.3. Mechanism of the Class II DHAP dependant aldolase, L-FucA from *E.* coli.[148]

3.2 DHAP aldolases from *S. cattleya*

Previous work in the research group has identified two DHAP aldolases from *S. cattleya* that were capable of utilizing FAld **40** as a substrate in conjunction with DHAP.[95] These aldolases were first identified by the observation that two diastereoisomers, 5-FDRulP **39** and 5-FDXulP **42** accumulated on incubation of FAld **40** and DHAP in cell free

extracts.[95, 97] Stereochemical analysis of DHAP aldolase products using non-fluorinated substrates indicated that L-FucA generates the (*3R, 4R*) stereoisomer (Scheme 3.1).[141, 142] This is consistent with the action of a L-fuculose aldolase, to generate 5-FDRulP **39,** which is a proven intermediate in flurometabolite biosynthesis in *S. cattleya*.[97]

L-FucA, as well as the other Class II aldolases, are capable of utilising a wide variety of aldehyde substrates. Studies have shown chloroacetaldehyde is accepted as a substrate by L-FucA from *E. coli*.[141] A suggested mechanism for FAld **40** incorporation in 5-FDRulP **39** synthesis is shown in Scheme 3.4.



Scheme 3.4. Suggested catalytic mechanism for fuculose aldolase using FAld **40** as a substrate.

Attempts to isolate the native L-FucA from active cell free extracts of *S. cattleya* failed, however in that effort, a L-FruA aldolase responsible for 5-FDXulP **42** production was

successfully purified.[14] It therefore became a research focus to identify the alternative L-FucA gene from *S. cattleya* in order to over-express this enzyme and explore its role in fluorometabolite biosynthesis.

3.3 Identification of an L-fuculose aldolase from *S. cattleya*.

In an attempt to identify the L-FucA gene a reverse genetics approach was used. A BLAST search using L-FucA from *E. coli* (E.C. 4.1.2.17) revealed several putative aldolases from a number of different bacteria including *Streptomyces coelicolor, Streptomyces avermilitis, Saccharopolyspora erythraea, Rubrobacter xylanophilus* and *Methylobacterium extorquens*. The peptide sequences for these aldolases are aligned in Figure 3.1. The alignment reveals that they all possess the key catalytic residues, Glu73 and Tyr112, identified from L-FucA from *E. coli*.[147]

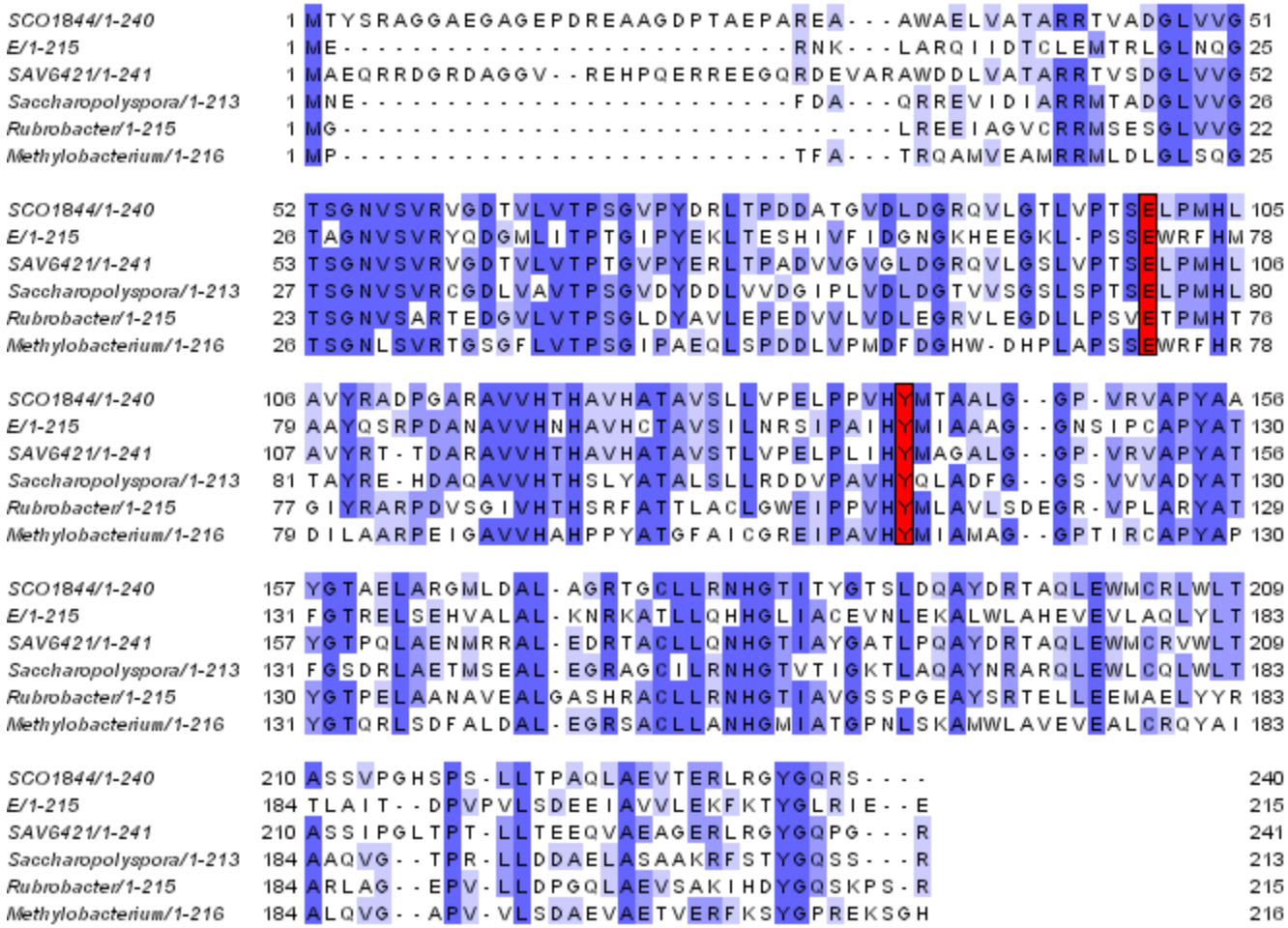


Figure 3.1. Alignment of the peptide sequences of L-FucA from *E. coli* and the putative L-FucAs SCO1844, SAV6421, *Saccharopolyspora erythraea*, *Rubrobacter xylanophilus* and *Methylobacterium extorquens* respectively. These homologs were identified by BLAST search. Conserved catalytic residues identified in L-FucA (Glu73 and Tyr113) are highlighted in red. Other conserved residues are highlighted in shades of blue; the dark blue residues are absolutely conserved throughout this family of aldolases.

The peptide sequence of L-FucA from *E. coli* was used to screen against the proteome databases of *S. coelicolor* and *S. avermilitis*, the two *Streptomyces* organisms which have been the subject of full genome sequence analysis.[135, 149] This search highlighted a gene, *SCO1844*, from *S. coelicolor* which had a 40% identity to *E. coli* L-FucA, and it also highlighted the *SAV6421* gene from *S. avermilitis* with 42% identity to this gene

(Figure 3.2). The *SCO1844* is a gene of 723 bp and encodes a protein of 240 amino acids with a molecular weight of 25 kDa (24988 Da). SAV6421 is a gene of 726 bp, encoding a protein of 241 amino acids and with a molecular weight of about 25.8 kDa (25891 Da). Both of these genes are annotated as putative fuculose aldolases. When the peptide sequences of these *Streptomyces* genes are aligned, highly conserved regions between the two become obvious. These regions include the putative catalytic residues (Figure 3.2).

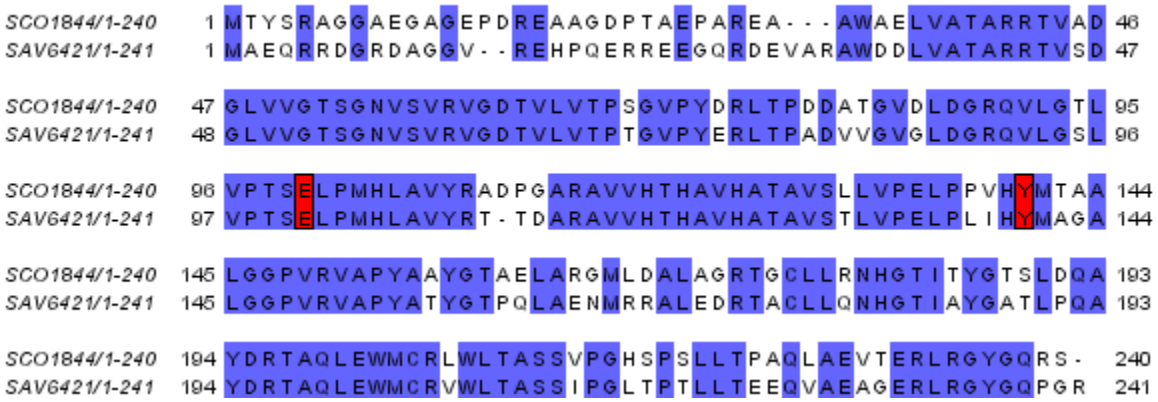


Figure 3.2. Peptide sequence alignment of SCO1844 and SAV6421 respectively showing 70% sequence identity. Catalytic residues are highlighted in red. Conserved residues are highlighted in blue.

3.3.1 Degenerate PCR primer design

Alignment of the peptide sequences of the putative L-FucAs, *SCO1844* and *SAV6421*, revealed that they possess 70% sequence identity and that there are many regions of homology. These regions are highlighted in blue in Figure 3.2. Conserved peptide sequences consisting of about 7 to 10 amino acids were identified from the alignment of these two proteins. These short peptide sequences were subject to reverse translation, in order to reveal the possible DNA sequences that could code for the selected amino acids. Reverse translation revealed that there is a lot of degeneracy in the putative DNA codes

i.e. there are many different combinations of DNA that could code for the amino acid sequence in question. The way in which DNA is translated to generate proteins differs from organism to organism. In order to best guess the DNA sequence present in *S. cattleya*, and design more accurate degenerate DNA primers for PCR, a codon usage table for *S. cattleya* was used.[150] The codon usage table documents all of the identified proteins from *S. cattleya* and the DNA sequences that encode them. Each amino acid is encoded by a DNA triplet, and the possible combinations of triplet DNA codes are expressed as a percentage of their occurrence in the genes of proteins from *S. cattleya*. It is therefore possible to interpret the output from reverse translation, in terms of the likelihood that these sequences would arise in the L-FucA gene from *S. cattleya*. Those DNA sequences that are rare in *S. cattleya* were disregarded in the design of candidate degenerate primers to reduce the degeneracy. The subsequent primers are tabulated in Table 3.1. DNA sequences were identified that occur most often in protein expression from *S. cattleya* for designing degenerate primers. The primers were designed in two groups, the forward and reverse primers respectively (Table 3.1). The forward primers were designed to the predicted 5'-3' sequence of the gene as deduced from the amino acid sequence from the N-terminal end of the anticipated protein. The reverse primers were designed as 3'-5' complements of the gene from amino acid sequences nearer the C-terminal domain of the putative protein.

3.3.2 Degenerate PCR for fuculose aldolase in *S. cattleya*

The forward and reverse degenerate primers were then used in combination with each other in PCR reactions using the Taq DNA polymerase in the presence of *S. cattleya*

genomic DNA, i.e. each forward primer was used in conjunction with each reverse primer in a trial and error manner.

Forward Primers	**Peptide Sequence**	**Degenerate Primer 5'-3'**
Fuc Ald A	LVVGTSGN	tsgtsgtsggsacswssggsaacg
Fuc Ald B	LDGRQVLG	cctsgacggsrgscacgtsctstac
Fuc Ald C	ELPMHLAVY	garctsccsatgcacctsgcsgtstac
Fuc Ald D	TSGNVSVRV	ccwssggcaacgtswssgtsggc
Fuc Ald E	LVPELP	ctsgtsccsgagctsccscys
Fuc Ald F	NVSVRVGD	saacgtswssgtsmgsgtsggsgac
Fuc Ald G	HTHAVHA	scacgcsgtscacgcsacsgc
Fuc Ald H	GVPY(D/E)RLTP	gtsccstacgasmgsctsacscc
Reverse Primers	**Peptide Sequence**	**Degenerate Primer 5'-3'**
Fuc Ald I	ERLRGYGQ	gtgsccgtasccscysagsckytc
Fuc Ald J	ETAQLEWMCR	cygcacatccaytcsaggtgsgcsg
Fuc Ald K	PVRVAPYA	cgtasggsgcsacgccsacsg
Fuc Ald L	LVPELP	srgsggsagctcsggsacsag
Fuc Ald M	QAYDRTAQ	ctgsgcsgtsckgtcgtasgcctg
Fuc Ald N	ALGGPVR	gcgtasggsgcsacscksacsg

Table 3.1. Degenerate primers designed from conserved regions of SCO1844 and SAV6421. S= g or c, W= a or t, R= a or g, M= c or a, Y= c or t, K= g or t.

The length of the expected PCR product was approximated using the distance in amino acid residues between targeted regions in the SCO1844 and SAV6421 peptide sequences shown in Table 3.2. The PCR products from these degenerate combinations were analysed using 1% agarose TAE gel electrophoresis alongside a 1kb DNA ladder. The

subsequent gels were imaged using a UV lamp. Figures 3.3 and 3.4 show examples of these degenerate PCR products and the subsequent DNA gel analysis.

Forward Primers / Reverse Primer	A	B	C	D	E	F	G	H
I	545	441	414	558	321	249	336	504
J	477	357	315	459	222	450	258	405
K	327	207	165		72	300	108	255
L	270	150	108	252	15	243	33	198
M	456	339	297	441	204	432	240	387
N	309	192	150	294	57	300	92	240

Table 3.2. The approximate expected length (base pairs) of amplified DNA fragments using different combinations of degenerate primers from Table 1. These were estimated using the peptide sequences of SCO1844 and SAV6421.

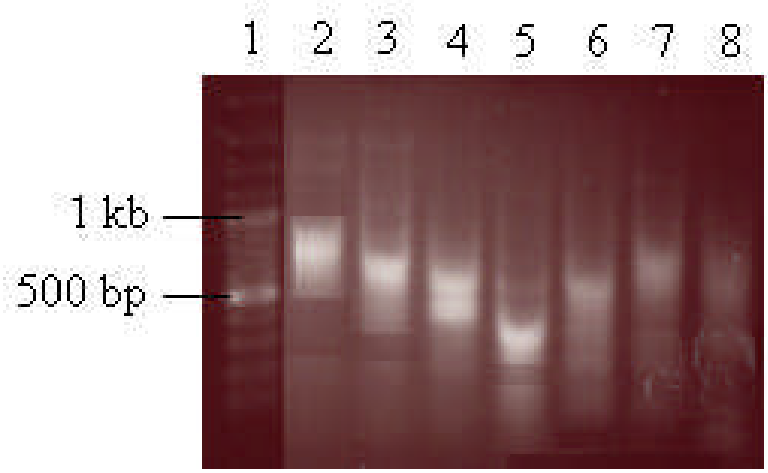


Figure 3.3. UV photograph of 1% Agarose DNA Gel containing degenerate PCRs for L-FucA from *S. cattleya* genomic DNA. Different lanes represent different combinations of forward and reverse primers: **1**= 100 bp DNA Ladder. **2**= Fuc Ald F and Fuc Ald M. **3**= Fuc Ald F and Fuc Ald N. **4**= Fuc Ald F and Fuc Ald J. **5**= Fuc Ald G and Fuc Ald M. **6**= Fuc Ald G and Fuc Ald N. **7**= Fuc Ald G and Fuc Ald I. **8**= Fuc Ald G and Fuc Ald J

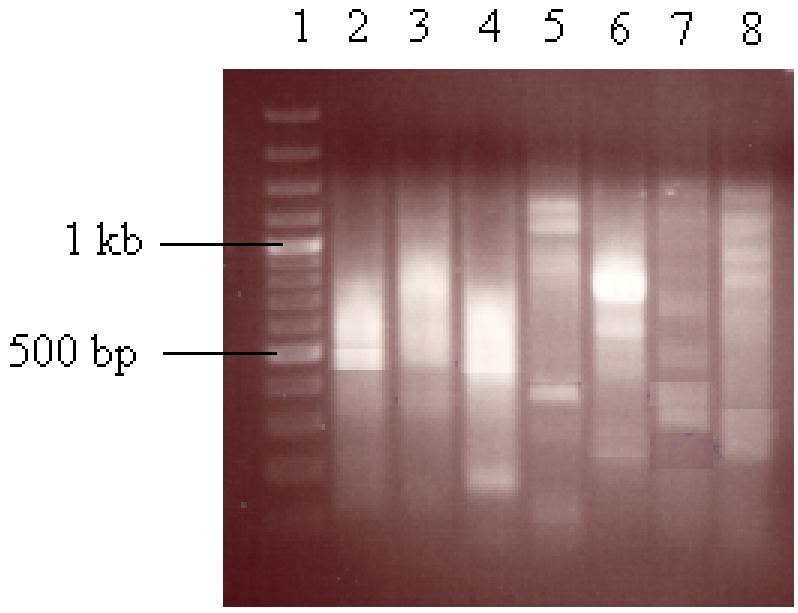


Figure 3.4. UV photograph of 1% Agarose DNA Gel containing degenerate PCRs fo L-FucA from *S. cattleya* genomic DNA. Different lanes represent different combinations of forward and reverse primers: **1**= 100 bp DNA Ladder. **2**= Fuc Ald A and Fuc Ald M. **3**= Fuc Ald A and Fuc Ald N. **4**= Fuc Ald A and Fuc Ald J. **5**= Fuc Ald B and Fuc Ald J. **6**= Fuc Ald C and Fuc Ald M. **7**= Fuc Ald C and Fuc Ald I. **8**= Fuc Ald C and Fuc Ald J.

Several candidate PCR products emerged from these experiments, and these were excised from the gel and purified into nuclease free water. The purified DNA was then ligated into pGem T easy vector (Promega, UK) using T4 DNA ligase. The recombinant plasmid was transfected into *E. coli* JM109 competent cells and positive clones were identified by ampicillin resistance and the blue/white colony assay.[151] Several white colonies from each PCR product were picked from the petri dish and inoculated into 5 ml LB media containing ampicillin, the cells were then grown overnight in 15 ml falcon tubes. Cells were then spun down and the supernatant was removed to reveal the cell pellet, from which the recombinant plasmids were purified into ultrapure water using the QIAprep Spin Miniprep Kit (Qiagen, UK) and prepared for DNA sequencing.

The resultant PCR product sequences were analysed by translating the confirmed DNA sequence to the corresponding amino acid sequence.[152] The peptide sequences were BLAST searched in an attempt to match them with peptide sequences from L-FucAs. This method has been successful in the past in our hands through the isolation of the isomerase gene from the genomic DNA of *S. cattleya*. However despite many attempts to isolate sequence associated with the L-FucA from *S. cattleya*, this trial and error method proved unsuccessful.

These L-FucA experiments have revealed the limitations of this method. High levels of conservation in the target amino acid sequences are important for the success of this approach and perhaps the target enzyme has a relatively low level of identity to the *S. coelicolor* and *S. avermilitis* enzymes. The full genome sequence of *S. cattleya* is currently being undertaken and there is a particular interest in identifying the gene responsible for L-FucA in this organism.

3.4 Amplification of a putative L-FucA gene from *Streptomyces coelicolor*

It remained a research focus to reconstitute the fluorometabolite pathway *in vitro* and therefore it was still necessary identify a surrogate L-FucA to reconstitute this pathway. We had two candidates in hand, thus over-expression of the putative L-FucA from *S. coelicolor*, *SCO1844* became an objective. It was clear from alignment of the peptide sequences from *E.coli* L-FucA and *SCO1844* that the catalytic residues identified in the *E. coli* L-FucA, are also conserved in *SCO1844* (Figure 3.5).[147]

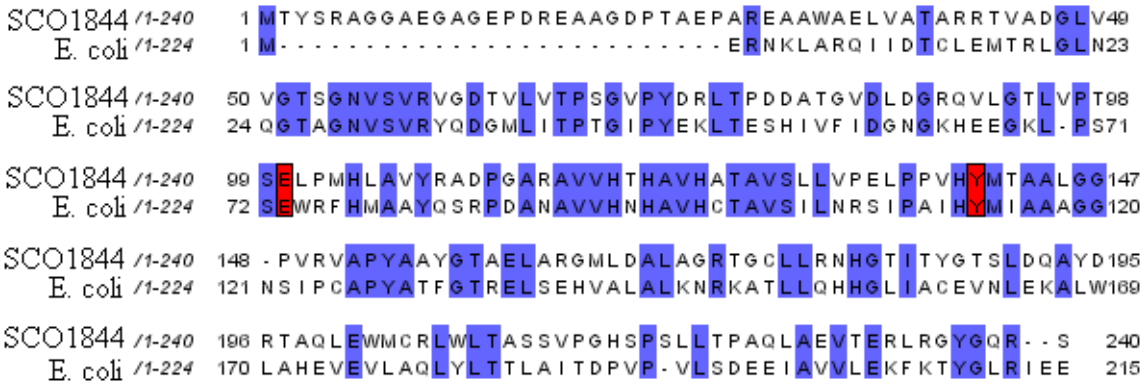


Figure 3.5. Peptide sequence alignment of L-FucA from *SCO1844* and *E. coli* L-FucA respectively. Catalytic residues Glu73 and Tyr113 are highlighted in red, conserved regions in blue.

Accordingly, specific DNA primers for SCO1844 were designed in the 5'-3' and 3'-5' direction with *EcoRI* and *Hind III* restriction sites respectively (Table 3.3). Genomic DNA was prepared from *S. coelicolor* and SCO1844 was amplified by PCR using the primers described in Table 3.3 and using the KOD DNA polymerase. This enzyme, isolated from the extreme thermophile, *Thermococcus kodakaraensis* KOD1, possesses high processivity levels similar to *Taq* polymerase with high-fidelity proofreading capacity similar to *pFu* polymerase. In the event a PCR product of ~700 bp was identified by DNA gel analysis (Figure 3.6). This product was then excised and purified to ~100 ng/µl of DNA.

Primer / Restriction Site	Sequence 5'-3'
Forward / EcoRI	CCTCCGCCGGAATTCATGACGTATTCGCGG
Reverse / Hind III	GAAGGAGCAAGCTTTCAGCTTCGCTGCCCG

Table 3.3. Specific DNA primers for amplification of SCO1844 from *S. coelicolor* genomic DNA.

The prepared PCR DNA, and the *E. coli* expression plasmid pHISTEV[21] were digested by *EcoRI* and *HindIII* restriction enzymes for 4 h. The DNA preparations were then purified again and introduced to each other in a 3:1 PCR:pHISTEV ratio in the presence of T4 DNA ligase for 16 h at 4°C. The ligation product was then transfected into competent *E. coli* BL21 Gold cells and colonies were selected for resistance to kanamycin. Resistant clones were then picked from the petri dish and subjected to colony PCR using the primers in Table 3.3. Those colonies that exhibited a PCR product of ~700 bp were then picked and inoculated in LB medium (5 ml) containing kanamycin in a 15 ml falcon tube (agitated for 16 h at 37ºC). The cells were then pelleted by centrifugation and the recombinant plasmid extracted and prepared for DNA sequencing using the QIAPrep Spin Miniprep Kit (Figure 3.7).

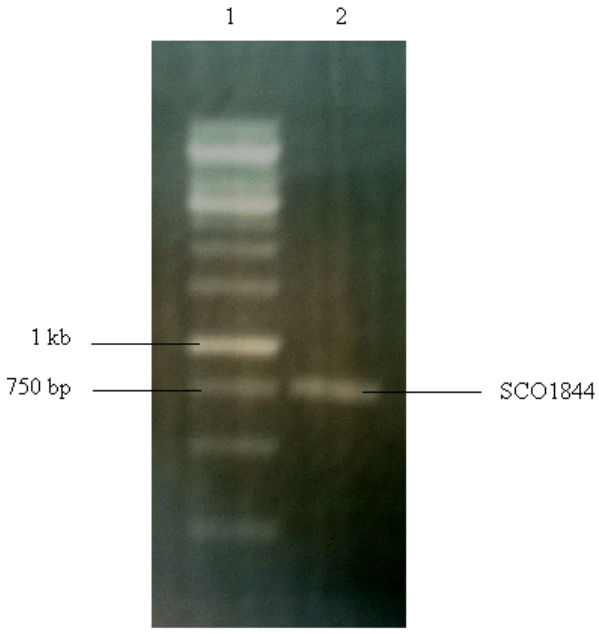


Figure 3.6. The 1% agarose gel containing the product of the PCR of SCO1844 from *S. coelicolor* genomic DNA using the primers from Table 3. 1= 1 kb DNA Ladder. 2= SCO1844 PCR.

acgggaaaattcccctctagaataatttgtttaactttaagaaggagatatacatatgtcgtactacaatcaccatcaccatcacgattacgatatcccaac
gaccgaaaacctgtattttcagggcgccatggctgatatcggatccgaattcATGacgtattcgcgggccggaggggcggaaggggccgggg
agccggaccgggaagcggccggggatccgaccgcggagccggcccgggaagcggcctgggcggaactcgtcgcgacggcccggc
ggacggtcgccgacggactggtggtgggcacctccggcaacgtctcggtgcgcgtcggcgacacggtcctggtcaccccgtcgggcgt
cccctacgaccggctgaccccggacgacgcgaccggcgtcgacctggacggacggcaggtgctcggcacccctcgtccccaccagcg
aactgcccatgcacctcgccgtgtaccgcgccgaccccggtgcccgggccgtcgtccacacccacgccgtgcacgcgacggcggtctc
gctgctcgtccccgaactcccgccggtgcactacatgaccgccgccctcggcggccccgtccgggtcgccccctacgcggcctacggc
accgccgaactcgcccgcggcatgctcgacgccctcgccggccgcaccggctgcctgctccgcaaccacggcacgatcacctacggc
acctccctcgaccaggcctacgaccgcaccgcccagctcgagtggatgtgccgcctgtggctgaccgcgtcctcggtgcccggccactc
gccgtccctgctgacgccggcccagctcgccgaggtgaccgaacgcctccgggggtacgggcagcgaagcTGAaagcttgcggccgc
actcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcatactagcat
acccccttgggcctctaaacgggtctgaggggttttgctgaaggagactatatcgatggcgatgggacgcgcctgagcggcgcataagcgcggcggtgt
ggtgttacgccaactgacgctacatgctgtgcttagccagctcttcgatttccatccttccgcacgttgaccggcttccgtcagta

Figure 3.7. Confirmed DNA Sequence of SCO1844-pHISTEV recombinant plasmid. Black = pHISTEV, Red = SCO1844.

3.4.1 Expression and purification from *E. coli* of a putative L-FucA from *Streptomyces coelicolor*

The presence of SCO1844 in pHISTEV was confirmed by DNA sequencing (Figure 3.7). The SCO1844-pHISTEV recombinant plasmid was then transfected into *E. coli* BL21 Gold competent cells and the SCO1844 protein was expressed after incubation with IPTG (1mM) for 16 h at 16ºC. Upon loading of the protein to the Ni^{2+} column, the nickel resin undergoes a very obvious colour change giving a brilliant turquoise. This is a very useful early indicator that SCO1844 has been expressed with a His_6-tag. Following Ni^{2+} chromatography, samples from the cell free extract, cell debris, supernatant, column flow through, column wash, and column elution were mixed with SDS dye at 95 ºC for 5 min. Samples were then loaded onto a 1 mm 4-12% Bis-Tris gel submerged in MES SDS running buffer for SDS PAGE analysis. The resulting SDS-PAGE revealed a band in the eluent fraction of ~25 kDa (Figure 3.8).

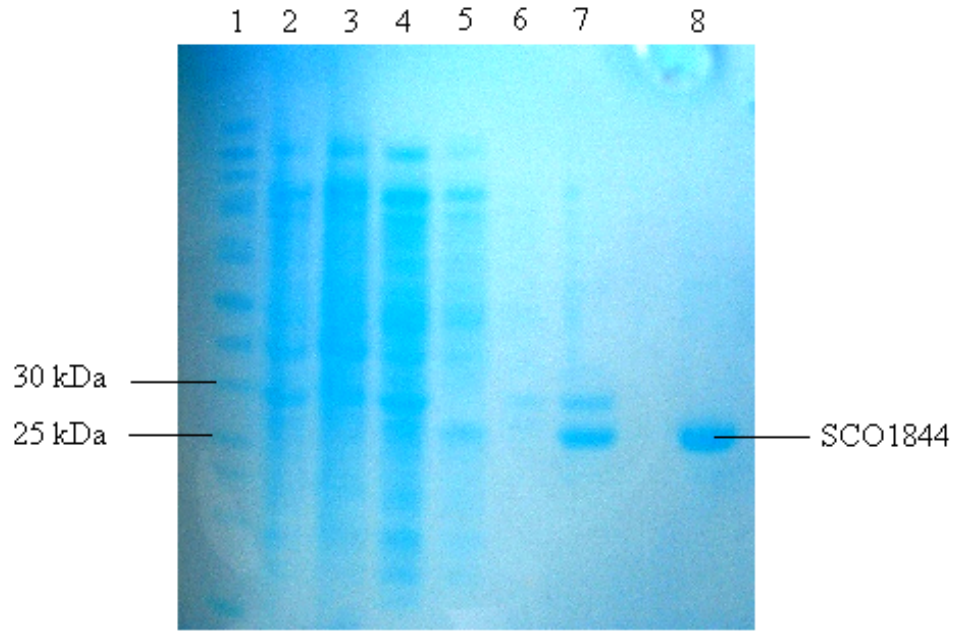


Figure 3.8. SDS PAGE of SCO1844 gene product purification from *E. coli* BL21 Gold. Lane 1= Protein Ladder. Lane 2= CFE. Lane. 3= Cell Debris. 4= Supernatant. 5= Column flow through. 6= Column wash. 7= Elution. 8= Gel Filtration.

The expression of SCO1844 was confirmed by in-gel tryptic digest and analysis of the resultant peptides by nanoLC-ESI MSMS (UltiMate (Dionex) and Q-Star Pulsar XL (Applied Biosystems)). The MS/MS data file generated was analysed using the Mascot 2.1 search engine (Matrix Science, London, UK) against an internal database consisting of a bacterial genome background to which the SCO1844 sequence (amongst others) had been added. The data was searched with tolerances of 0.2 Da for the precursor and fragment ions, using trypsin as the cleavage enzyme up to one missed cleavage was assumed. Carbamidomethyl modification of cysteines was a fixed modification and L-methionine oxidation was selected as a variable modification. The identity of the SCO1844 protein product gel band was confirmed with a Mascot Score of 673.

The fraction containing the SCO1844 protein was concentrated (2ml) and subjected to FPLC gel filtration (Figure 9). Gel filtration using phosphate buffer (10mM, pH 7.8) revealed that the active protein eluted between 67 and 77 ml (the highlighted region of Figure 9) with a protein concentration of 1.3 mg/ml. SDS-PAGE analysis of the pooled fractions containing SCO1844 revealed that following size exclusion chromatography, relatively pure protein had been attained (Figure 3.9). Data from size exclusion chromatography indicated that the soluble SCO1844 protein is dimeric, consistent with other Class II aldolases.[147, 148]

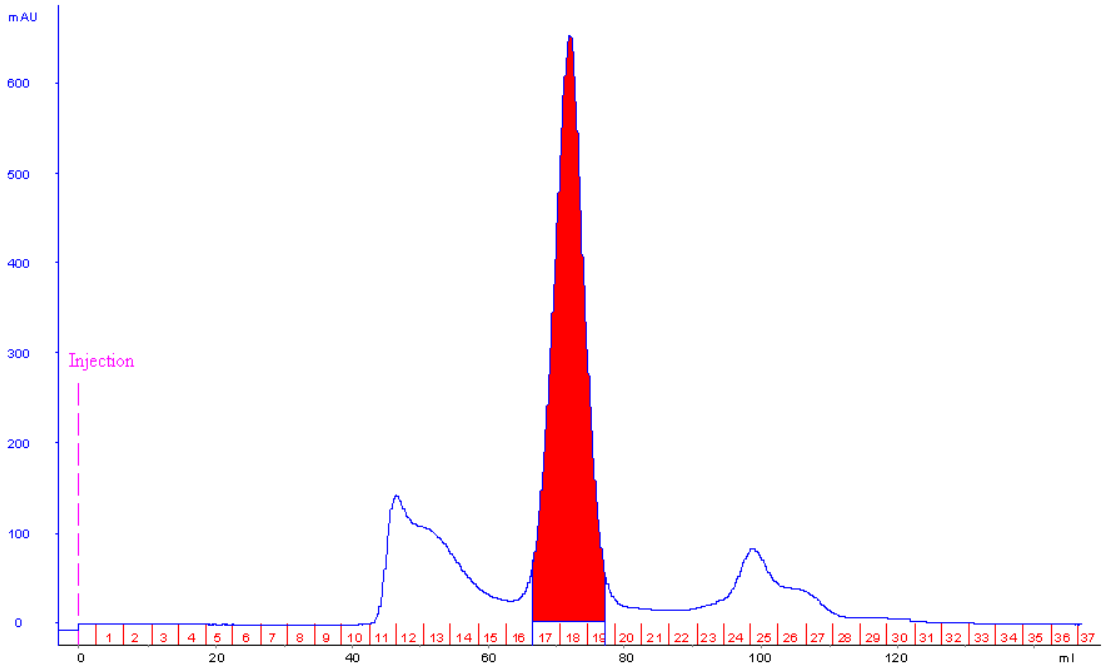


Figure 3.9. Chromatogram obtained by size exclusion using a Superdex 200 column (120 ml) after injection of a sample (2 ml) from Ni^{2+} chromatography of SCO1844. Revealing that the protein is a dimer. Highlighted area= Elution volume (~72 ml).

3.4.2 Enzymatic assay of the SCO1844 Protein

With the purified SCO1844 protein in hand, the objective was to examine its aldolase activity against appropriate substrates, i.e. could the SCO1844 protein perform the aldol reaction using FAld **40** as a substrate as suggested in Scheme 4? Accordingly an incubation of DHAP (1mM, Sigma Ltd UK) with synthetic FAld **40** (1 mM) generated from fluoroethanol[153] in the presence of SCO1844 (0.1mg) (37 °C, 6 h) was explored. The reaction was terminated by heat deactivation (95 °C, 5 min), and the reaction solution was centrifuged (12,000 rpm, 2 min). The subsequent supernatant was removed and the volume made up to 700 µl with ultrapure water, before adding D_2O (100 µl). The sample was then subjected to ^{19}F NMR, and a typical spectrum is shown in Figure 3.10.

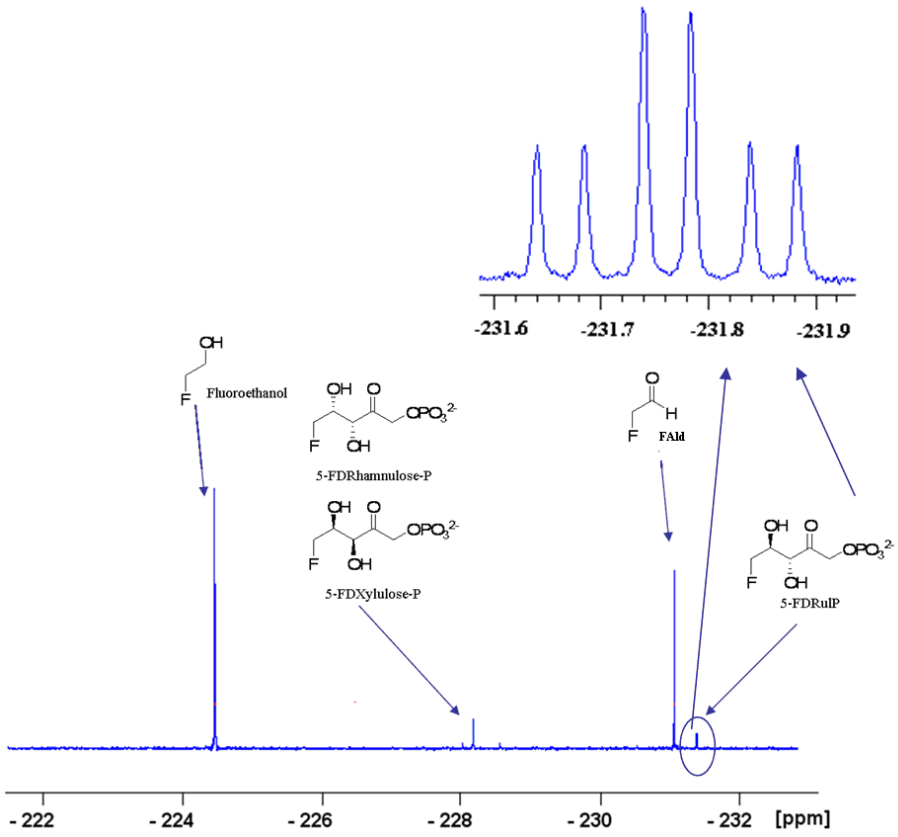


Figure 3.10. ^{19}F NMR spectra of the incubation of SCO1844 with DHAP (1mM) and FAld **40** (1 mM) for 6 h at 37 °C. Aldol products 5-FDRulP **40** and 5-FDXulP **42**/ 5-FDRhuP **41** are clearly identifiable.

^{19}F NMR analysis of the aldol reaction with SCO1844 L-FucA indicated residual FAld **40** and fluoroethanol. Fluoroethanol originates from the starting preparation of FAld **40** by incomplete oxidation. However, two new ^{19}F NMR signals were observed at -228.15 ppm (dt, $^{2}J_{F,H}$ 47.0 Hz and $^{3}J_{F,H}$ 16.0 Hz) and -231.36 ppm (dt, $^{2}J_{F,H}$ 47.0 Hz and $^{3}J_{F,H}$ 16.0 Hz) in a 2:1 ratio respectively (Table 3.4). These signals were absolutely absent in control experiments without protein. The signal at -231.36 clearly correlates to 5-FDRulP **39** which had already been established as an intermediate in the fluorometabolite pathway in *S. cattleya*.[97]

The major product with the chemical shift -228.15 is difficult to assign definitively. 5-FDXulP **42**, the (*3S, 4R*) diastereoisomer to 5-FDRulP **39,** has previously been identified as a product of a fructose aldolase purified from cell free extracts of *S. cattleya* (Scheme 3.1) and possesses a similar chemical shift by ^{19}F NMR.[97, 141, 142] This product may also be the alternative (*3R, 4S*) diastereoisomer, 5-fluorodeoxyrhamnulose-1-phosphate **41** (5-FDRhuP) (Scheme 3.1). We tentatively suggest that this product is the 5-FDRhuP **41** diastereoisomer. In order to enzymatically generate the 5-FDXulP **42** diastereoisomer, configurational inversion at C3 would be required. As it is accepted that the absolute configuration at the C3 position is conserved upon reaction with electrophiles this outcome appears unlikely.[144, 148, 154, 155, 156, 157] Also, if an epimerisation event was occurring to generate 5-FDXulP **39**, a diastereoisomer product ratio of 1:1 would be expected, in this case a 2:1 ratio is observed experienced (Table 3.4). This may be exacerbated especially with "alien" FAld **40** as a substrate which may promote some level of epimerisation of the products.

Generation of the 5-FDRhuP **41** diastereoisomer would be consistent with the widely observed lack of full stereospecificity of fuculose aldolases. Aldol products with the opposite configuration at the C4 position are common. The stereointegrity of these enzymes is dependant on the orientation of the aldehyde substrate in the active site pocket. C4 stereochemistry is determined by the presentation of the aldehyde's *si* or *re* face for attack by the enzyme-DHAP-enediolate complex (Scheme 3.3). It is conceivable that this could account for the production of 5-FDRhuP **41**. Previous observations in the literature suggest that the major product of the SCO1844 aldol reaction is 5-FDRhuP **41**, however the two possible outcomes are difficult to distinguish as they possess the same ^{19}F NMR characteristics. It may be that the signal at -228.15 ppm (Figure 3.10) is a result of a mixture of these enantiomers.

	10 mM Phosphate Buffer pH 7.6 37°C, 6h Integration.	**10 mM Phosphate Buffer pH 7.6 RT, 24h Integration.**	**10 mM Phosphate Buffer pH 7.6 4°C, 24h Integration.**	**10 mM Phosphate Buffer pH 7.6 + 10 µM $ZnSO_4$ 37°C, 6h Integration.**	**10 mM Phosphate Buffer pH 7.6 + 1 mM EDTA 37°C, 6h Integration**
Fluoroethanol	1.0	1.0	1.0	1.0	1.0
FAld 40	0.79	0.40	0.50	0.76	0.79
5-FDRulP 39	0.07	0.21	0.30	0.04	0
5-FDRhuP 41/ 5-FDXulP 42	0.15	0.25	0.06	0.08	0
Diastereomeric Ratio (5-FDXulP 42/5-FDRhuP 41 :5-FDRulP 39)	2.1:1.0	1.2:1	1:5	2.0:1.0	--

Table 3.4. The products of the incubation of SCO1844 with FAld **40** with DHAP at different temperatures and in the presence of $ZnSO_4$ (10 µM) and EDTA (1 mM) in comparison with a control experiment. Figures are a value of the integrated area under signal peaks of the products, against fluoroethanol as a standard.

3.4.2.1 The aldol reaction by SCO1844 at different temperatures

Incubation temperature shown to effect the stereointegrity of L-FucAs in the aldol direction.[157 158] To this effect, three different aldol reactions with SCO1844, FAld **40** and DHAP were set up identically as before. They were then incubated separately, one at 37 ºC for 6 h, one at room temperature for 24 h and the last at 4 ºC for 24 h. The samples were terminated by heat deactivation (95 °C, 5 min), and the reaction solution was centrifuged (12,000 rpm, 2 min). The subsequent supernatant was removed and the volume made up to 700 µl with ultrapure water, before adding D_2O (100 µl). The sample was then subjected to ^{19}F NMR, and a typical spectrum is shown in Figure 3.11 and the subsequent product integrals are tabulated in Table 3.4.

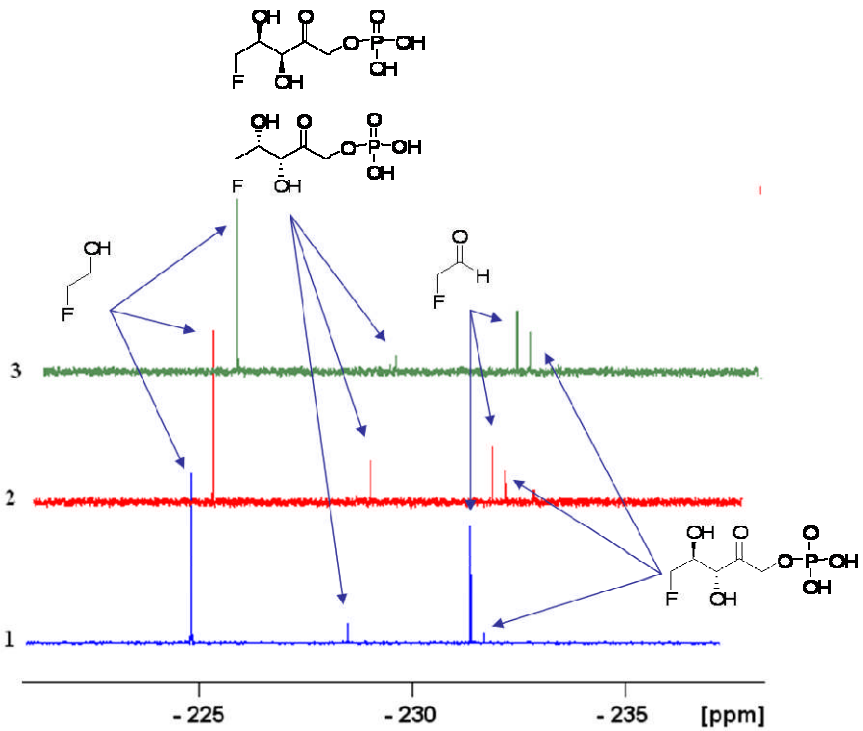


Figure 3.11. $^{19}F\{^1H\}$ NMR of the aldol addition of FAld **40** and DHAP at different incubation temperatures. 1= 37 ºC for 6 h, 2= room temperature (RT) for 24 h and 3= 4 ºC for 24 h.

The diastereoisomeric product ratio of SCO1844 in the aldol addition of FAld and DHAP is profoundly affected by the temperature of that incubation. At 37 ºC the diastereisomeric ratio is more than 2:1 with the major product being the "wrong" diastereoisomer, 5-FDRhuP **42** (or 5-FDXulP **41**). At room temperature, in a longer incubation, this ratio is nearer 1:1, but the major product is again the alternative configuration. However, upon incubation at 4 ºC, the diastereisomeric ratio is 5:1 with 5-FDRulP **39** as the major product. This effect is consistent with knowledge about the effect of temperature on L-FucAs, that reducing the temperature significantly alters improves the diastereoisomeric ratio of the product towards *3R, 4R*, the expected product of fuculose aldolases. [158]

3.4.2.2 Retro-aldol assay for SCO1844

The ability of the SCO1844 protein product to catalyse the aldol reaction between FAld **40** and DHAP to generate 5-FDRulP **39** has been demonstrated. SCO1844 was now assayed for its ability to catalyse the retro-aldol reaction, generating FAld **40** and DHAP from 5-FDRulP **39**. There are no commercial or synthetic routes to 5-FDRulP **39** available, and to date the only way to generate this compound for such as assay is enzymatically from SAM **34** and fluoride ion (see Scheme 3.5). The fluorinase and PNP genes from *S. cattleya* have been reported, and they have each been successfully over expressed in *E. coli*. The availability of an isomerase capable of converting 5-FDRP **38** to 5-FDRulP **39** (described in Chapter 2) allows us to reconstitute the fluorometabolite

pathway *in vitro* from SAM **34** and fluoride ion in the presence of these three enzymes to generate 5-FDRulP **39**.

The putative retro-aldol reaction of a fuculose aldolase will result in the generation of FAd **40** and DHAP. The assay for the SCO1844 in this case involves incubation of the SCO1844 protein with the fluorinase, PNP and isomerase in the presence of SAM **34** and fluoride ion. The reaction was followed by ^{19}F NMR. The fluorinase, PNP and isomerase genes were inserted into *E. coli* expression vectors pET28(b) and pHISTEV and transfected into *E coli* (BL21 Gold) competent cells. Expression and purification of these proteins was carried out similarly to that for SCO1844. The enzymes were each purified to final concentrations of ~1 mg/ml in phosphate buffer (10 mM, pH 7.6). Equimolar amounts (0.1 mg) of these proteins were then incubated together in the presence of SAM **34** (2 mM) and fluoride ion (50 mM) for 16 h at 37ºC. A control experiment was set up with SCO1844 excluded from the reaction, and the experiments were run simultaneously. The reactions were stopped by heat deactivation (95ºC, 5 min) and centrifuged (12,000 rpm, 2 min). The resulting supernatant was made up to a volume of 700 µl, and D_2O (100 µl) was added and the product mixtures were then analysed by ^{19}F NMR. A typical ^{19}F NMR product spectra of a reaction and a control are shown in Figure 3.12.

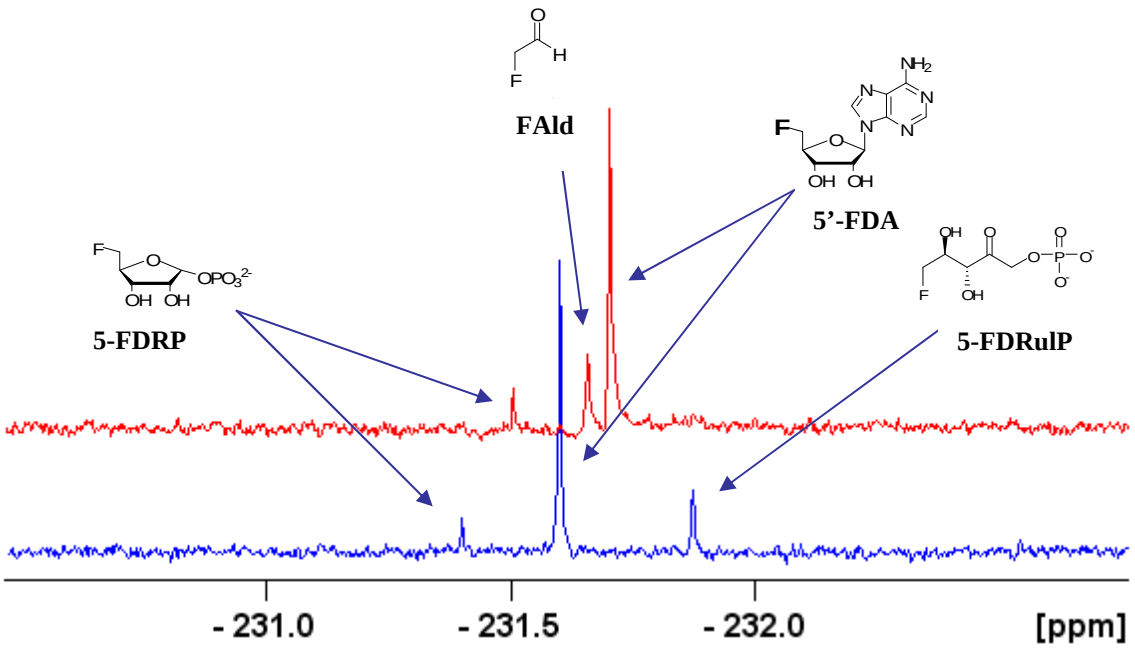


Figure 3.12. ^{19}F $\{^{1}H\}$ NMR spectra of the aldol reaction of SCO1844 in an *in vitro* fluorometabolite pathway experiment starting from SAM **34** and fluoride ion. The blue spectrum represents a control experiment, without the SCO1844 protein. The red spectrum represents an experiment with the SCO1844 protein incubated alongside the fluorinase, PNP and isomerase enzymes from *S. cattleya*.

Figure 3.12 clearly shows the generation of FAld **40** from 5-FDRulP **39**, in the presence of the SCO1844 protein product. In the control experiment where the SCO1844 protein was excluded from the reaction, then no FAld **40** is generated and 5-FDRulP accumulates. Clearly SCO1844 is responsible for that transformation.

3.4.2.3 Inhibition of SCO1844 by EDTA

The catalytic activity of fuculose aldolases is dependant upon the presence of Zn^{2+} at the active site (Scheme 3.3). They are therefore inhibited by the presence of EDTA, which sequesters the active site Zn^{2+}, inhibiting catalysis. The response of the SCO1844 protein

to the presence of EDTA would give a further insight into its activity, as it has only previously been annotated as a putative fuculose aldolase. Accordingly SCO1844 (0.1 mg) in phosphate buffer (10 mM, pH 7.6) was incubated with synthetic FAld **40** (1 mM) and DHAP (1 mM) in the presence of EDTA (1 mM) (6 hours, 37 °C). The reaction was stopped by heat deactivation (95 °C, 5 min), followed by centrifugation (12,000 rpm, 2 min). The resulting supernatant was removed and the volume made up to 700 µl using ultrapure water. The mixture was then added to D_2O (100 µl) and subject to ^{19}F NMR. An example of the resulting spectrum is shown in Figure 3.13 and the corresponding product profiles are tabulated in Table 3.4.

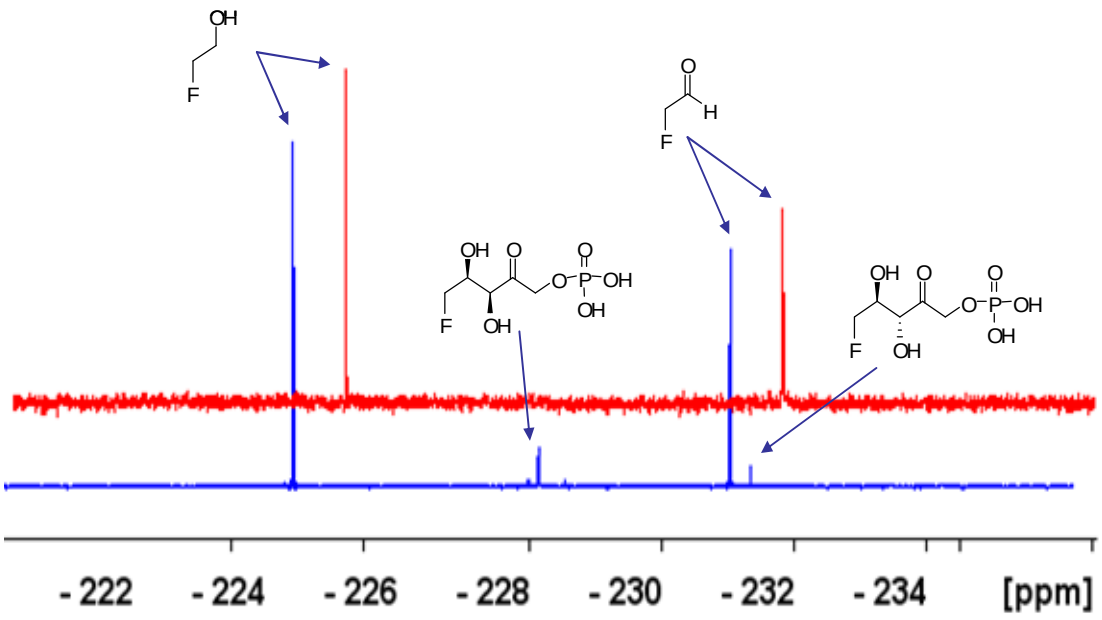


Figure 3.13. $^{19}F\{^{1}H\}$NMR spectrum of SCO1844 incubated with FAld **40** (1 mM) and DHAP (1 mM) for 6 h at 37 ºC. The red spectra is SCO1844 incubated with EDTA (1 mM). The blue spectra is a control experiment. Product profiles are tabulated in Table 3.4.

Incubation of SCO1844 in the presence of 1 mM EDTA apparently abolished any observable catalytic activity, arresting the production of both the 5-FDRulP **39** and 5-FDXulP **42**/ 5-FDRhuP **41** (^{19}F NMR, Figure 3.13). This is consistent with the idea that catalytic activity of SCO1844 is dependant upon a divalent ion cofactor, most likely in the form of Zn^{2+}. These results are also consistent with those expected for Class II aldolases, and are further evidence that SCO1844 is an enzyme of this type.

Previous studies involving the over-expression and purification of the L-FucA aldolase from *E. coli* have revealed that Ni^{2+} affinity chromatography has an inhibitory effect on the enzyme. Crystal structure studies of fuculose aldolase from *E. coli* have revealed that a Zn^{2+} ion is coordinated to three histidine residues (Scheme 3.3).[147] Inhibition probably occurs as the Ni^{2+} interacts with the active site histidines, stripping the enzyme of Zn^{2+}. There are many reports of the restoration of such activity after incubation of the enzyme with up to 10mM, $ZnSO_4$ solution after Ni^{2+} purification.[159, 160] However, upon elution of SCO1844 from Ni^{2+} affinity chromatography there was no apparent inhibition or loss of activity exhibited (Figure 3.13).

3.4.2.4 Inhibition of SCO1844 by Zn^{2+}

L-FucA is also reported to be inhibited by Zn^{2+} in solution even at low concentration. That is, too much Zn^{2+} has a negative effect on catalytic activity. It is therefore necessary to dialyse the protein into a Zn^{2+}-free buffer in order to restore activity.[161] In the light of this it was important to establish a role for Zn^{2+} in the activity of the SCO1844 aldolase. Accordingly the SCO1844 enzymes was taken up (0.1 mg) in phosphate buffer (10 mM pH 7.6) and incubated with DHAP (1 mM) and FAld **40** (1 mM) with and without the

addition of $ZnSO_4$ (10 µM) for 6 hours at 37 ºC simultaneously. The reactions were then stopped by heat deactivation (95 ºC, 5 min) and centrifugation (2 min, 12,000 rpm). The volume of the resulting supernatant was made up to 700 µl with ultrapure water. D_2O (100 µl) was then added and the resulting mixture was subject to ^{19}F NMR analysis. A typical spectrum is shown in Figure 3.14 and product distributions are elaborated in Table 3.4.

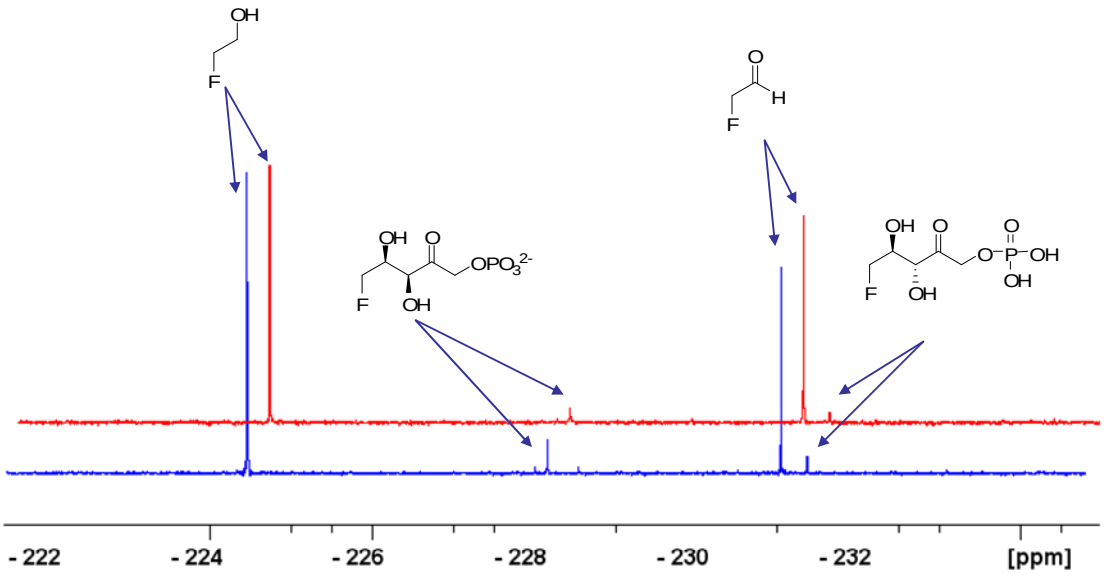


Figure 3.14. $^{19}F\{^1H\}$ NMR of the incubation of SCO1844 with FAld **40** (1 mM) and DHAP (1mM) for 6 hours at 37°C. The blue spectrum represents the control. The red spectrum has $ZnSO_4$ (10 µM) added.

There is some evidence for the inhibitory effect of Zn^{2+} as shown in Table 3.4. Integration of the signals assigned to 5-FDRulP **39** and 5-FDXulP **41**/ 5-FDRhuP **42** reveals that the presence of just 10 µM $ZnSO_4$ in solution reduces the activity of SCO1844 by about a half compared to a buffer without any added Zn^{2+}. The same experiment was attempted with 50 and 100 µM $ZnSO_4$ in solution. Both of these experiments resulted in the immediate precipitation of the SCO1844 protein, and

complete abolition of activity. It is noteworthy that the inhibition by Zn^{2+} does not significantly affect the diastereoisomeric ratio of the products of the reaction.

3.5 Conclusions

The putative fuculose aldolase gene SCO1844 from *S. coelicolor* was identified as a surrogate aldolase for the purpose of reconstituting the fluorometabolite pathway *in vitro*. The SCO1844 gene was amplified from genomic DNA using PCR and inserted into an *E. coli* expression system, from which the protein was efficiently overexpressed and purified. The active SCO1844 protein is a dimer, susceptible to EDTA inactivation. It is apparently inhibited by Zn^{2+} in solution, consistent with our knowledge of Class II fuculose aldolases. Incubation of the SCO1844 protein with its donor substrate DHAP and a FAld **40** resulted in the production of 5-FDRulP **39** as well a diastereoisomer, 5-FDXulP **42** or 5-FDRhuP **41**. A lack of stereospecificity was exhibited by the enzyme in the aldol direction. The protein also efficiently catalysed the reverse reaction, generating FAld **40** from 5-FDRulP **39** in an *in vitro* reconstitution experiment.

Attempts to isolate the gene for L-FucA, putatively involved in fluorometabolite biosynthesis, from *S. cattleya* genomic DNA were unfortunately unsuccessful. The designing of degenerate primers in a similar manner to the method employed in identifying an isomerase from *S. cattleya* was unable to detect the L-FucA gene. Identification of this gene will become a focus after sequencing of the *S. cattleya* genome, which is currently underway.

4 The *in vitro* reconstitution of fluoroacetate and 4-fluorothreonine biosynthesis

The availability of gene clusters, responsible for natural product assembly is opening up the possibility of total natural product synthesis by biotransformation. This is emerging as a new and alternative approach to organic synthesis. An added benefit of using this approach is that reaction conditions are mild, and the enzyme catalysed reactions are extremely stereoselective and can give rise to stereochemical complexity. The *in vitro* reconstitution of natural product pathways has already been used to generate clinically important polyketides such as enterocin.[162] It has also been used in the characterization of the metabolic pathways of halogenated natural products, identifying the necessary co-factors and steps involved in halogenation.[40-42] Following the identification of the fluorinase from *S. cattleya*, work within the research group has been concentrated on the identification of the intermediates and the subsequent enzymatic steps in the biosynthetic pathway of FAc **8** and 4-FT **33**. Recently, the enzymatic synthesis of radiolabelled [^{18}F] molecules using the fluorinase as the C-F bond catalyst was achieved for PET analysis.[112, 113] Reconstitution of the fluorometabolite *in vitro* pathway would also open up the possibility of generating novel [^{18}F]-labelled compounds for this technique and thus became a research focus.

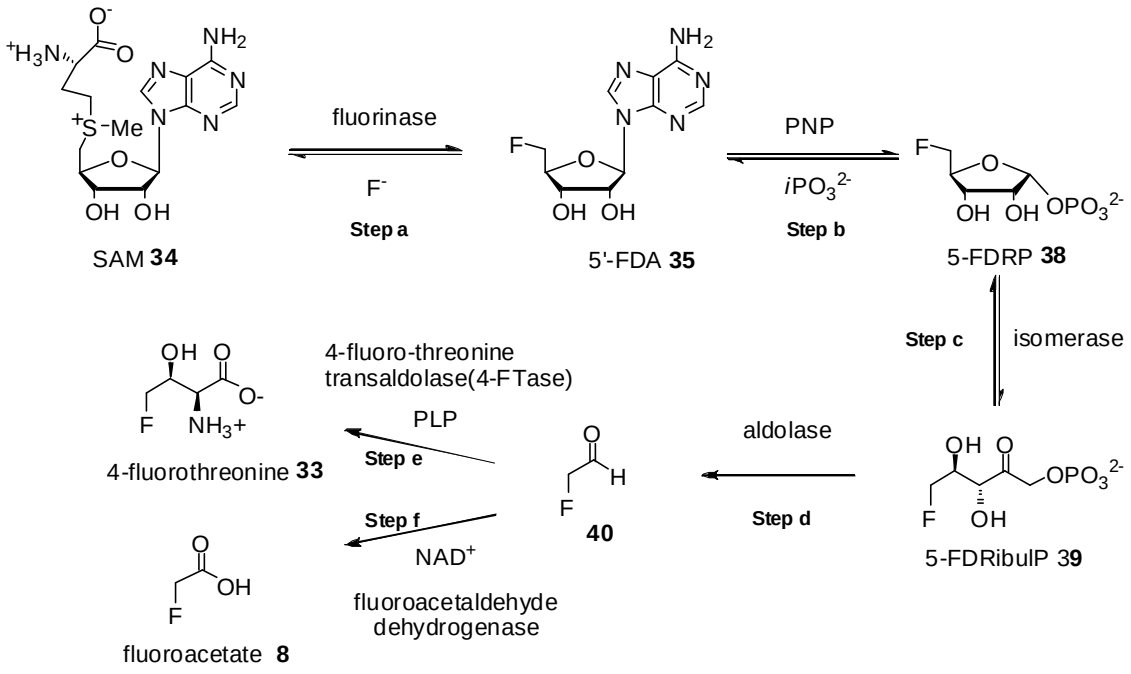
The identification of an isomerase from *S. cattleya* capable of converting 5-FDRP **38** to 5-FDRulP **39** (Step c, Scheme 4.1) and a fuculose aldolase from *S. coelicolor* capable of generating FAld **40** from 5-FDRulP **39** (Step d, Scheme 4.1) has been described in Chapters 2 and 3 respectively. Aldehyde dehydrogenases are commercially available (Sigma Ltd, UK) (Step f, Scheme 4.1) and with the successful over expression of the PLP transaldolase (Step e, Scheme 4.1) prospects of reconstituting the entire fluorometabolite pathway, *in vitro* emerged.

4.1 The fluorometabolite pathway in *S. cattleya*

The 10kb gene cluster was identified by J. Spencer at Cambridge in 2005, was described in Chapter 1.[83] Adjacent to fluorinase were a number of genes which appear to be involved in the biosynthesis and regulation of the fluorometabolite pathway (Figure 1.5). The fluorinase *flA* gene is in the middle of the cluster and the immediate upstream gene *flB*, has been shown to express a purine nucleotide phosphorylase (PNP) enzyme which is selective for the conversion of 5'-FDA **5** to 5-FDRP **6**[83, 95] (step b, Scheme 4.1). Several regulatory and resistance genes were also identified although their exact roles are unclear. However, the genes for the remaining three biosynthetic enzymes (enzyme steps **c-e**, isomerase, aldolase and PLP transaldolase, Scheme 4.1) are not in the gene cluster, and thus only the *flA* (fluorinase) and *flB* (PNP) gene products, are available by PCR and over expression from this cluster sequence.

With the isomerase from *S. cattleya* and the surrogate fuculose aldolase from *S. coelicolor* in hand then an enzyme for all steps for the synthesis of FAc **8** and 4-FT **33**

was available. The next phase of the research aimed to recombine these enzymes *in vitro* to affect a complete biotransformation of these fluorometabolites from fluoride ion. FAld **40** has been established as the last common intermediate feeding into each of the fluorometabolites **8** and **33**. [153] An $NAD(P)^+$ dependent fluoroacetaldehyde dehydrogenase has been purified[8] which oxidises FAld **40** to fluoroacetate (FAc) **8**. Separately the pyridoxal phosphate (PLP) dependant enzyme described in Chapter 1 was available, which mediates a trans-aldol reaction between L-threonine and FAld **40** to generate 4-FT **33**.[99]



Scheme 4.1. Current status of fluorometabolite biosynthesis in *S. cattleya* showing metabolic intermediates and the enzyme **steps a-f** and co-factors involved.

4.2 *In vitro* reconstitution of fluorometabolite biosynthesis

The biosynthetic pathway from fluoride ion and SAM **34** to FAc **8** in *S. cattleya* requires five enzymes (Scheme 4.1). The *flA* and *flB* genes from the 10kb gene cluster of *S. cattleya* code for the fluorinase and the PNP enzymes which catalyse the first two steps in fluorometabolite biosynthesis.[83] The efficient over-expression of the fluorinase enzyme has previously been described[82] and the enzyme is readily available and stable. Attempts at over-expression of the PNP enzyme in *E. coli* by PCR amplification of the *flB* gene from genomic DNA (*S. cattleya*), were only partially successful. Although the protein could be expressed successfully it was largely insoluble[83, 95] and it proved difficult to obtain sufficiently soluble protein for biotransformation assays. To get around this problem, the *flB* gene was fused to a modified pMAL vector, pLOU, coding for a maltose binding protein.[103] This PNP was used in subsequent biotransformations.

The expression and purification of the isomerase from *S. cattleya* and the fuculose aldolase from *S. coelicolor* have been described previously in Chapters 2 and 3 respectively. The fluorinase, PNP, isomerase and fuculose aldolase were all over-expressed in *E. coli* and purified by Ni-affinity and size exclusion chromatography to ~1 mg/ml in phosphate buffer (PBSA). The identities of these over-expressed enzymes was confirmed by SDS-PAGE and MS-MS analysis. The resulting SDS-PAGE gel is shown in Figure 4.1.

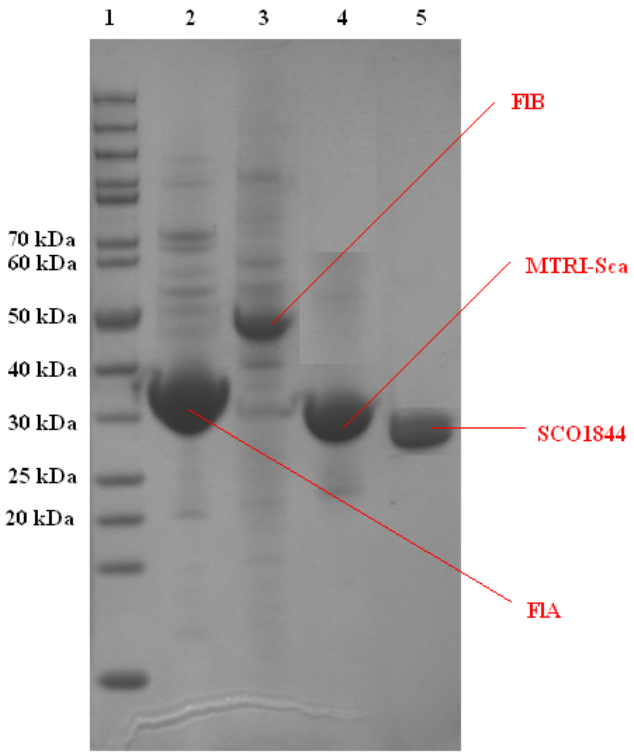


Figure 4.1. SDS-PAGE of enzymes, used in reconstitution experiments. 1. Protein molecular weight markers (Fermentas); 2.The fluorinase (*FlA*); 3. The PNP (*FlB*); 4. MTRI (*MTRI-Sca*); 5. Fuculose aldolase (*SCO1844*). The identity of the proteins was confirmed by nanoLC-ESI MSMS (UltiMate (Dionex) and Q-Star Pulsar XL (Applied Biosystems)of a tryptic digestion of the partially purified protein.

4.2.1 *In vitro* FAc 8 biosynthesis

Aldehyde dehydrogenase from *S. cerevisiae* (Sigma Ltd) was suspended in phosphate buffer (10 mM, pH 7.8). A solution of $NAD(P)^+$ (Sigma Ltd) was adjusted to a final concentration of 20 mM. With all of the enzymes and co-factors in hand, it was possible to combine them and to follow the production of FAc **8** by $^{19}F\{^1H\}$ NMR. All of the pathway enzymes were added into an eppendorf tube (1.5 ml) to a final concentration of 0.1 mg/ml. They were incubated with SAM **34** (1.4 mM, Sigma Ltd), KF (35 mM, Sigma Ltd), $NAD(P)^+$ (1 mM, Sigma Ltd) for 6 h at 37 °C. During the reaction, aliquots (100 µl) were removed at 0, 1, 2, 3, 4, 6 and 24 h. The reactions were stopped by heat deactivation

at (95 °C, 5 min) followed by centrifugation of (2 min,12,000 rpm). The supernatant was then made up to a volume of 700 µl, to which D_2O (100 µl) was also added. The resulting mixture was then subjected to $^{19}F\{^1H\}$ NMR analysis (Figure 4.2).

Figure 4.2 shows that after 2 h, two organofluorine signals with chemical shifts of -231.55 ppm and -217.35 ppm emerged. The major signal at -231.55 ppm was confirmed as 5'-FDA **35** by add-mixing with a synthetic standard. The minor signal at -217.3 ppm was confirmed as FAc **8** by add-mixing with a reference compound (Sigma Ltd). The product ratios from the spectra are tabulated in Table 4.1, calculated by integration of the signals.

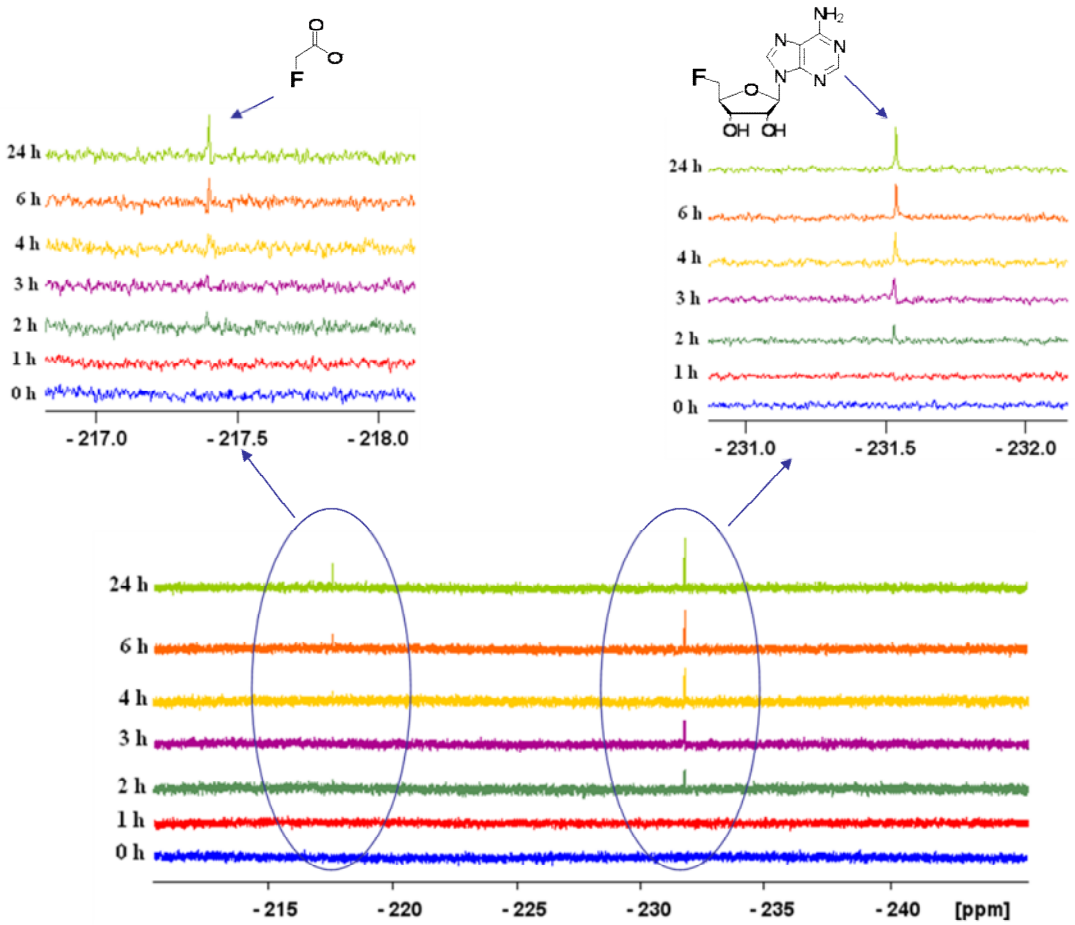


Figure 4.2. $^{19}F\{^{1}H\}$ NMR spectra of the time course experiments from the reconstitution of the FAc **8** biosynthesis pathway of *S. cattleya in vitro*. The reaction mixture was assayed at 0, 1, 2, 3, 4, 6 and 24 h.

Time (h)	FDA 35:FAc 8
0	0
1	0
2	5.5:1
3	4:1
4	2.6:1
6	2.7:1
24	2.2:1

Table 4.1. The ratio of the products from the *in vitro* FAc **8** biosynthesis with time.

The results from Figure 4.2 and Table 4.1 show that reconstitution of the FAc **8** biosynthetic pathway has been successful. With each enzyme in equimolar concentrations, only two fluorinated products are observed by ^{19}F NMR, corresponding to 5'-FDA **35** and FAc **8**. The accumulation of 5'-FDA **35** as the major product in all of the time course experiments reveals that the PNP (Step b, Scheme 4.1) is a bottle neck in this reaction. After 3-4 h of reaction, the fluorinase (Step a, Scheme 4.1) has reached an equilibrium as evidenced by the accumulation of 5'-FDA **35**. After 3 h, the ratio of 5'-FDA **35**: FAc **8** begins to decrease as more 5'-FDA **35** is converted to FAc **8**.

There is no evidence from ^{19}F NMR of the accumulation of any other fluorometabolite intermediates in the reaction medium. In these reactions, any 5'-FDA **35** converted to 5-FDRP **38** by the PNP enzyme is immediately converted to 5-FDRulP **39** by the isomerase, which in turn instantly undergoes the retro-aldol reaction catalysed by the fuculose aldolase to produce FAld **40**. The advent of FAld **40** production, in the presence of the aldehyde dehydrogenase from *S. cerevisiae* and high concentrations of its co-factor, $NAD(P)^+$, drives the equilibrium from FAld **40** to FAc **8** in an irreversible reaction. After 24 h the product ratio of 5'-FDA **35**:FAc **8** has leveled out at around 2:1.

4.2.2 *In vitro* reconstitution of the 4-FT pathway

The enzyme preparations (Steps a-d, Scheme 4.1) involved in the first four steps of the *in vitro* biosynthesis of FAc **8** were also used in attempts to generate 4-FT **33** starting from SAM **34** and fluoride ion. The recent identification, over-expression and purification of the PLP-dependant transaldolase (FTase), responsible for the generation of 4-FT **33** from

FAld **40** and L-threonine from *S. cattleya* has previously been described. The FTase gene was then subcloned into the pXY2000 E coli-*Streptomyces* shuttle vector with restriction sites of NdeI and EcoRI. The resultant plasmid pXY-ScaFTase was transfected into protoplasts of *S lividans* TK24[180] (by Dr Hai Deng, University of St Andrews[99]). The *FTase* gene was introduced into the *Streptomyces lividans* TK24 strain *via* the *E. coli* - *Streptomyces* shuttle vector pXY2000.[163] Apramycin resistant clones were grown in the YEME medium and protein over-expression was induced by the addition of thiostrepton (10μg/mL).[163] The enzyme was then partially purified by affinity column chromatography. Enzyme expression in *S. lividans* was low, however expression was confirmed by MS-MS sequencing of the partially purified protein (Figure 4.3).

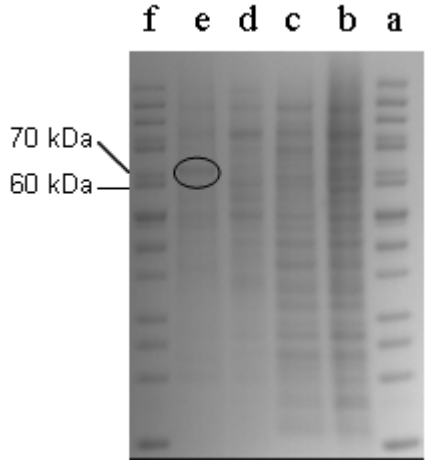


Figure 4.3. SDS-Page of the progressive purification of 4-FTase over-expressed in *S. lividans*. a. Protein molecular weight markers (Fermentas). b. Cell-free extract. c. Cell-free extract supernatant; d. Cell-free extract precipitant. e. Eluent after Ni-affinity column. f. Protein molecular weight markers (Fermentas). The identity of the protein was confirmed by nanoLC-ESI MSMS.

The over-expression of *Streptomyces* genes eg. in *E. coli* often leads to insoluble protein (inclusion bodies). Proteins that exhibit this are often expressed in *Streptomyces* hosts, and even then is expression poor. The FTase is typical of this, however the over

expressed protein (Figure 4.3) was suspended in phosphate buffer (PBSA) and concentrated to ~0.25 mg/ml. This preparation was subsequently used in the bio-transformations.

4.2.2.1 *In vitro* 4-FT biosynthesis

All of the pathway enzymes were added into an eppendorf tube (1.5ml) to a final concentration of 0.1 mg/ml. They were incubated with SAM **34** (1.4 mM, Sigma), KF (35 mM, Sigma), PLP (0.7 mM, Sigma) and L-threonine (35 mM, Sigma) for 16 h at 37 °C. The reaction was stopped by heat inactivation (95 °C, 5 min) followed by centrifugation (2 min at 12,000 rpm). D_2O (100 µl) was then added to the supernatant and subject to $^{19}F\{^1H\}$ NMR analysis. A typical $^{19}F\{^1H\}$ NMR spectra is shown in Figure 4.4.

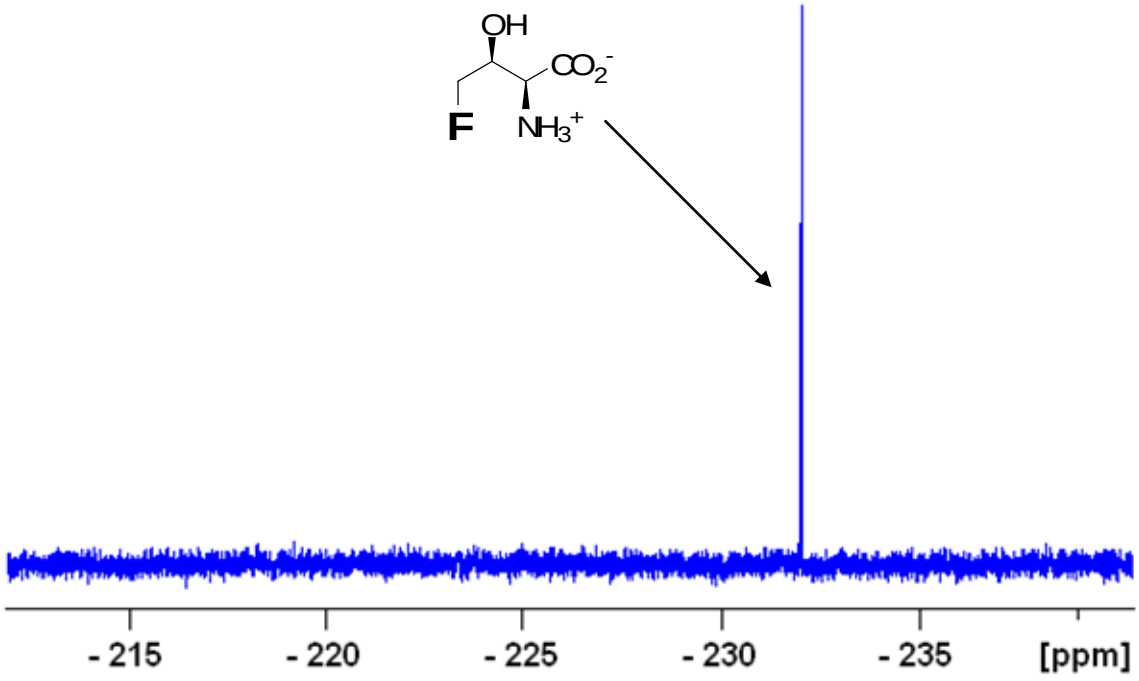


Figure 4.4. $^{19}F\{^1H\}$ NMR spectra of 4-FT **33** generated *in vitro*.

The $^{19}F\{^{1}H\}$ NMR spectra revealed that the combination of all of the constituent enzymes results in an efficient conversion with a single organo-fluorine product (-232.0 ppm). This signal corresponds to the production of 4-FT **33**. The identity was confirmed by add-mixing with a synthetic sample of 4-FT **33**. Unlike the reconstituted FAc **8** pathway, there are no other fluorometabolites detected by $^{19}F\{^{1}H\}$ NMR analysis from this experiment. This suggests that the FTase reaction pulls the equilibrium in the direction of 4-FT **33** synthesis. The absence of other fluorinated intermediates suggests that the reaction was still occurring after 16 h incubation. An equilibrium for 4-FT **33** generation by the FTase was not reached, this was most likely because of the introduction of high concentrations of L-threonine (35 mM) in the reaction medium.

To further confirm the identity of 4-FT **33** the product solution was subjected to ^{19}F-NMR (500MHz) but without $\{^{1}H\}$-decoupling and the coupled signal compared to that of a synthetic sample of 4-FT **33**. The resulting spectra is shown in Figure 4.5. This revealed that the organofluorine product had an identical ^{19}F-NMR signal at -232.0 ppm with a characteristic multiplicity (d.t $^{2}J_{HF}$ 46.9 Hz, $^{3}J_{HF}$ 25.0 Hz) to that of the synthetic standard of 4-FT **33**.

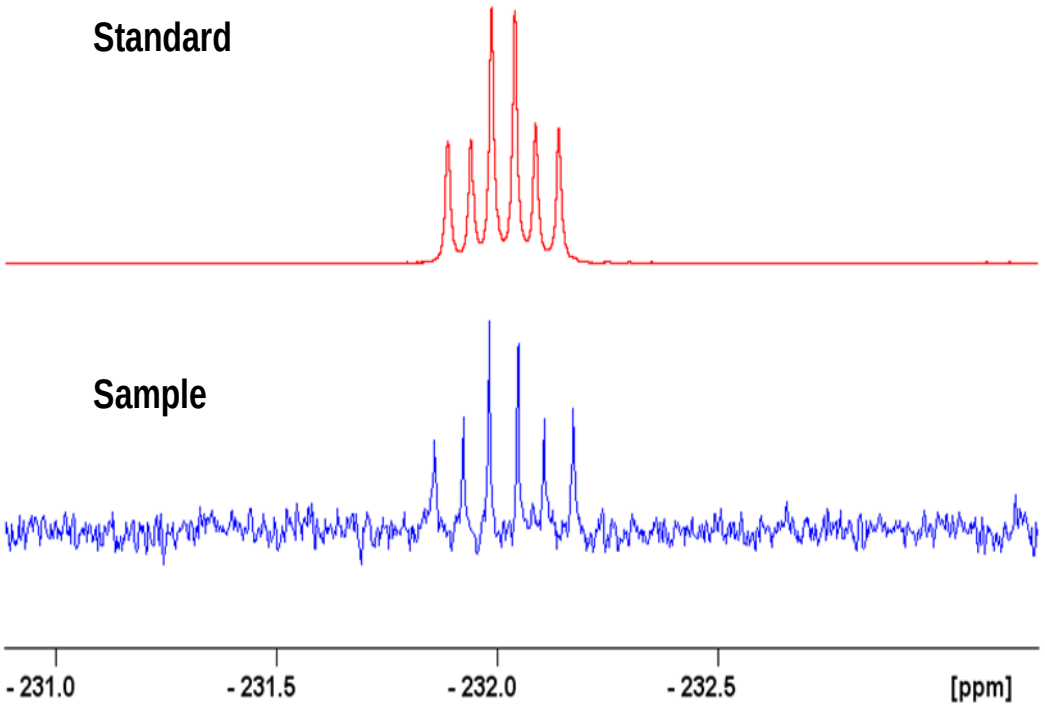


Figure 4.5. ^{19}F NMR of 4-FT **33** generated by the *in vitro* reconstitution of the fluorometabolite pathway compared with a synthetic standard of 4-FT **33** *(dt,* δ_F -232.0, $^2J_{HF}$ 46.9 Hz, $^3J_{HF}$ 25.0 Hz).

In order to confirm unambiguously the generation of 4-FT **33**, a sample of the reaction mixture was subject to GC-MS analysis (Dr J.T.G Hamilton, Queens University, Belfast). The sample was lyophilized and then treated with N-methyl-N-(trimethylsilyl) trifluoroacetamide (60 min, 100°C). This treatment per-trimethylsyilylates the amino acid and then GC-MS analysis was carried out on a silica capillary column. The resulting total ion chromatogram (TIC) and mass spectra are shown in Figures 4.6 and 4.7 respectively.

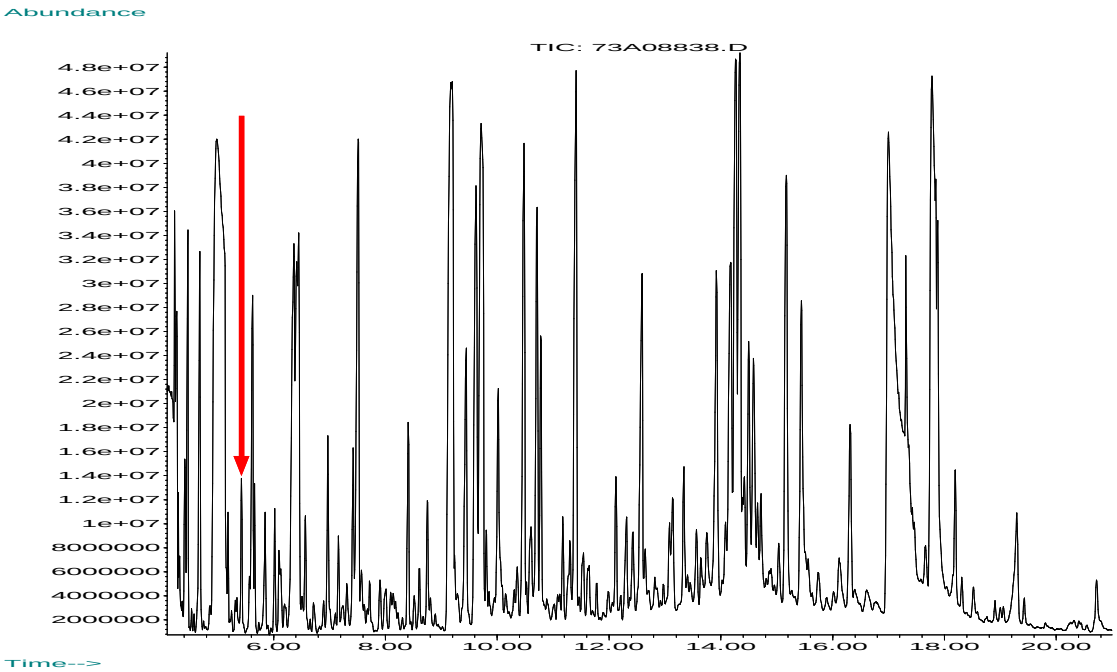


Figure 4.6. GC-MS total ion chromatogram of the persilylated 4-FT **33**. The red arrow indicates the presence of 4-FT **33** determined by comparison with a synthetic 4-FT **33** standard.

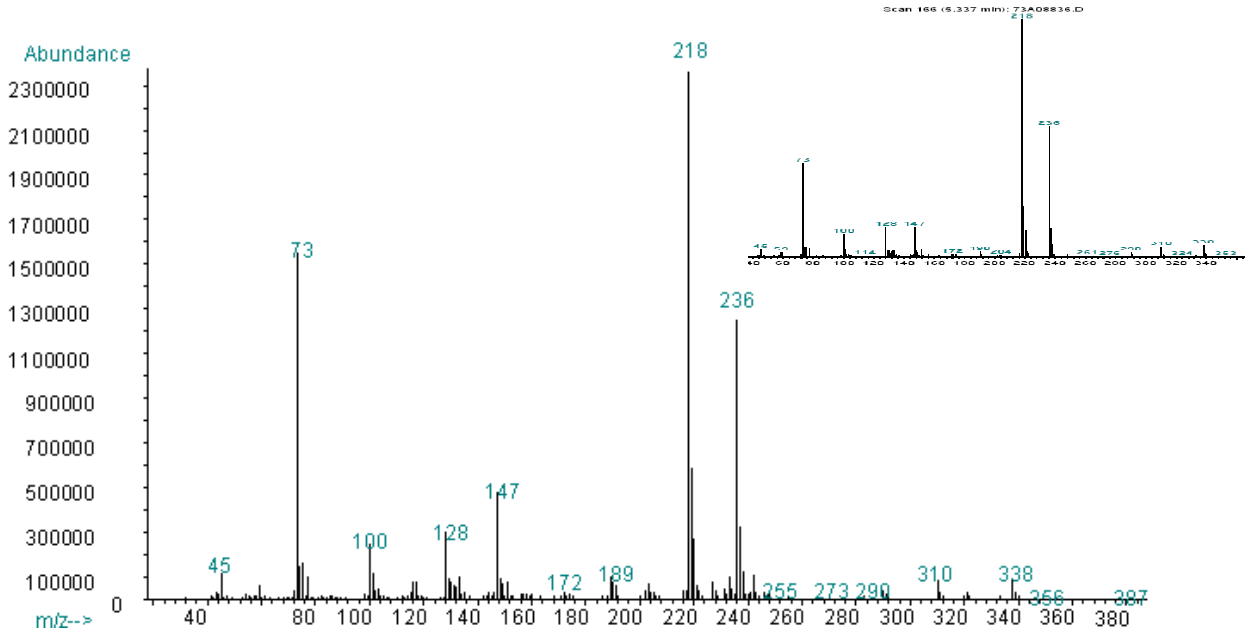


Figure 4.7. GC-MS mass spectra of the persilylated 4-FT **33**. Two mass ions are prevalent, 218 amu and 236 amu. Inset is the spectra generated from a 4-FT **33** standard.

The GC-MS mass spectrum (Figure 4.7) of the persilylated derivative of 4-FT **33** revealed predominant mass ions at 218 and 236 amu respectively. It has previously been reported that after such derivatisation, 4-FT **33** undergoes cleavage to form two different mass fragments on GC-MS analysis (Figure 4.9). It is well established that the mass ion at 218 amu is an indicator for α- amino acids. The mass ion at 236 amu is indicative of the presence of 4-FT **33**.[103]

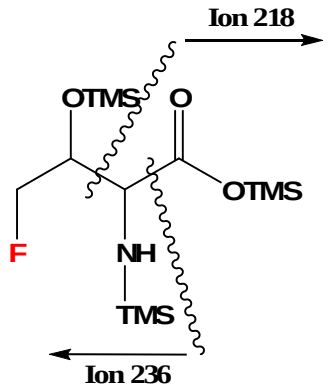


Figure 4.9. The established fragmentation of the 4-FT **33** molecule after derivitisation by persilylation, and the corresponding masses.[103]

The combination of ^{19}F NMR and GC-MS has unambiguously confirmed the generation of 4-FT **33** by the reconstitution of the fluorometabolite pathway from SAM **34** and fluoride ion in a one-pot reaction.

4.2.2.2 *In vitro* 4-FT 33 reconstitution control experiments

A series of experiments were now conducted where either all of the enzymes were combined, or for control reactions, one enzyme was omitted from the *in vitro*

recombination biotransformations. Identities of intermediates were re-confirmed by add-mixing reference compounds of 4-FT **33**, FAld **40**, 5'-FDA **35**, 5-FDRP **38**, 5-FDRulP **39** and 5'-FDI **36** into product solutions of the relevant experiments for further analysis by $^{19}F\{^{1}H\}$NMR.

4.2.2.2.1 SAM and fluoride ion omission

All of the enzymes and co-factors except SAM **34** were combined (F^-, PLP, L-threonine) for 16 h at 37 ºC. In another similar experiment, all of the enzymes and cofactors (in PBSA buffer) were combined except for fluoride ion and incubated similarly. The reactions were stopped by heat deactivation at (95 ºC, 5 min), and centrifuged (12,000 rpm, 2 min). The resulting supernatant was removed and D_2O added (100 µl) before being subject to $^{19}F\{^{1}H\}$ NMR analysis. As expected, there was no organo-fluorine production, confirming that both SAM **34** and fluoride ion are necessary for fluorometabolite production in these *in vitro* experiments.

4.2.2.2.2 Fluorinase omission

All of the enzymes and co-factors (SAM **34**, F^-, PLP, L- threonine) except the fluorinase (*FlA*) were combined and incubated for 16 h at 37 ºC. The reaction was stopped by heat deactivation at 95 ºC for 5 min, and centrifuged (12000 rpm, 2 min). The resulting supernatant was removed and added to D_2O (100 µl) before being subject to $^{19}F\{^{1}H\}$ NMR analysis. Omission of the fluorinase arrests organo-fluorine production, because the crucial C-F bond forming enzyme is removed from the reaction. These results show that

the fluorinase, and its associated substrates are critical for C-F bond formation and fluorometabolite production in the reconstitution experiments.

4.2.2.2.3 PNP omission

Removing the PNP enzyme from the bio-transformation was then explored. To this effect all of the enzymes and co-factors (SAM **34**, F^-, PLP, L-threonine), except the PNP (*FlB*), were combined and incubated for 16 h at 37 ºC. The reaction was stopped again by heat inactivation, and the resulting supernatant was analysed as before by $^{19}F\{^1H\}$ NMR analysis (Figure 4.10).

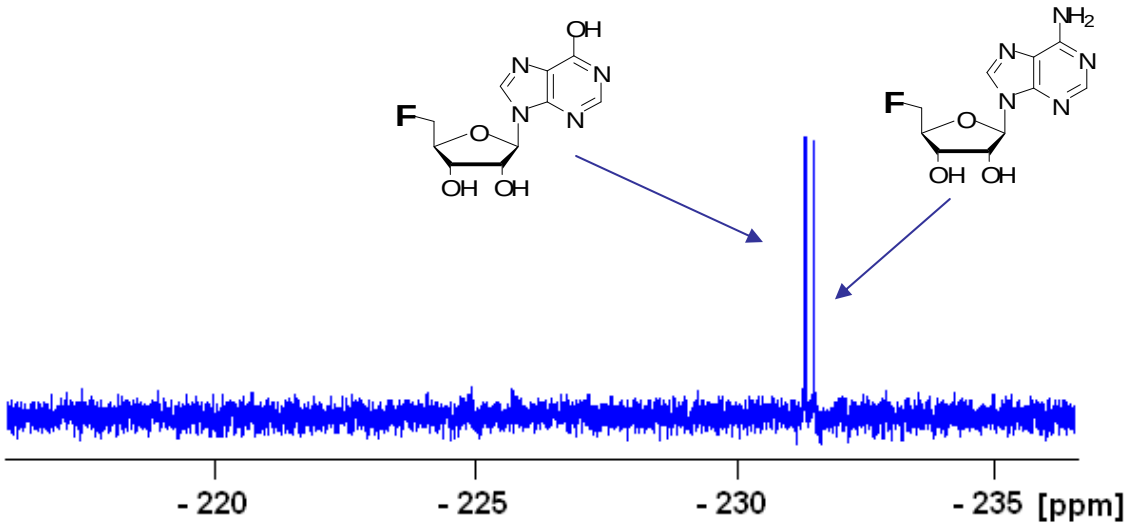


Figure 4.10. The $^{19}F\{^1H\}$ NMR spectrum of the products of the reconstituted 4-FT **33** reaction with PNP omitted from the reaction.

$^{19}F\{^1H\}$ NMR analysis of this reaction reveals two organofluorine signals, with chemical shifts of -231.36 and -231.52 ppm in a ratio of 0.88:1 respectively. The major signal at -231.52 ppm is consistent with the generation of 5'-FDA **35**, which is expected to

accumulate in this experiment. The minor signal at -231.36 corresponds to 5'-FDI **36**, a product which arises from enzymatic deamination of 5'-FDA **35**[93] by adenosine deaminase. This conversion has previously been identified in cell free extracts of *S. cattleya*, where it is particularly active. In this case it arises from a low level activity of the deaminase in the *S. lividans* FTase preparation (see 4.2.2.2.6).

4.2.2.2.4 Isomerase omission

Analysis of the reconstituted pathway, with the omission of the isomerase enzyme, was then explored similar to the experiments described above. All of the enzymes and co-factors (SAM **34**, F^-, PLP, L-threonine) except the isomerase (*MTRI-Sca*) were combined. The reaction was stopped after 16 h by heat deactivation, and the resultant $^{19}F\{^1H\}$ NMR spectrum is shown in Figure 4.11.

Figure 4.11 shows the $^{19}F\{^1H\}$ NMR spectrum indicating accumulation of three organofluorine peaks with the chemical shifts -231.32, -231.33 and -231.51 ppm and with a product ratio of 0.43:1:0.34 respectively. The major product of the reaction at -231.33 ppm was attributed to 5'-FDI **36**, which as in the previous experiment is generated by the enzymatic deamination of 5'-FDA **35** (-231.51 ppm). The reduction in the intensity of the signal for 5'-FDA **35** relative to that of 5'-FDI **36** is consistent with the latter being a shunt product,[93] and that 5'-FDA **35** is the substrate for the PNP enzyme. The third signal at -231.32 ppm corresponds to 5-FDRP **38**, the result of phosphorolytic cleavage of the adenosine base of 5'-FDA **35** by the PNP enzyme and the logical result of this incomplete biotransformation.

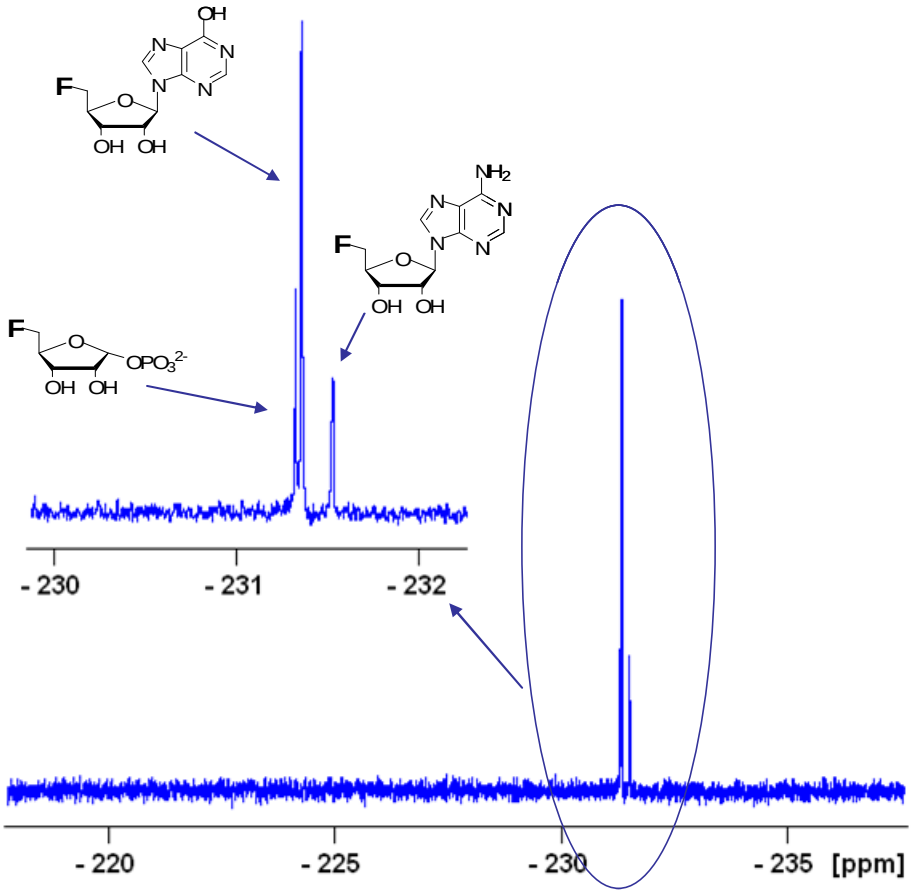


Figure 4.11. The $^{19}F\{^{1}H\}$ NMR spectrum of the 4-FT **33** bio-transformation with the isomerase omitted from the reaction.

4.2.2.2.5 Fuculose aldolase omission

The role of the surrogate fuculose aldolase from *S. coelicolor* in the reconstitution experiment was then investigated. In a similar vein, all of the enzymes and co-factors (SAM **34**, F^-, PLP, L-threonine) except the fuculose aldolase (*SCO1844*) were combined and incubated for 16 h at 37 ºC. The reaction was stopped and the resulting supernatant was analysed by $^{19}F\{^{1}H\}$ NMR. The spectrum is shown in Figure 4.12.

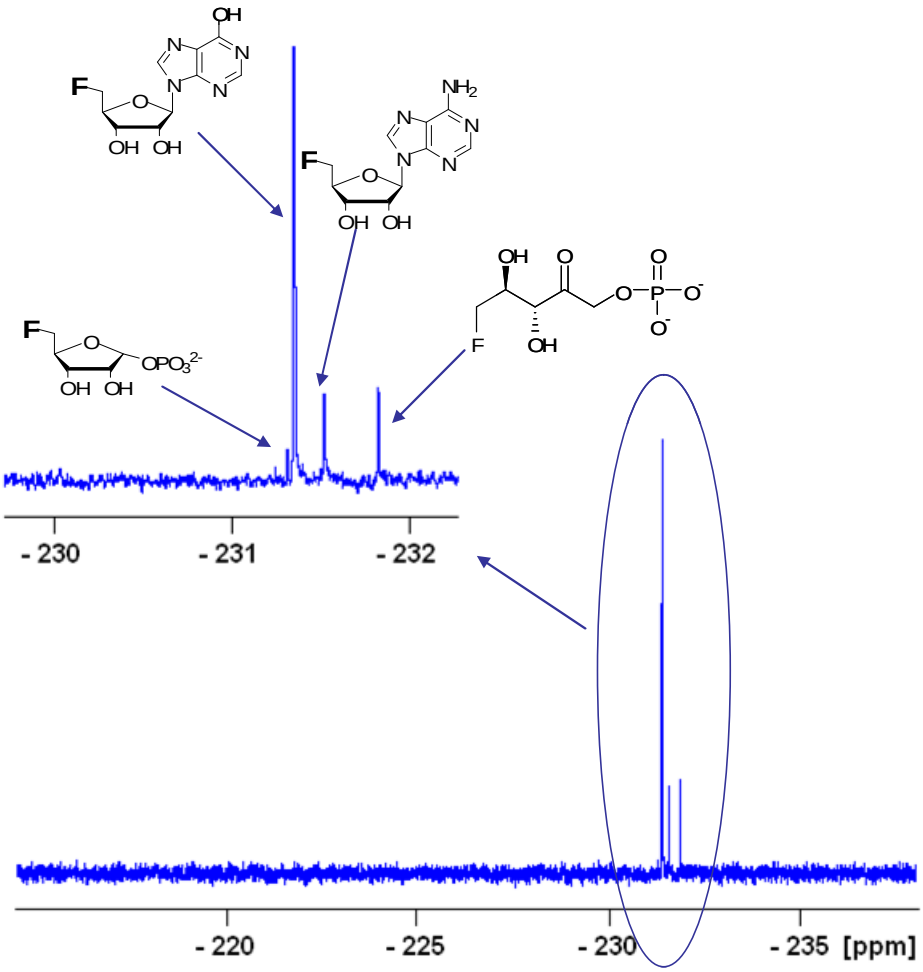


Figure 4.12. The $^{19}F\{^{1}H\}$ NMR spectrum of the 4-FT **33** bio-transformation with fuculose aldolase (*SCO1844*) omitted from the reaction.

The resultant $^{19}F\{^{1}H\}$ NMR spectra revealed the accumulation of four organofluorine signals, at -231.30, -231.34, -231.51 and -231.82 ppm and the product ratios are tabulated in Table 2. The minor signal at -231.30 ppm corresponds to 5-FDRP **38**, the product of the enzymatic phosphorylation of 5'-FDA **35** by the PNP (*FlB*) enzyme. The signal at -231.34 ppm is the major product of this reaction and corresponds to 5'-FDI **36**, which indicates the activity of an adenosine deaminase acting on 5'-FDA **35** (-231.51 ppm). The accumulation of 5'-FDA **35**, reveals that the PNP enzyme is once again rate limiting in

the reaction. The final signal, with a chemical shift of -231.82 ppm, corresponds to the production of 5-FDRulP **39**, the product of ring opening and isomerisation of 5-FDRP **38** catalysed by the isomerase (*MTRI-Sca*). The generation of 5-FDRulP **39** appears to be efficient as there is only a small amount of residual 5-FDRP **38** detected in the reaction mixture and, as the isomerase is a reversible reaction,[131] in this experiment all of the intermediates accumulate up until 5-FDRulP **39**.

4.2.2.2.6 PLP-dependant transaldolase omission

Finally the 4-FT **33** forming enzyme, FTase was left out of the reaction in order to determine its role in the pathway. The reaction was carried out similarly to those previously, but with PBS buffer used in the place of FTase. The reaction was again stopped by heat deactivation and D_2O (100 µl) was added to the resulting supernatant for $^{19}F\{^1H\}$ NMR analysis (Figure 4.13).

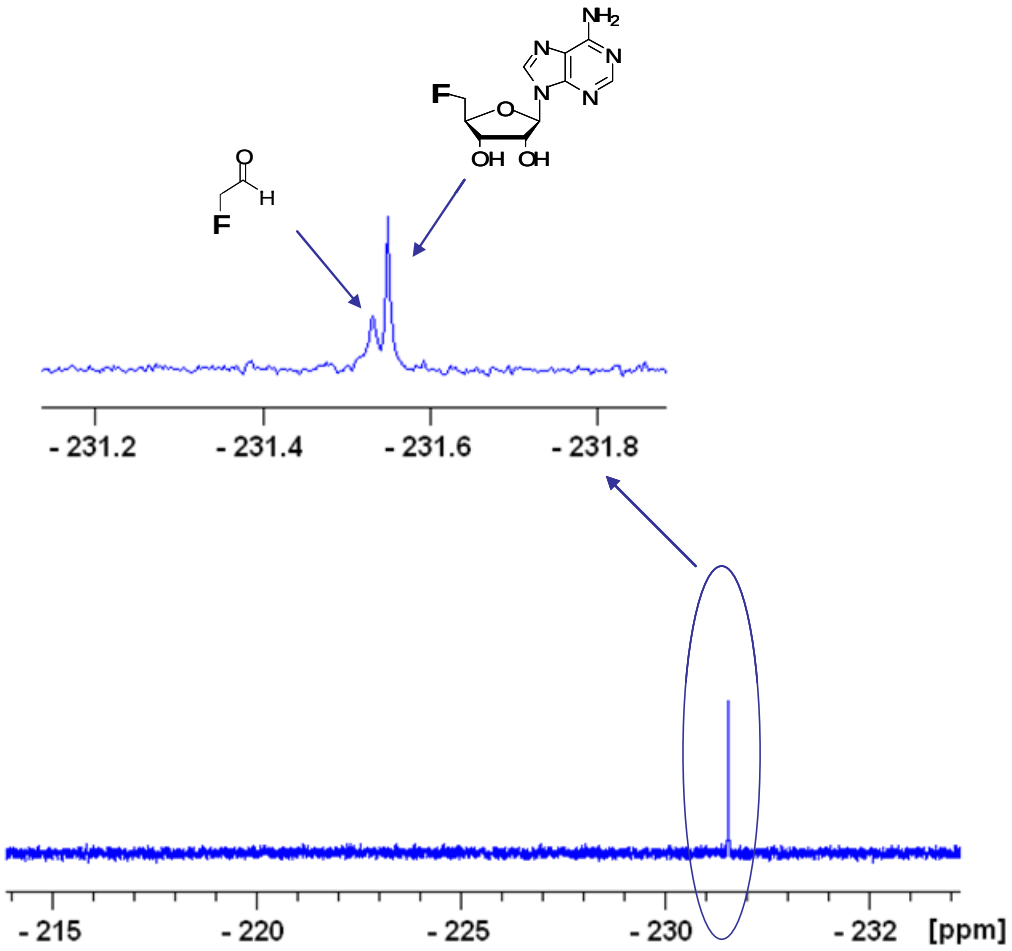


Figure 4.13. The $^{19}F\{^{1}H\}$ NMR spectrum of the 4-FT **33** bio-transformation with the PLP-dependant transaldolase (*SCO1844*) omitted from the reaction.

The resultant $^{19}F\{^{1}H\}$ NMR spectra revealed the accumulation of just two organofluorine signals, at -231.53 ppm and -231.55 ppm, the product ratios are detailed in Table 4.2. The minor product, with a chemical shift of -231.53 ppm was shown to be FAld **40** by add-mixing a synthetic standard with the reaction products. The major product (-231.55 ppm) corresponds to the accumulation of 5'-FDA **35**, which was again confirmed by add-mixing with a synthetic reference sample. The accumulation of 5'-FDA **35** identifies the

PNP enzyme as a bottle neck in these reactions, limiting the downstream conversion to other intermediates in the pathway.

The accumulation of FAld **40** is the expected outcome of this experiment, with no downstream enzyme to convert FAld **40**, the last common intermediate between FAc **8** and 4-FT **33** to the final fluorometabolites. However unlike the previous control experiments, there is no accumulation of other intermediate compounds between 5'-FDA **35** to FAld **40**. This may be rationalised by the results revealed in Chapter 3, whereby the SCO1844 fuculose aldolase product is shown to be extremely efficient at generating FAld **40** *via* the retro-aldol reaction of 5-FDRulP **39**. However, it is not so efficient in the reverse (aldol) direction. In fact, this fuculose aldolase is more efficient in generating an alternate diastereoisomer to 5-FDRulP **39**, 5-FDXulP **42** or 5-FDRhuP **41** at temperatures above 4 ºC. There is no evidence from $^{19}F\{^{1}H\}$ NMR of an alternate diastereoisomer being generated in this reaction and it is therefore deduced that this is because the equilibrium for this aldolase activity significantly favours the production of FAld **40** over the reverse reaction to 5-FDRulP **39** and 5-FDXulP **42**/ 5-FDRhuP **41**. As a direct result, the action of this enzyme pulls the equilibrium of the entire pathway towards the generation of FAld **40**, reducing the accumulation of upstream intermediates on the pathway. This in turn prevents the other enzymes from catalysing reverse reactions on their accumulated substrates. It is also noteworthy that no 5'-FDI **36** is generated in this reaction, suggesting that the adenosine deaminase activity displayed in the other reconstitution control experiments is contained in the FTase enzyme preparation.

4.2.2.3 Summary

The $^{19}F\{^{1}H\}$ NMR spectra from all of the reconstitution experiments are summarized in Figure 4.17, and the subsequent product ratios are tabulated in Table 4.2.

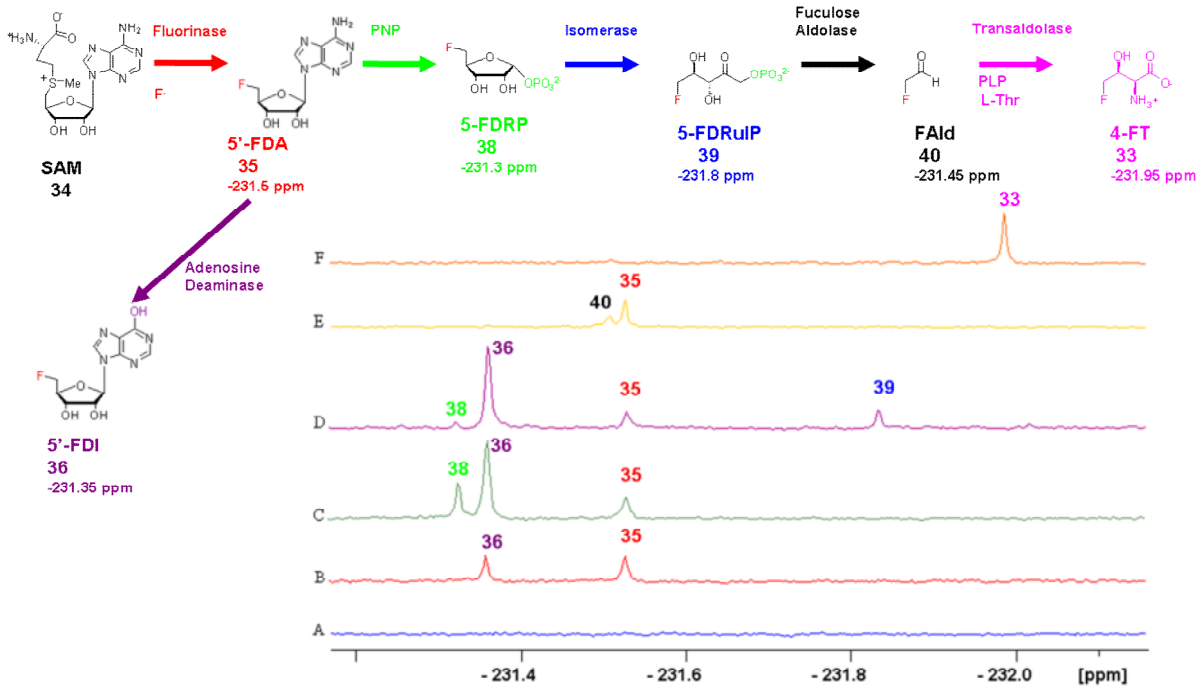


Figure 4.17. A summary of the **r**esultant $^{19}F\{^{1}H\}$ NMR spectra of *in vitro* reconstituted biotransformations when fluoride ion was incubated at 37°C for 16h with cloned and over-expressed enzyme combinations. Control experiments were carried out by removing one enzyme each in a stepwise manner. **Experiment A**; minus the **fluorinase**. **B**; minus the **PNP**. **C**; minus the **isomerase**. **D**; minus the fuculose aldolase (*S. coelicolor*). **E**; minus the **PLP transaldolase** (expressed in *S. lividans*). **F**; Complete pathway.

Fluorometabolite / Experiment	5'-FDA	5-FDRP	5-FDRulP	FAld	4-FT	5'-FDI
Fluorinase	-	-	-	-	-	-
PNP	1.0	-	-	-	-	0.88
Isomerase	0.34	0.43	-	-	-	1.0
Fuculose Aldolase	0.21	0.05	0.16	-	-	1.0
PLP Transaldolase	1.0	-	-	0.56	-	-
Complete Pathway	-	-	-	-	1.0	-

Table 4.2. The ratios of the products from the reconstitution control experiments, with one component of the pathway removed and the complete pathway experiment.

The product ratios from the reconstitution experiments (Table 4.2) illustrate the consequences of the stepwise removal of the enzymatic components of the *in vitro* fluorometabolite pathway. It is clear that in the presence of the fluorinase, the identified intermediates along the fluorometabolite pathway accumulate only upon incubation with the necessary enzyme preparations.

The activity of an adenosine deaminase, a low-level contaminant in the FTase transaldolase preparation, removes more than half of the 5'-FDA **35** generated by the fluorinase to the shunt product 5'-FDI **36**.[93] As 5'-FDI **36** cannot be metabolized by the PNP from *S. cattleya*, this product accumulates, directing organic fluorine away from the pathway. However in the presence of all of the component enzymes, and with high

concentrations of L-threonine (35 mM) and PLP (0.7 mM), the equilibrium of the PLP-dependant transaldolase is driven towards the 4-FT **33** product effectively making it irreversible. It is interesting to note that 5'-FDI **36** is not an identified product in whole cell bio-transformations of *S. cattleya*, but is often observed in CFE's. As such this prevents the accumulation of preceding fluorometabolites intermediates, and consequently inhibits the generation of 5'-FDI **36** as 5'-FDA **35** does not accumulate.

4.3 Conclusions

The biotransformation of the fluorometabolites FAc **8** and 4-FT **33** has been achieved by the *in vitro* reconstitution of the enzymatic steps that effect their synthesis in *S. cattleya*. The fluorinase, PNP and MTRI identified from *S. cattleya* were over expressed in *E. coli* and purified. A putative fuculose aldolase was identified from *S. coelicolor* as a surrogate enzymatic step to that in *S. cattleya*. The fluorinase from *S. cattleya* is the key C-F bond forming enzyme and incubation with the PNP, MTRI and the fuculose aldolase led to the production of FAld **40**, the last common intermediate in FAc **8** and 4-FT **33** biosynthesis. The gene for the FTase, responsible for FAld **40** conversion to 4-FT **33** in the presence of L-methionine, was recently identified, sequenced and cloned, by Dr Hai Deng (University of St Andrews). This enzyme was over expressed in *S. lividans* and the partially pure protein product was used to affect 4-FT **33** synthesis in the reconstitution experiments. The commercially available aldehyde dehydrogenase from *S. cerevisae* was used in conjuction with the other pathway enzymes to successfully generate FAc **8**.

The ability to recombine all of the enzymes to reconstitute the biotransformation of 4-FT **33** in particular, from fluoride ion opens up the possibility of utilising the approach to prepare 4-[^{18}F]-FT from [^{18}F]-fluoride. This is currently being explored to affect a radiolabelled synthesis of 4-[^{18}F]-FT.

5 The development of a novel fluorinase assay

5.1 Introduction

The low catalytic rate of the fluorinase enzyme makes it unsuitable for large-scale biocatalysis applications.[76, 112, 113] It is therefore attractive to explore methods by which the activity of fluorinase can be increased, for more wide scale application.

5.1.1 Mutagenesis of the fluorinase: Application of the C-F bond

One of the hotspots in modern biochemistry is the accelerated evolution of proteins,[164, 165] particularly improving enzyme performance, through directed evolution or rational design. Directed evolution is the "low frequency introduction of random mutations in a gene of interest".[165] Rational design is the changing of a specific residue in an enzyme, identified using structural and mechanistic information. Mutating the gene of a wild type enzyme to increase its activity or stereoselectivity has attracted a lot of interest and has yielded some successes.[177-179]

5.1.2 Crystallography and computational studies: Site-directed mutagenesis of the fluorinase

Crystal structure and QM/MM calculations[77] have established the putative hydrogen bonding networks important for catalysis and the integrity of the active site pocket of the fluorinase. Rational design extending from these analyses is unlikely, however, to yield site-directed mutants with improved kinetics. The active site possesses intricate H-bonding networks which are essential for catalysis. It is more likely that mutations beyond the active site, may have a more profound effect on the catalysis. Random mutagenesis can lead to enzyme mutants with increased activity, caused often by a single residue change which would not be predicted to effect catalysis. In these cases it is much harder to predict the effect of any one mutation on catalysis. The identification of individual residues outside the active site using this method would then form the basis of saturation mutagenesis at that particular residue, which may yield further improvements in reactivity.

5.2 A novel assay for fluorinase activity

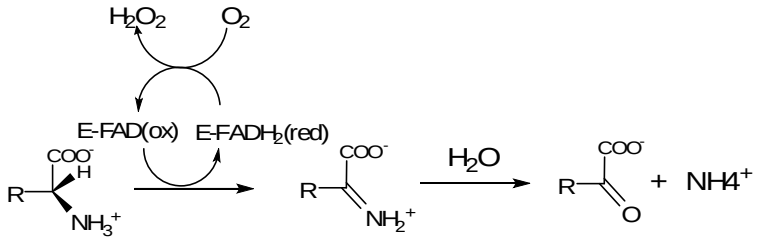
To succeed at random mutagenesis for enzyme evolution, certain pre-requisites must be satisfied.[165] The most crucial of these is the requirement for a reliable high-throughput assay that is capable of identifying mutant enzymes with increased activity. By its nature, random mutagenesis produces many thousands of mutants. The vast majority of these mutants are negative, with only very few likely to produce a positive effect. The development of an assay that is capable of screening thousands of mutants at once in order to identify the positive mutants is critical.

5.2.1 L-amino acid oxidase and horseradish peroxidase-coupled assay for L-methionine detection.

A novel assay has been developed by Ingenza[177-179] that incorporates an L-amino acid oxidase (LAAO), horseradish peroxidase (HRP) and the dye 3',3-diaminobenzidine (DAB). This assay is capable of producing a colour change as a direct indicator of L-methionine concentration in solution, it was applied to the fluorinase reaction in order to detect the product L-methionine, which is co-produced with 5'-FDA **35**.

5.2.2 L-amino acid oxidases

LAAOs are homodimeric flavoenzymes containing non-covalently bound FAD as a cofactor. They catalyse the stereospecific oxidative deamination of amino acid substrates to the corresponding α-keto acids with the concomitant production of ammonia and hydrogen peroxide (Scheme 5.2).[166]



Scheme 5.2. A general mechanism for the reaction catalysed by LAAOs. [166]

These enzymes are widely distributed in many different organisms and have been purified from the bacteria *Rhodococcus opacus*[166] and snake venom.[167] In each case their crystal structures have been determined.[166, 168]

The LAAOs of bacterial, fungal and plant species are utilized in nitrogen starved conditions during which time their expression is upregulated. Upon LAAO expression amino acids, purines, nitrate, proteins and/or peptides are metabolized in the absence of any readily metabolisable nitrogen sources e.g. ammonium, glutamine and glutamate.

The function of the LAAOs identified from snake venom is poorly understood. It is thought that they may play a role in apoptosis[169, 170], interact with platelets responsible for blood clotting or they may act directly as toxins.[171] LAAOs are the major component of many snake venoms and as such they are relatively easy to purify. The snake venom LAAO from *Crotalus adamanteus*, the venomous rattlesnake, is commercially available.

5.2.3 Horseradish peroxidase

Horseradish peroxidase (HRP) is a prototypical hemoprotein peroxidase which is isolated from the roots of horseradish. This protein is capable of catalysing the oxidation of small organic substrates in the presence of hydrogen peroxide. Biological oxidation reactions catalysed by HRP involve high-oxidation state Fe intermediates (Figure 5.1).

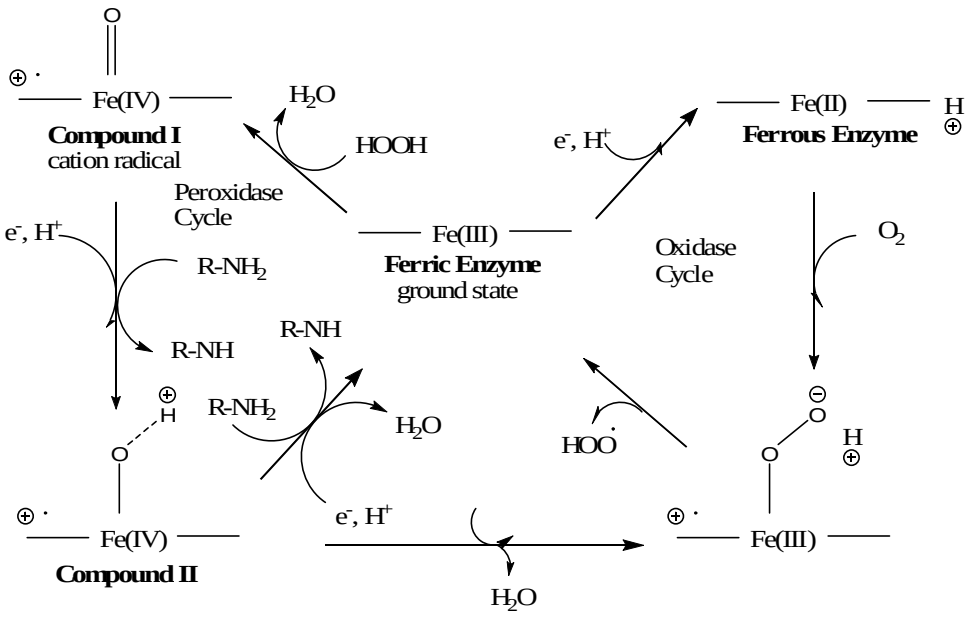
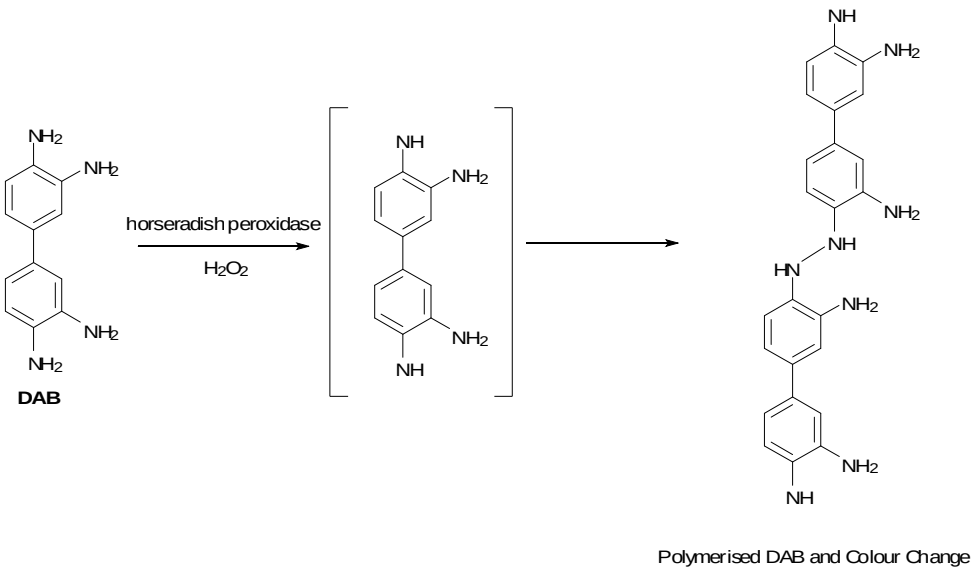


Figure 5.1. Oxidation states of Fe in the catalytic cycle of horseradish peroxidase. R= DAB.[172]

The ferric state (ground state) of the HRP enzyme reacts with H_2O_2 to give compound I, a two-electron oxidized species in which the heme is oxidized to a ferryl porphyrin π cation radical. Compound I undergoes successive reductions by small molecule substrates to first generate compound II and then again to return to the ground state. One turn of the peroxidase cycle by HRP generates two oxidized substrate molecules (Figure 5.1), which initiates a colour change in a colorimetric assay. A typical substrate for the colorimetric assay of HRP is 3,3-diaminobenzidine (DAB) (Scheme 5.2). Such dye molecules are used in immunohistochemical staining, often in the identification of cancers.[173] It is noteworthy that this reaction occurs in the absence of small molecules as substrates.[177-179]

5.3 LAAO and HRP-coupled liquid phase assay of L-methionine concentration

The combination of an LAAO and HRP in the presence of L-methionine and the colorimetric substrate DAB was investigated as a possible visual assay for the fluorinase. The assay is based upon the production of L-methionine as a rate determining side-product of the fluorination reaction. L-Methionine is a substrate for LAAO, producing H_2O_2 and NH_3 as a result of oxidative deamination.[167] The resultant H_2O_2 is then converted by the horseradish peroxidase, which acts to catalytically oxidise two DAB molecules. Oxidised DAB dimerises, causing a change in colour to an insoluble brown product (Scheme 5.2). The rate and intensity of the colour change is indicative of the rate of reaction.



Scheme 5.2. The catalytic oxidation and dimerisation of DAB by HRP.

The assay was applied to fluorinase activity either as a purified solution (liquid phase) to follow the fluorination reaction in real time or expressed intracellularly (solid phase) in order to assay mutant fluorinase clones expressed in *E. coli*.

5.3.1 The liquid phase assay

For the liquid phase assay the colour change caused by DAB dimerisation was monitored by UV spectroscopy (480 nm, with an extinction coefficient of 5,500 M^{-1} [174]). The colour change was related to the specific activity of the fluorinase. Snake venom LAAO from *C. adamanteus* (Sigma Ltd, UK) was used in these experiments. DAB with a metal enhancer (cobalt) was also purchased, which is intended to intensify the colour change caused by oxidation of DAB by H_2O_2 and HRP.[175]

5.3.1.1 Liquid phase assay controls

In order to establish the proof of principle of the liquid phase assay, two sets of control experiments were carried out. L-Methionine (Sigma Ltd, UK) was suspended in ultra-pure water (100 mM). Commercially available LAAO and HRP were suspended in phosphate buffer (10 mM, pH 7.5) to final concentrations of 5 mg/ml and DAB solution was prepared using the manufacturers instructions, into ultrapure water (one tablet, 25 ml).

In eleven separate experiments LAAO, HRP (0.5 mg/ml respectively) and DAB solution (40 µl) were added into a UV cuvette. L-Methionine (100 mM) was then added to

different final concentrations (0, 0.0125, 0.025, 0.05, 0.1, 0.125, 0.2 0.25, 0.5, 1, 2 mM respectively). The solutions were then made up to a final volume of 1 ml using ultrapure water, and the reactions were measured after incubation in the dark after 10 and 30 min in a spectrophotometer (480 nm) (Figure 5.2).

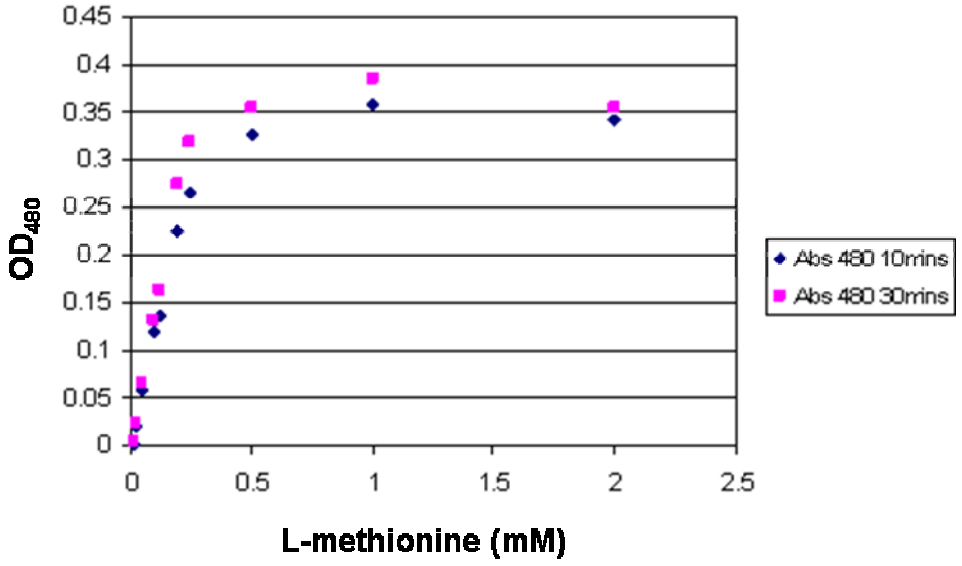


Figure 5.2. A graph showing the colour change of DAB monitored at 480 nm using the components of the liquid phase assay in the presence of different L-methionine concentrations . Values were taken after 10 (blue) and 30 min (pink) incubation at room temperature, with logarithmic error values of R_2= 0.93 and 0.90 respectively.

The data was presented in Figure 5.2 and shows that the UV response of oxidized DAB is directly linked to the concentration of L-methionine in the solution. After 10 min, the relationship between the the OD value and L-methionine concentration exhibits a typical saturation curve with Michaelis-Menten kinetics. As such a logarithmic curve can be fitted to these results, with low error (R_2=0.93). This curve behaves as a one-substrate reaction, indicating that the two-enzyme coupled production of oxidized DAB from

L-methionine is working efficiently in a combined system. The rate of change in OD becomes non-linear in these assays at more than ~0.25 mM of L-methionine in solution. The colour change of these experiments peak at ~0.35 (OD_{480}) at an L-methionine concentration of 0.5 mM. This suggests that other factors, such as DAB availability and the concentration of enzymes are affecting the rate of colour change in the presence of L-methionine concentrations higher than 0.25 mM.

After 30 min incubation, the relationship between L-methionine concentration and OD is nearly identical to that after 10 min and a log curve can again be fitted to the data (R_2=0.93). There are very small differences in the OD values attained after 10 and 30 min, suggesting that the colour change reaction has concluded after 10 min. It can be concluded that the DAB dimerisation in this assay, is capable of indicating L-methionine concentration close to real time.

A further six experiments were then set up in an identical manner using the same components with different L-methionine concentrations (0, 0.0125 0.025 0.05 0.1, 0.2 mM). In the subsequent assays, the DAB preparation was used which included a metal enhancer (Co^{2+}), reported to increase the intensity of the OD response of DAB oxidation. The reactions were measured after incubation in the dark, again for 10 and 30 min at RT and DAB dimerisation was monitored by UV (480 nm). The data are presented in Figure 5.3.

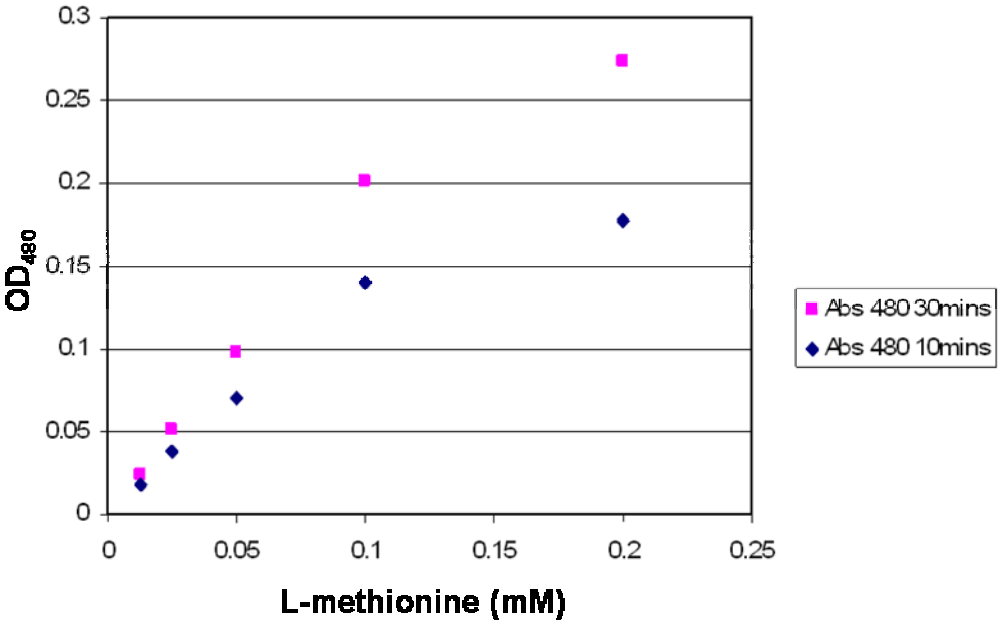


Figure 5.3. The OD_{480} of DAB with Co^{2+} a metal enhancer using the components of the liquid phase assay in the presence of different L-methionine concentrations . Values were taken after 10 (blue) and 30 m (pink) incubation at room temperature, with logarithmic error values of R_2= 0.96 and 0.95 respectively.

This liquid phase assay with DAB (Co^{2+}) with a range of L-methionine concentrations for 10 min with a curve fitted revealed lower errors than the previous experiments (R_2= 0.96). The L-methionine concentration range in these experiments covered the linear range of the OD_{480} response previously determined in Figure 5.4 above. Conversions after 30 min showed a similar shaped curve to that observed after 10 min incubation with low error (R_2= 0.95), however the OD_{480} response was more intense. The OD_{480} values that are recorded in Figure 5.3 at 10 min are significantly lower than those at 30 min suggesting that DAB oxidation is slower, in the presence of Co^{2+}. However, at similar conversions, the colour change is more intense.

The data from a comparison of DAB vs DAB (Co^{2+}) after both a 10 min and 30 min incubation are compared in Figures 5.4 and 5.5 below.

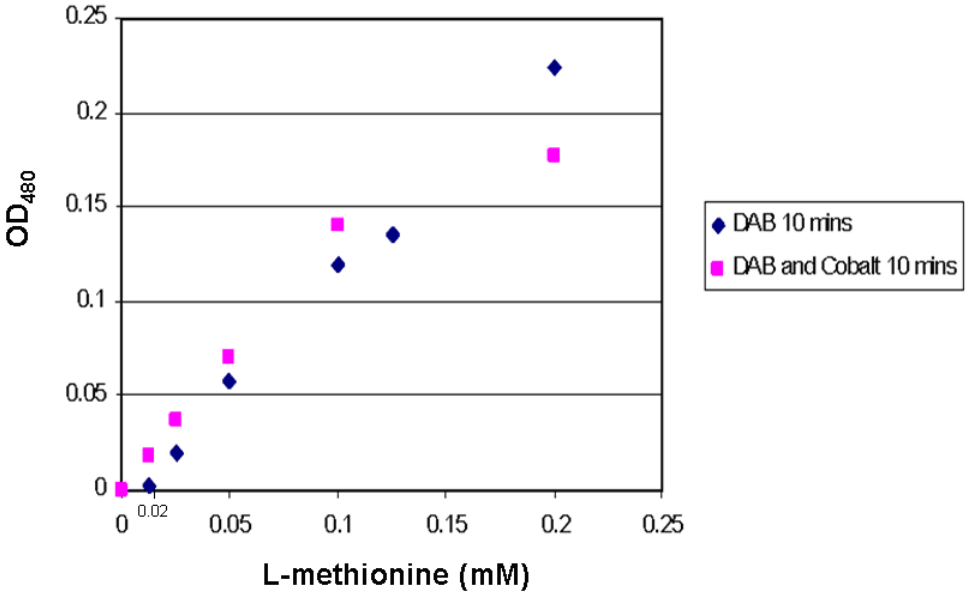


Figure 5.4. A comparison of the linear colour change of DAB (blue) and DAB (Co^{2+}) (pink) using the components of the liquid phase assay in the presence of different L-methionine concentrations . Values were taken after 10 min incubation at RT, with linear error values of R_2= 0.96 and 0.95 respectively.

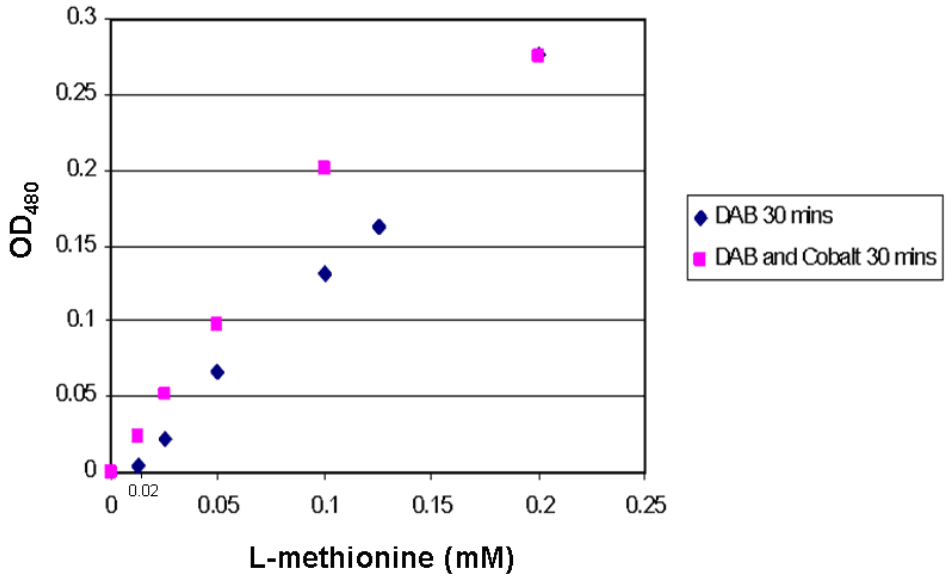


Figure 5.5. A comparison of the linear colour change of DAB and DAB (Co^{2+}) using the components of the liquid phase assay in the presence of different L-methionine concentrations. Values were taken after 30 min incubation at RT, with logarithmic error values of R_2= 0.99 and 0.95 respectively.

It is clear that Co^{2+} has an enhancing effect after both 10 and 30 min incubation. This is particularly apparent at L-methionine concentrations below 0.1 mM. DAB (Co^{2+}) produces an OD_{480} change that is ~30 % greater than DAB alone. Figures 5.4 and 5.5 also shows that using DAB alone, there is a limit to L-methionine detection in these assays at about 0.02 mM. This limit of detection is not apparent when using DAB with the metal enhancer. However, the linear response of DAB (Co^{2+}) is curtailed in comparison to DAB alone i.e it becomes non-linear at lower L-methionine concentrations. The response of DAB alone maintains its linearity at L-methionine concentrations higher than 0.1 mM, and as the DAB with metal enhancer response diminishes the two experiments have near identical OD_{480} at 0.2 mM.

5.3.1.1.1 Conclusions

From these experiments it can be concluded that the LAAO-HRP coupled assay in solution is capable of detecting L-methionine through the oxidation of DAB in a single substrate Michaelis-Menten response. LAAO and HRP are sufficiently coupled to allow quantitative analysis of the OD_{480} response to L-methionine. It has been determined that DAB oxidation in these reactions occurs within at least 10 min of incubation with L-methionine, and holds prospects for a fluorinase assay. The experiments were also compared with a second dye substrate, DAB with Co^{2+} as an enhancer. The OD_{480} response was increased marginally with Co^{2+}, however this modification had a lag time and a less linear relationship at L-methionine concentrations above 0.1 mM relative to DAB alone. For these reasons, this second dye was not considered a good candidate for

future assays. With this model study in place, an assay of the fluorinase was now explored.

5.3.1.2 A fluorinase assay

In order to explore the DAB oxidation as a UV-based fluorinase assay, the fluorinase enzyme was over expressed and purified to ~10 mg/ml into phosphate buffer (10 mM, pH 7.8) as described previously. It was combined with commercially available LAAO and HRP which were dissolved in phosphate buffer (10 mM, pH 7.5) to final concentrations of ~5 mg/ml. A DAB solution was prepared in ultrapure water (25 ml) and then LAAO, HRP (0.5 mg/ml respectively), DAB solution (40 µl), SAM **34** (1 mM), fluoride (20 mM) were combined in a 1 ml UV cuvette for a zero time reading and the fluorinase (0.6 mg) was added to initiate the reaction. Control experiments were set up simultaneously, with the fluorinase, SAM **34** and fluoride omitted individually from the reactions. All of the subsequent reactions were then incubated at room temperature for 90 min and monitored by UV (480 nm) every 2 min. The time versus OD_{480} data is presented in Figure 5.6 below.

Figure 5.6 clearly indicates that fluorinase activity can be detected using this novel assay. Incubation of the fluorinase with SAM **34** and fluoride ion led to the generation of oxidized DAB over time in a linear relationship to a maximum 0.38 OD_{480} after 120 min, which corresponds to ~1 mM L-methionine. Control experiments revealed that in the absence of one of the components of the fluorinase reaction, DAB oxidation was significantly reduced. In the absence of the fluorinase, some DAB oxidation is apparent,

corresponding to about a third (0.11 OD_{480}, 120 min) of that achieved in the presence of the fluorinase. This reveals that there is a significant background reaction causing DAB oxidation in these assays. In the absence of fluoride ion, the oxidation of DAB is also diminished (Max OD_{480} 0.11), revealing that the presence of fluoride ion in the reaction is critical in order to effect the oxidation of DAB in these experiments. The final control experiment, with SAM **34** omitted from the reaction, reveals a very low level of DAB oxidation (0.025 OD_{480}, 120 min). This indicates that the SAM **34** preparation is the source of the background reaction exhibited in the other control experiments.

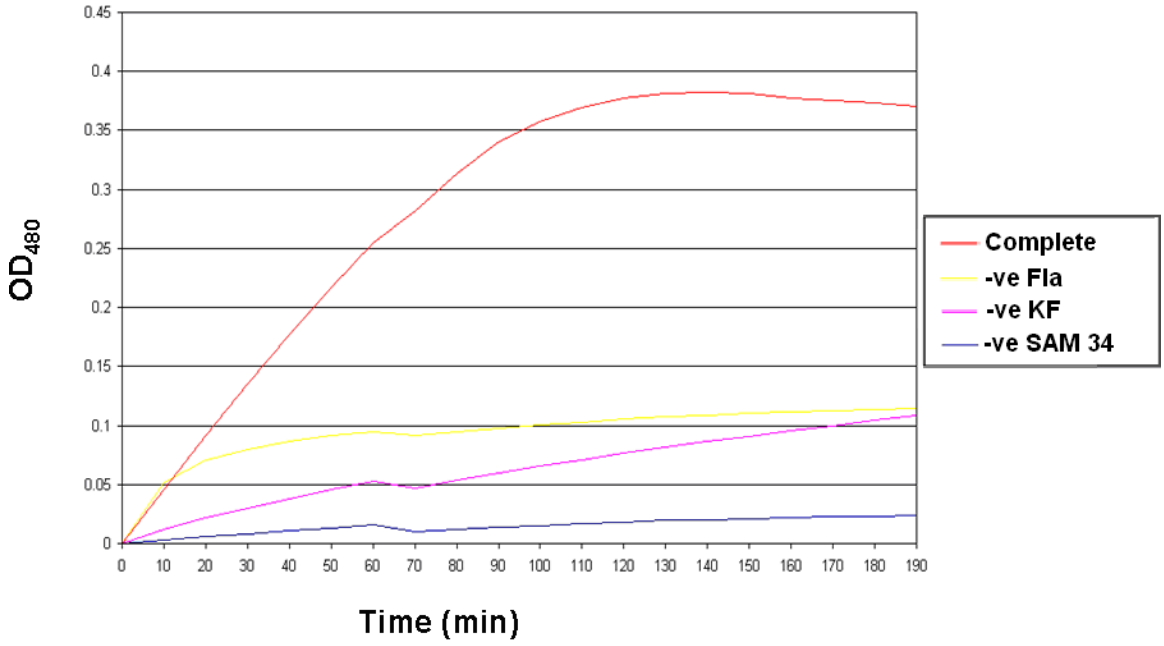


Figure 5.6. Fluorinase assays monitoring the Absorbance of DAB versus time. Red= Complete reaction. **34**. Yellow= minus fluorinase. Pink= minus KF. Blue= minus SAM **34**.

It is possible that the background reaction is caused by contaminating L-methionine in the SAM **34** sample, as a result of SAM **34** breakdown over time. Commercially available SAM **34** is only ~90 % pure, and according to the manufacturer will degrade at

a rate of 10 % every hour at room temperature in solution. Therefore there is a time-dependant release of SAM **34** degradation products in the course of these reactions, which may mask the release of L-methionine as a consequence of the activity of the fluorinase. This effect probably results in the exceptionally high estimation of L-methionine from this assay (1 mM). In the control experiment with fluoride ion omitted, the presence of the fluorinase enzyme may have some stabilizing effect on SAM **34**. In this case, the initial rate of SAM-derived DAB oxidation is reduced compared to the control in which fluorinase is omitted from the reaction. This may also suggest that the presence of F^- has a role in SAM **34** degradation. As a result it is difficult to accurately assess the effect of SAM **34** degradation on the final results of these liquid phase assays.

5.3.1.2.1 Liquid phase assay of the fluorinase: Pre-treated SAM

In order to reduce the effect of SAM **34** degradation in these reactions, attempts to "clean up" SAM **34** prior to its use in these experiments were undertaken. It was noted previously that the HRP enzyme is capable of reducing H_2O_2 in the absence of DAB or equivalent substrates.[177-179]

A pre-treatment was envisaged in order to oxidise the contaminating L-methionine contained in the SAM **34** preparation before adding the DAB substrate. This should reduce the background DAB oxidation experienced in the liquid phase assays. To this effect, all subsequent SAM **34** solutions used were first incubated with the LAAO and HRP enzymes for 20 min at room temperature before the remaining assay components

were added (Figure 5.7). Thus any L-methionine should be oxidized, prior to the fluorinase reaction assay.

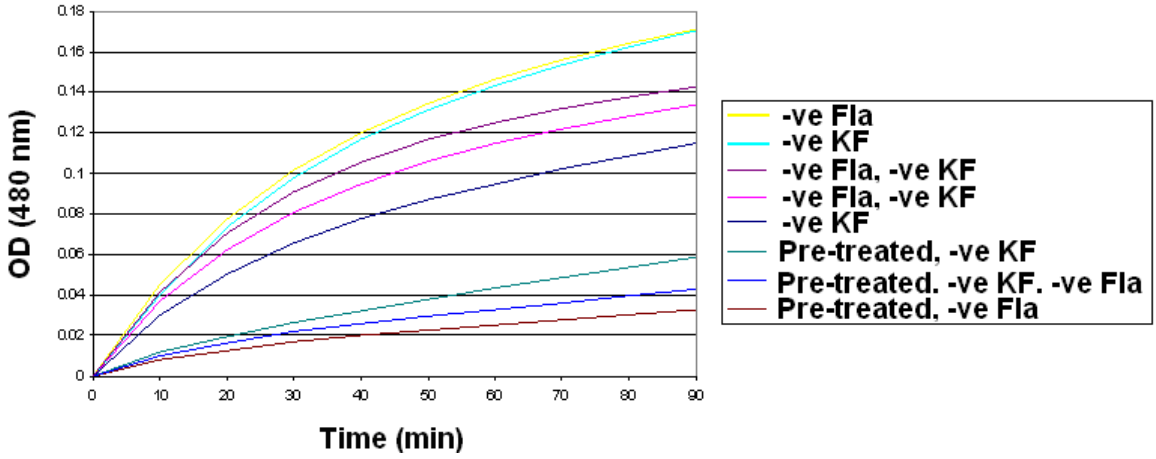


Figure 5.7. Comparison of controls with SAM **34** and pre-treated SAM **34**. Each control has a component of the fluorinase-liquid phase assay removed.

The pre-treatment has a clear effect and was capable of reducing the effect of SAM-derived oxidation by approximately 66 %. Indeed the pre-treatment reduced the background reaction rates to those controls in the absence of SAM **34**. This experiment reinforced the idea that the background reaction was indeed due to contaminating L-methionine in the SAM **34** preparation.

Following the relative success of the "clean-up" of SAM **34** it was felt that this method could be used as a quantitative real time assay of the fluorinase reaction. To this effect, pre-treated SAM **34** solution was prepared as previously described and the liquid phase assay components were set up as described above. The assays were conducted with

different SAM **34** concentrations (0, 0.25, 0.5, 1 and 2 mM) and incubated at room temperature. DAB oxidation was followed by UV (480 nm) every 10 min and the resulting colour change (OD_{480}) over time is shown in Figure 5.8 below.

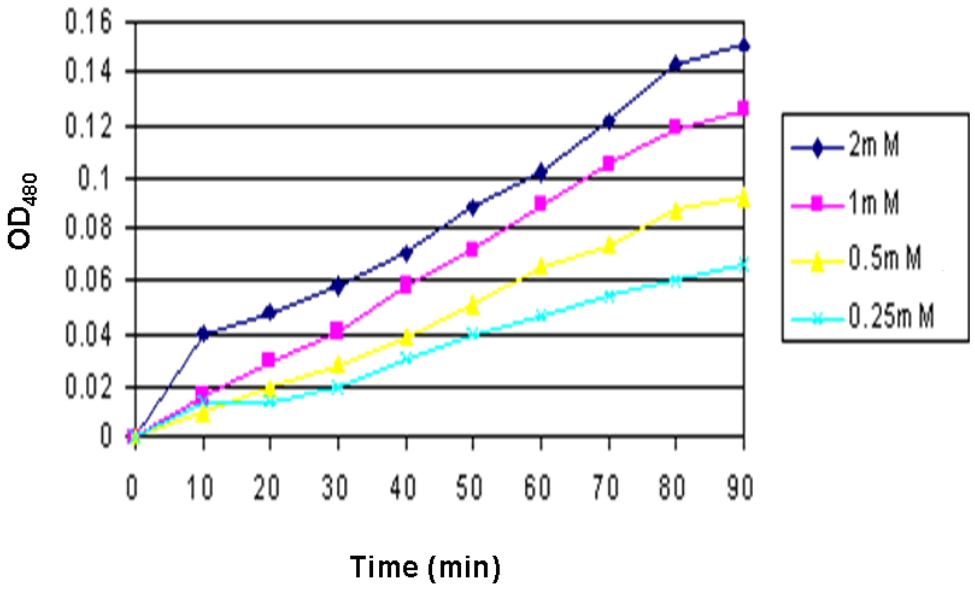


Figure 5.8. The UV absorbance of DAB at 480 nm over time in liquid phase assays containing different concentrations of SAM **34**.

Figure 5.8 shows the OD_{480} change over time of the fluorinase reaction at different concentrations of SAM **34**. In all of the reactions there is a burst in OD_{480} after 10 min, which is attributed again to residual products of SAM **34** breakdown affecting a colour response as previously described. After the initial burst in colour change the OD_{480} response establishes a secondary linear relationship. This response is attributed to the activity of fluorinase. Despite cleaning up SAM **34**, residual L-methionine is present, masking the initial activity of fluorinase in these experiments. As expected, the overall OD_{480} change, and consequent fluorinase activity, is influenced by SAM **34**

concentration. It is possible to determine a difference in the rate of the fluorinase reaction in these liquid phase assays.

In order to validate this assay it was then compared with an established fluorinase assay which uses HPLC to detect 5'-FDA **35**,[176] in the assay mixtures. Samples (10 µl) were taken from the experiments described in Figure 5.8 after 30, 60 and 90 min respectively. These samples were subject to heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min), before being made up to 100 µl with ultrapure water. The resulting mixture was subject to HPLC analysis to determine the FDA **35** concentration compared to DAB oxidation. The resultant 5'-FDA **35** concentration calculated from these samples were plotted against the measured OD_{480} value attained in the liquid phase assay of that sample (Figure 5.9).

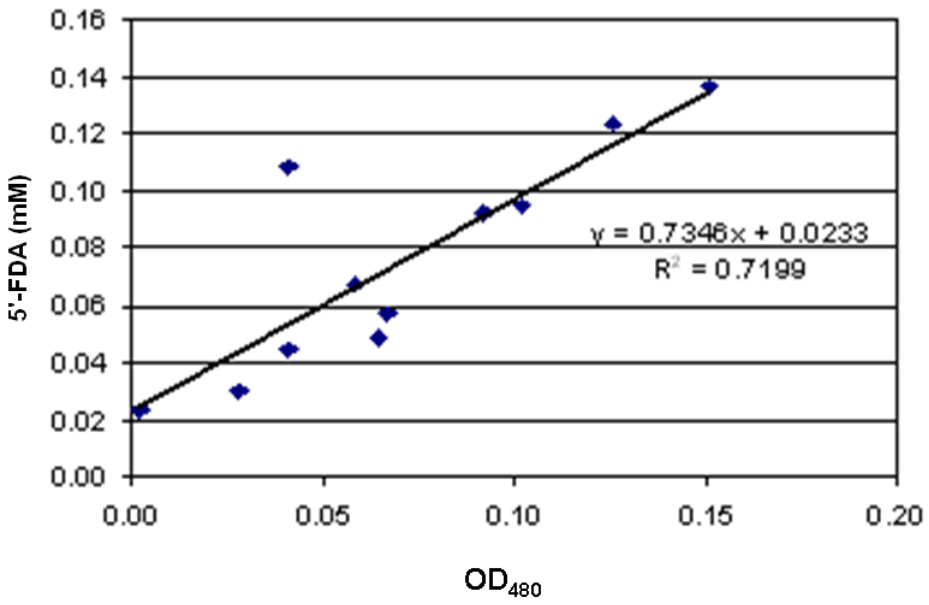


Figure 5.9. The OD_{480} response of the liquid phase assay plotted against the concentration of FDA **35** in samples analysed by HPLC. R^2=0.72.

The results of HPLC determination of 5'-FDA **35** concentration (by HPLC) was compared to the OD_{480} (L-methionine oxidation) in Figure 5.9. A line of best fit reveals an approximately linear relationship although its statistical significance (R^2= 0.72) is poor. However we can determine from these results that there is a relationship between 5'-FDA **35** production by the fluorinase and the OD_{480} generated by the liquid phase assay. This data also reveals that there is a threshold limit of around 0.02 mM 5'-FDA **35** (and therefore L-methionine) before a change in OD_{480} is effected. This is consistent with the DAB response recorded in previous analysis (Figures 5.4 and 5.5), using L-methionine standards. Discrepancies in the response of the liquid phase assays compared with the 5'-FDA **35** concentration are most likely due to the background effects of SAM **34** degradation and the preparation of sample for HPLC analysis.

The quantitative analysis of the kinetics of the fluorinase reaction was explored using the liquid phase assay with pre-treated SAM **34**, however results were an order of magnitude different from those attained using the more accurate HPLC methodology.[76, 87]

5.3.2 Conclusions

A new fluorinase assay has been developed measuring L-methionine release by LAAO, HRP and DAB. It has been demonstrated that there is a linear relationship between DAB oxidation and L-methionine concentration. Coupling of this assay with the fluorinase (L-methionine released) was achieved. Comparisons with control assays gave confidence that colour change of DAB dimerisation in the liquid assay is induced by the fluorinase activity.

There is a significant background rate due to SAM **34** degradation products, which prevented the accurate analysis of fluorinase activity in solution. “Cleaning up” SAM **34** using LAAO and HRP in the absence of a dye substrate is effective at reducing the initial background rate which is ~30 % of untreated preparations. However, SAM **34**, which is unstable at room temperature, continues to degrade throughout the reaction. This, coupled with the relatively low activity of the fluorinase, makes it difficult to determine fluorinase activity accurately. The measured K_m of SAM **34** using these assays (790 μM) is two orders of magnitude higher than that determined using the more accurate HPLC methodology.

A comparison of DAB oxidation (OD_{480}) with the 5’-FDA **35** concentration in solution also revealed high error rates for the new assay. Discrepancies in these values suggests that the liquid phase assay is not capable of generating reliable quantitative data, most likely again due to a background reaction, low catalytic turnover of the fluorinase at a non-optimal temperature and the threshold L-methionine concentration for DAB oxidation. In practice, it is difficult to generate reproducible quantitative results using this liquid phase assay. However it has some utility for determining the real time activity of the fluorinase after purification without the need for HPLC analysis.

This assay may have a role if adapted to visually highlight *E. coli* colonies expressing mutant fluorinase clones. Such an assay is required to screen after random mutagenesis, to identify improved activity by high-throughput methods.

5.3.1.3 Solid phase assay for the fluorinase

A solid phase assay to determine fluorinase activity was explored, based upon the same principles as the liquid phase assay. It was developed in conjunction with Ingenza, a company in Edinburgh, involved in the evolving amine oxidases. Solid phase assays are used by Ingenza to identify active mutant clones of these oxidases, expressed by individual colonies of *E. coli* cells (Figure 5.10).[177, 178, 179]

Figure 5.10. Photograph of *E. coli* cells transfected with mutant amine oxidase clones. Positive mutants are identified by DAB oxidation, which is produced in a rate-determining manner.[178]

It was envisaged that coupling the production of L-methionine, a side product of the fluorinase reaction, to an adapted version of the liquid phase assay, could provide a method for screening mutated fluorinase in *E. coli* colonies. The fluorinase gene, subjected to random mutagenesis, would be entered into an *E. coli* expression vector and used to transform competent *E. coli* cells. In the presence of IPTG, the mutant genes can then be expressed to generate the corresponding enzyme mutant. Each individual *E. coli* colony represents one mutant copy of the fluorinase gene. As a result, following the

expression of the mutant enzymes, those with enhanced activity can be identified using the colourimetric solid phase assay, as shown in Figure 5.10 above. The solid phase assay was developed for the high-throughput screening of fluorinase mutants, as it is capable of screening thousands of mutants simultaneously (Figure 5.11).

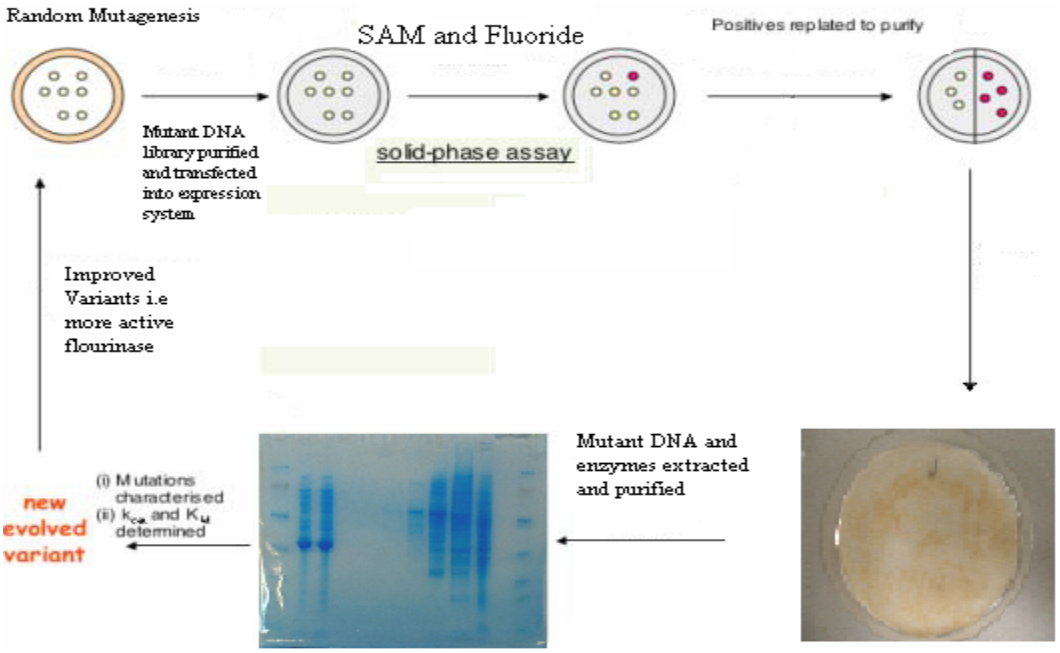


Figure 5.11. Schematic process of the solid phase assay of many thousands of mutant fluorinase plasmids, generated by random mutagenesis. It may be possible to undergo several rounds of mutation to generate mutants with significantly improved activity identified by the rate-determining production L-methionine leading to the oxidation of DAB.

Unlike the liquid phase assay, a solid phase assay does not require enzyme purification and enables many rounds of mutagenesis and analysis over a short period of time. In this assay, competent *E. coli* cells are transformed with an expression plasmid containing a mutated gene plated on antibiotic-containing LB Agar plates, with IPTG to drive the expression of the mutant enzyme. The resulting cell colonies can then be faithfully removed from the agar surface with filter paper and exposed to cold-shock conditions,

which crack some of the cells open, making their cell membranes leaky. This exposes the over expressed mutant enzyme to the assay mixture containing substrates for the desired reaction and the LAAO, HRP and DAB components for the colorimetric assay. Once exposed to the assay mixture, in the case of fluorinase, L-methionine will be produced during the generation of 5'-FDA **35** from SAM **34** and fluoride ion. The released L-methionine should initiate the dimerisation of DAB in the immediate area of the colony. Colonies which have expressed active fluorinase should be identifiable by a change in colour, attributed to the oxidation of DAB. These colonies can be picked and used as a starting point for protein over-expression and purification. New mutant genes can then be characterized and also subject to further rounds of mutagenesis for iterative improvement.

In order to determine the proof of principle for this assay, competent *E. coli* cells transformed with the *Fla*-pET28(a) vector were used to test the ability of the solid phase method to identify individual clones expressing active fluorinase. Competent cells were transformed with the fluorinase plasmid and plated onto LB Agar plates which contained kanamycin and IPTG (1 mM, 16 h, 37 °C). An identical, control experiment was set up simultaneously without IPTG, preventing fluorinase expression.

Colonies from each plate were then faithfully removed from the surface of the Agar plates using sterile filter paper and subject to cold shock in liquid nitrogen (20 s) before being allowed to thaw at room temperature. A wash solution (1 ml) was prepared containing LAAO and HRP (0.75 mg) and fluoride ion (125 mM) and soaked into a fresh filter disc and the excess fluid removed. The discs containing cells were subject to the

wash solution, in order to remove any L-amino acids liberated by the cold shock procedure not associated with fluorinase activity.

After a 30 min incubation a second solution was prepared (2 ml), containing LAAO, HRP (0.25 mg respectively), fluoride ion (187.5 mM), SAM **34** (5 mM) and DAB (250 µl). This solution was soaked onto two sterile filter paper discs, in fresh petri dishes and the excess fluid was removed. The filter discs containing the cell colonies were placed over the top, ensuring contact across the whole surface of the discs, and incubated at 37 ºC. The plates were monitored every 15 min for colour change and the assays were removed from incubation after 30 min. A typical result is shown in Figure 5.12. Identical assays were also set up using DAB (Co^{2+}) in the place of DAB. Typical results of these experiments are shown in Figure 5.13.

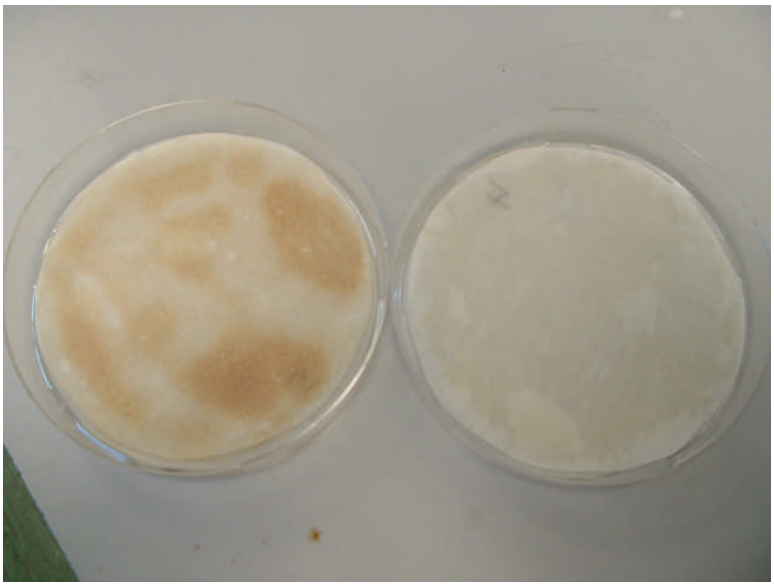


Figure 5.12. Two typical solid phase assays. The plate on the left shows colour change caused by active fluorinase, and on the right is a control where fluorinase has not been expressed by removing IPTG from the growth media.

The solid phase assay of the fluorinase appeared partially successful in that colonies induced to express the fluorinase enzyme were capable of oxidizing DAB when compared to control experiments. However the fluorinase-expressing colonies were not so well distinguished by this procedure certainly by comparison with Ingenza assays monitoring amine oxidase-expressing *E. coli* colonies. There is also a significant level of background coloration reaction which occurred in the control plates.

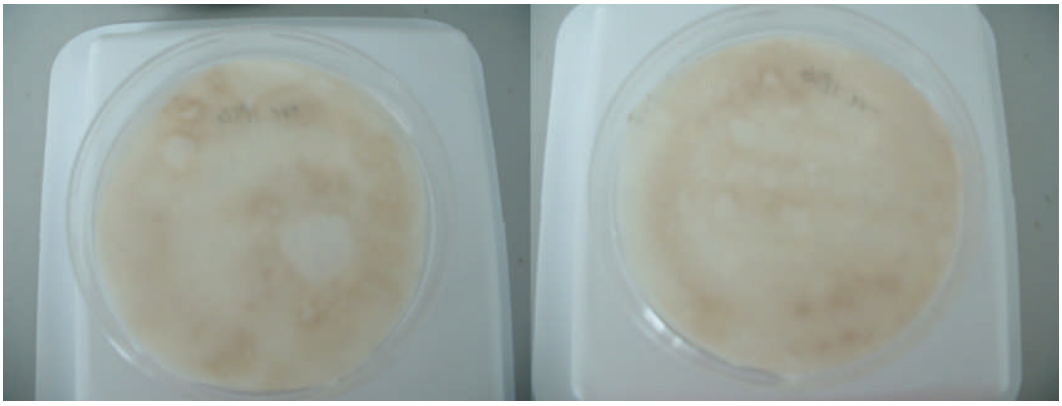


Figure 5.13. The solid phase assay using DAB and a metal enhancer on the left, control is on the right (without IPTG).

In the assays using DAB (Co^{2+}), there was no significant difference between the control and expressed fluorinase experiments. This suggests that there is either an adventitious side-reaction taking place, most likely with L-methionine derived from SAM **34** degradation, or that L-methionine produced by the fluorinase reaction is diffusing from the point at which it is generated. These effects are more prevalent when DAB (Co^{2+}) was used as the substrate for HRP. As a result, DAB alone was used for all subsequent experiments.

5.3.1.3.2 Solid phase assay of fluorinase using pre-treated SAM

Optimization of the liquid phase assay revealed a significant background reaction rate, caused by the contaminating degradation of SAM **34**. It proved possible to reduce the background reaction by up to 60 % by a pre-treatment with LAAO/HRP. Thus in an attempt to improve the solid phase assay, a similar pre-treatment was explored. In these experiments, the cells and solutions were set up exactly as previously described, except that in the assay mixture, SAM **34** was preincubated (30 min, RT) in the presence of LAAO and HRP in order to remove any residual L-methionine in the SAM **34** sample. Typical results of these assays are shown in Figure 5.14 below.

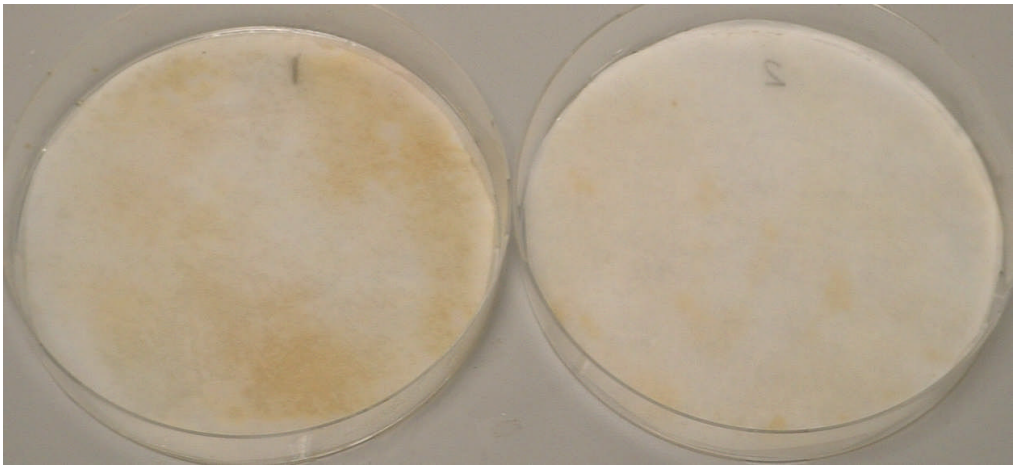


Figure 5.14. Two typical solid phase assays with pre-treated SAM **34**. The assay on the right-hand side, is a control (without IPTG).

The solid phase assays using pre-treated SAM **34** revealed a significant reduction in DAB oxidation when compared with typical assays shown in Figure 5.12. Clearly the majority of DAB oxidation occurring in the previous assays was actually generated by the

oxidation of SAM **34** breakdown products, rather than the fluorinase. In these experiments it was also apparent that DAB oxidation occurs in the control reactions, without fluorinase. This background reaction may be attributed to low-level SAM **34** breakdown during the course of the experiment at 37 ºC. There appears to be some difference between the control experiments and those assays where the fluorinase is expressed. This indicates that the fluorinase reaction is affecting DAB oxidation. Unlike Ingenza's assays, these assays do not identify individual colonies with active fluorinase. It appears that that there is significant diffusion of L-methionine, generated by the fluorination reaction, away from the fluorinase-expressing colonies.

5.3.1.3.3 L-methionine diffusion in agar assays

The main difference between the solid phase assay for amine oxidases compared with these fluorinase assays is that, in these experiments, LAAO is added in solution. The amine oxidase enzymes under assay were capable of producing H_2O_2 directly from various amine substrates.[177-179] As these oxidases were expressed intracellularly, any amine substrates would be oxidized in and around the colonies themselves, generating localized H_2O_2. This results in lighting up individual colonies when HRP causes the oxidation of DAB. In the fluorinase assays, the LAAO is in solution with HRP and this leads to diffusion of the L-methionine from its source. Consequently DAB oxidation does not highlight the individual colonies so effectively.

To exemplify the influence of L-methionine diffusion, an experiment was set up, whereby all of the components of the solid phase assay (LAAO, HRP, DAB) were

combined with 2 % agar in a petri dish to form an agar plate solid phase. Sample discs (5 mm diameter) of filter paper were then soaked in L-methionine (1 mM), and then dried and placed carefully onto the agar. A typical result is shown after a 30 min incubation at RT in Figure 5.15 below.

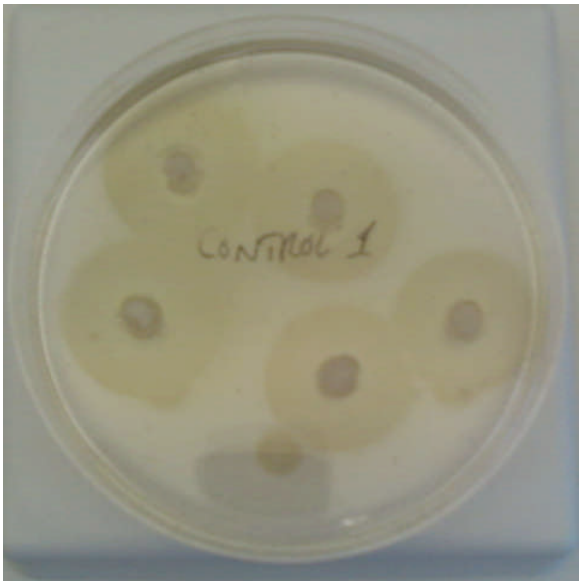


Figure 5.15. A solid phase agar assay incubated with filter paper discs soaked in L-methionine at room temperature for 30 min.

Figure 5.15 shows that even in the presence of 2 % agar, and using a dried localized source of L-methionine, DAB oxidation by the LAAO-HRP coupled assay is very diffuse. It can be assumed that the same effect is occurring in the solid fluorinase assays and that a localized LAAO is necessary for the identification of individual *E. coli* colonies.

5.4 Conclusions

A significant colour change was identified using the solid phase fluorinase assay. However, it was found that this colour change was also occurring, to a lesser extent in control reactions, most probably as a result of SAM **34** degradation. Despite reduction in the background reaction by pre-treatment of SAM **34**, identification of individual colonies was not possible using this assay. It was found that in the LAAO-HRP assay, L-methionine diffuses from its origin before being oxidized by LAAO. In previous work using this technique, the oxidation of amines is carried out by amine oxidases which have been expressed intracellularly. As a result amines are only oxidized in the immediate area of the colony, concentrating DAB oxidation to the colony itself (Figure 5.10).

Attempts to limit the diffusion of L-methionine using agar were not successful, and colour change in control experiments revealed that this effect was significant. The identification of a bacterial LAAO from *R. opacus* capable of the oxidation of L-methionine,[166] opened the possibility of expressing this gene on a low copy plasmid alongside mutant fluorinase. However, this LAAO does not express in *E. coli* and work by Ingenza revealed that this LAAO also acts on SAM **34** as a substrate (personal communication), something that does not appear to occur using the snake venom LAAO. Therefore this line of research was not pursued.

In conclusion, the LAAO-HRP assay in both the liquid and solid forms is capable of detecting fluorinase activity. However this effect is masked by the adventitious effect of SAM **34** degradation in solution, an effect observed in both forms of the assay. In the

solid phase assay, diffusion of L-methionine across the assay prevented identification of individual colonies expressing active fluorinase. As a result, this assay was judged not suitable for identifying mutant fluorinase clones in a high-throughput manner. The identification of an LAAO gene, which can be expressed in *E. coli* alongside the fluorinase gene, may make this technique a viable tool in the future. This would enable the solid phase assay of the fluorinase to be comparable to the successful assays previously used by Ingenza.

Conclusion and Future Work

The work in this thesis has detailed the successful validation of the intermediates and enzymes involved in the fluorometabolite pathway in *S. cattleya*. This was achieved by effecting an *in vitro* biosynthesis of the fluorometabolites using recombinately over expressed enzymes, identified from *S. cattleya* and also surrogate enzymes from other organisms. Consequently, 4-FT **33** and FAc **8** were generated from fluoride ion and SAM by metabolism of 5'-FDA produced by the fluorinase. Concurrently, the accumulation of intermediates involved in fluorometabolite biosynthesis was also achieved. This work also identified an isomerase enzyme at the genomic level from *S. cattleya* that is capable of utilising 5-FDRP **38** to generate 5-FDRulP **39**. The reannotation of a homologous enzyme found in *S. coelicolor* was also accomplished. A fuculose aldolase, also from *S. coelicolor*, has also been identified and characterised.

The fluorinase has been used extensively as a biocatalyst to generate novel [^{18}F]-labelled PET compounds. The successful *in vitro* generation of 4-FT **33** in particular has opened a route to an [^{18}F]-labelled version of this amino acid. Future work is needed to optimise the expression and reaction conditions of the reconstitution experiments to generate enough material to be a valid tool for PET. Also the ability to accumulate intermediates in the fluorometabolite pathway, opens up these new compounds to PET trials. This application of the fluorinase will also drive further efforts to identify a robust high-throughput assay for the fluorinase and the identification of the mechanism of the isomerase, and its potential role in *S. cattleya* will also be an interesting topic to explore.

6 Experimental

6.1 General Methods

All commercial reagents, chemicals or enzymes were purchased from Sigma Biochemicals, Fluka, Promega, Novagen and Fermentas unless otherwise stated. The following commercial enzymes were used. 5'-Adenylic acid deaminase (E.C. 3.5.4.6, from *Aspergillus species*, A1907, 0.11 units/mg) and immobilised PNP (*E. coli*, donated by GlaxoSmithKline, E.C. 2.4.2.1). *pFu* DNA polymerase (E.C. 2.7.7, M7741), GoTaq™ DNA polymerase (Promega, M5122), KOD polymerase (Novagen, 71085-3). PCR primers were designed in house and purchased from MWG Biotech. The competent cells BL21(DE3), BL21 (DE3) GOLD, C43(DE3), BL21 Star^{TM}, Rosetta 2 (DE3), and DH5α were purchased from Invitrogen in 50 µl aliquots.

All microbiological work was carried out in a Gallenkamp flowhood under sterile conditions unless otherwise stated. Glassware, media and consumables were sterilised by autoclaving. Centrifugation (>1000 µl) was carried out on a Beckman JA instrument at 14,000 rpm, at 20,000 (JA 25.50) or at 9,000 rpm (JLA 9.100). An Eppendorf 5415C centrifuge was used for microcentrifugation of volumes less than 1000 µl. Cell free

extracts for protein purification were acquired through sonication using a Sonics and Materials Inc., Vibra Cell. Ultra-pure water was collected from a USF Elga Maxima water supply system.

6.1.1 High Performance Liquid Chromatography (HPLC)

HPLC was carried out on a Varian Prostar system, consisting of a solvent delivery system (230, Prostar), a dual wavelength UV-Vis detector (325, Prostar) and a Prostar 400 autosampler. An analytical hypersil 5 µm C-18 column (250x10 mm, Phenomenex) was used at a flow rate of 1 ml/min. Sample volumes of 100 µl were used, of which 20 µl was automatically injected. Solvents were HPLC grade and filtered before use. The mobile phases consisted of two solvents, A, 50 mM KH_2PO_4: acetonitrile (95:5) and solvent B, 50 mM KH_2PO_4: acetonitrile (80:20). Runs were monitored at 254 nm by gradient elution over 30 min from 0% B to 100% B.

6.1.2 ^{19}F NMR Spectroscopy

^{19}F NMR analyses were performed on Bruker Avance 500 MHz (operating at 470 MHz) or Varian unity 500 MHz (operating at 470 MHz) spectrometers. All ^{19}F NMR spectroscopy was carried out using D_2O (~ 10 %) as an internal reference. Chemical shifts are given in ppm and coupling constants (*J*) are given in Hertz (Hz). Spectral coupling patterns are designated as follows; d: doublet and t: triplet. Spectra were analysed using TopSpin™ V. 2.1 (Bruker BioSpin).

6.2 Growth and maintenance of *S. cattleya* on agar

Streptomyces cattleya NRRL 8057 was originally supplied by Prof. D. B. Harper at the Queens University of Belfast, Microbial Biochemistry Section, Food Science Department, Belfast. Cultures were maintained on agar plates containing soybean flour (2 % w/v), mannitol (2 % w/v), agar (1.5 % w/v) and tap water. The plates were incubated at 30 ^{o}C for 28 days or until sporulation could be detected. The resultant static cultures were stored at 4 ^{o}C for future use.

6.2.1 Culture medium and growth conditions of *S. cattleya* [72]

Streptomyces cattleya seed and batch cultures were grown in conical flasks (500 ml) containing chemically defined medium (90 ml). The medium was prepared as follows. Sterile ultra-pure water (450 ml) was added to ion solution (150 ml), filtered carbon solution (75 ml), (see Section 5.1.4), sterile phosphate buffer (75 ml, 150 mM, pH 7.0) and sterile potassium fluoride (3 ml, 0.5 M). The seed cultures were prepared by transferring spores from a static culture as described above, and added to a conical flask (500 ml) containing chemically defined medium (90 ml). After incubation for 6 d at 28 ^{o}C on an orbital shaker (180 rpm), an aliquot (0.3 ml) of spores was used to inoculate the batch cultures. The batch cultures were incubated at 28 ^{o}C, on an orbital shaker at 180 rpm for 6-8 d.

6.2.2 Media for growing *S. cattleya*

6.2.2.1 Ion solution

The following reagents were added to ultra-pure water (900 ml).

NH_4Cl	6.75 g
NaCl	2.25 g
$MgSO_4.7H_2O$	2.25 g
$CaCO_3$	1.13 g
$FeSO_4.7H_2O$	0.113 g
$CoCl_2.6H_2O$	0.045 g
$ZnSO_4.7H_2O$	0.045 g.

The solution was sterilised by autoclaving prior to use.

6.2.2.2 Carbon source solution

The following reagents were added to ultra-pure water (900 ml).

glycerol (45 g)

monosodium glutamate (22.5 g)

myo-inositol (1.8 g)

para-aminobenzoic acid (450 μl of freshly prepared solution 1 mg/ml)

The solution was sterilised by filtration into pre-sterilised Schott bottles.

6.3 Preparation of resting cell cultures of *S. cattleya*

After 6 days of growth, *S. cattleya* cells were harvested by centrifugation (9,100 rpm / 25 min) and the resulting pellet was washed three times with phosphate buffer (50 mM, pH 6.8). After the final wash, the bacterial pellet was stored at –80 ^{o}C or could be used directly for genomic DNA extraction using the Wizard™ genomic DNA purfication kit (Promega).

6.4 Transformation of Competent Cells

Chemically competent *E. coli* cells were purchased from Invitrogen (BL21 (DE3), C43 (DE3) and DH5α) or derived from them. To transform these cells, each aliquot (50 µl) was mixed gently on ice and allowed to thaw for 3-5 min. After thawing, 200 ng of plasmid DNA solution (~1-2 µl) was added directly to each aliquot, stirred gently and returned to ice for 5 min. The cells were then exposed to 42 ºC for exactly 30 s, before returning to ice for 2 min. SOC medium (250 µl) was added to each vial and the vials were then incubated at 37 ºC whilst shaking at 250 rpm for 60 min. An aliquot (50 µl) of the transformed cells was then added to agar plates containing antibiotic determined by the plasmid vector that was used, distributed using a plate spreader, and maintained at 37 ºC for 16 h. Single colonies from these plates were then added to LB medium (10 ml) (containing 0.01 % antibiotic) and shaken on an orbital shaker for 16 h at 37°C. From this culture, aliquots (1 ml) were taken and glycerol (100%) added to a concentration of 50% and the aliquots stored at -80°C.

6.5 Over expression vectors

The *E.coli* expression vectors PET28(a) (novagen) and pHISTev (from Dr. H. Liu, University of St Andrews) were used to over express and purify enzymes used in this thesis. For expression of the PNP (*FlB*) from *S. cattleya*, the pLou vector (from Dr. L.Major, University of St Andrews) was used, encoding a maltose binding protein (*malE*) to help with folding of the expressed protein. They each generate a His_6 tag, enabling nickel column purification, as described below. For degenerate PCR, the pGem™-T Easy vector (Promega) was used in conjunction with GoTaq polymerase (Promega) for preparation for DNA analysis. For expression in *S. lividans*, the pXY200 *E. coli-Streptomyces* shuttle vector was used (see Figure 6.1).

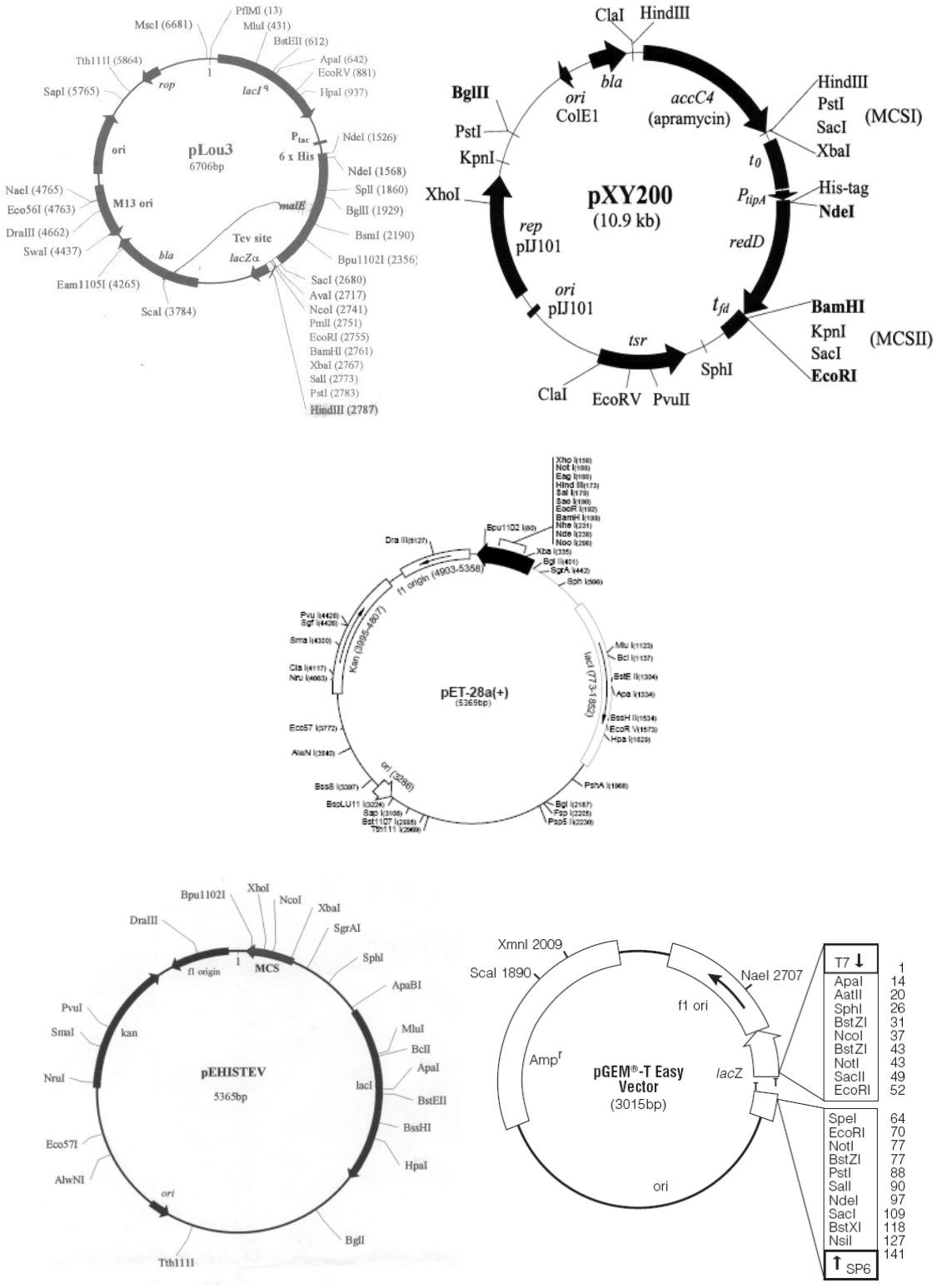


Figure 6.1. Plasmid vectors used for overexpression and DNA analysis in this thesis.

6.6 Nickel column chromatography

Cells (5 g) were resuspended in phosphate buffer (100 mM, pH 7.8), containing imidazole (10 mM), and stirred for 30 min at 4 °C. The suspension was then sonicated, x10 at 60 cycles for 1 min, and the cell debris removed by centrifugation (9,000 rpm for 20 min) and then the supernatant retained as a CFE. A sample of the cell debris was resuspended and retained for SDS page analysis. Purification was carried out on a $NiSO_4$ charged resin.

Sample Buffer (1l): 10 mM Phosphate Buffer pH 7.8
10 mM Imidazole

Loading Buffer (500 ml): 10 mM Phosphate Buffer pH 7.8
30 mM Imidazole
Eluting Buffer (500 ml): 10 mM Phosphate Buffer pH 7.8
500 mM Imidazole

The column was pre-equilibriated with 4 to 5 column volumes of Sample Buffer. The protein sample was loaded and the column washed by 4 to 5 volumes of Loading Buffer to remove endogenous proteins. Elution was carried out with 4 column volumes of Eluting Buffer. Protein concentration was monitored by Nanodrop and SDS PAGE and the His_6-tag was cleaved with thrombin (Sigma T-4648, 1 unit/mg protein) by incubation for 16 h at 4 ºC.

6.7 Fast Performance Liquid Chromatography

6.7.1 Size Exclusion Chromatography

Protein purification was carried out on an ACTA Basic system. Size exclusion chromatography was performed using a High Load 16/60 Superdex 200 column (Amersham Biosciences). The column was equilibriated using phosphate buffer (10mM, pH 7.8) at 1 ml/min for 4 or 5 column volumes. Protein samples were reduced to 2 ml using 10,000 MW microcentrifuge membranes and injected. Elution was monitored at 280 nm at 1ml/min in samples of 4 ml.

6.7.2 Anion exchange chromatography

Protein purification was carried out on an ACTA Basic system. Size exclusion chromatography was performed using a HiTrap™ Q HP 100 ml column (Amersham Biosciences). The column was equilibriated using phosphate buffer (10 mM, pH 7.8) at 1 ml/min for 4 or 5 column volumes. Protein samples were reduced to 2 ml using 10,000 MW microcentrifuge membranes and injected. Elution of the protein was achieved by increasing the buffer B (1 M NaCl) from 0 to 30 % (i.e. 0-0.3 M) over 30 min, monitored at 280 nm at 1 ml/min in samples of 4 ml.

6.7.3 Desalting chromatography

Protein purification was carried out on an ACTA basic system at room temperature. Size exclusion chromatography was performed using a HiTrap™ Q HP 100 ml column (Amersham Biosciences). The column was equilibriated using 10 mM phosphate buffer

pH 7.8 at 1ml/min for 4 or 5 column volumes. Protein samples were reduced to 2 ml using 10,000 MW microcentrifuge membranes and injected. Elution of the protein was monitored at 280 nm at 1 ml/min in samples of 1 ml and salt removal monitored by conductivity.

6.8 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed on an Invitrogen XCell *SureLock*^{TM} mini-cell apparatus connected to an Amersham Pharmacia biotech EPS 301 power supply operating at a constant current of 125mA for 35 min. NuPAGE^{TM} Bis-Tris 10 well gels containing 4-12 % of acrylamide were used.

Protein samples for SDS-PAGE analysis were prepared by adding 5 µl of NuPAGE™ LDS sample buffer (Invitrogen, NP0007) to 20 µl of protein sample, and the protein was denatured at 100 ºC for 3 minutes. A sample (10-20 µl) was then added to the sample wells of the pre-cast NuPAGE^{TM} Bis-Tris gel. Prestained pageruler™ protein ladder (SM0661, Fermentas) was used as a guide for MW determination.

The gel was stained by soaking and was agitated in Coomassie blue G250 dye for 30 min. Destaining was achieved by submerging in water and microwaving for 10 min at full power. Stain solution is composed of coomassie blue G250 (2.0g), methanol (400 ml), glacial acetic acid (70 ml) and ultra pure water (530 ml).

6.9 DNA Gel Electrophoresis

DNA gels were performed on a HOEFER™ HE33 mini horizontal submarine unit (Amersham Biosciences Ltd), using a power pac 300 (Bio Rad), operating at a constant current of 110 V for 30 min. DNA gels were prepared by adding 1 % w/v agarose to 1X TAE solution prepared as described below. The solution was heated until liquid at full power in a microwave, before cooling to 40 ºC in a water bath. 5 µl ethidium bromide was then mixed with ~40 ml of this solution and set into a DNA gel casting tray, using a comb to generate 12 wells. DNA samples were thoroughly mixed with blue/orange loading dye (Promega, cat no. G190A) and Generuler™ 1 kb and 100 bp ladders dyed similarly and used as a standard. DNA gels were analysed under UV light, and bands excised using a scalpel for purification by the SV Wizard™ Gel and PCR Cleanup Kit (Promega) according to manufacturers instructions.

6.9.1 TAE Buffer

For 1 litre of 50 x TAE buffer the following reagents were added:

2M Tris base

1M glacial acetic acid (100%) (57.19 ml = 1 mole)

100 ml 0.5 M Na_2 EDTA (pH 8.0)

H_2O up to 1000 ml

6.10 Polymerase chain reaction (PCR)

PCR reactions were carried out on a TC-512 PCR machine (Techne). DNA primers were designed in-house with appropriate restriction enzyme sites and ordered from MWG Biotech. They were received in a freeze-dried form, and subsequently dissolved into nuclease free water to a final concentration of 100 pM/µl. This stock solution was further diluted to 20 pM/µl, and 1 µl each of forward and reverse primer solutions were used in the PCR reactions. PCR programmes were determined by the length of insert and the DNA polymerase used for amplification. DMSO was added to final concentration of 6 % in all PCR reactions with *Streptomyces* DNA as a template.

6.11 MS-MS Mass Spectrometry

Proteins identified by SDS-PAGE analysis were excised and subject to MS-MS by A. Houston and Dr. C. Botting (University of St Andrews). The identity of excised protein bands was confirmed by in-gel tryptic digest and analysis of the resultant peptides by nanoLC-ESI MSMS (UltiMate (Dionex) and Q-Star Pulsar XL (Applied Biosystems)). The MS/MS data file generated was analysed using the Mascot 2.1 search engine (Matrix Science, London, UK) against an internal database consisting of a bacterial genome background to which the FTase sequence (amongst others) had been added. The data was searched with tolerances of 0.2 Da for the precursor and fragment ions. Trypsin was used as the cleavage enzyme with up to one missed cleavage assumed. Carbamidomethyl modification of cysteines was selected as a fixed modification and L-methionine oxidation as a variable modification.

6.12 GC-MS Mass Spectrometry

GC-MS analysis was carried out by Dr. J. Hamilton (Queen's University, Belfast). Lyophilised samples were per-trimethylsilylated by addition of N-methyl-N-(trimethylsilyl) trifluoroacetamide and heating for 60 min at 100 °C. GC-MS analysis was performed on an Agilent 5890 GC instrument which was directly attached to an Agilent 5973A mass selective detector (MSD). The GC was equipped with an Ultra 1 fused-silica capillary column (Agilent Technologies; 12 m × 0.25 mm × 0.17 μm). The oven temperature was programmed to hold for 1 min at 100 °C and then ramped at 10 °C/min to 300 °C. The injector and transfer line temperatures were set at 250 °C and the per-trimethylsilylated sample (1 μl) automatically injected in the splitless mode. The MSD was operated in the full scan mode measuring ion currents between *m/z* 30 and 500 amu.

6.13 Fluorinase and purine nucleotide phosphorylase expression.

E. coli BL21 (DE3) Gold cells were transformed with the pET28(a) plasmid containing the *FlA* gene. The pLou plasmid construct containing the *FlB* gene was transformed similarly. 20 μl of cell stock containing 50 % glycerol was added to 20 ml LB containing 0.05% (100mg/ml) kanamycin and incubated at 37 °C for 16 h. Aliquots (2.5 ml) were transferred to 2 l flasks containing 750 ml LB and 0.05 % ampicillin and incubated until the solution reached an O.D of 0.6 at 600 nm. The flasks were then cooled to 4 °C before induction by 0.01% IPTG (100 mg/ml) at 16 °C for 16 h. The cells were harvested by

centrifugation at 9,000 rpm for 20 min, and then the supernatant was discarded and the cell pellet either stored at -80°C or used directly for further protein purification.

6.14 PLP dependant transaldolase expression and purification

The plasmid pXY-ScaFTase was transfected into protoplasts of *S. lividans* TK24 by Dr Hai Deng, followed by the standard procedure.[180] The transfected *S lividans* protoplasts were then plated in SFM medium with $MgCl_2$ (10 mM) at 30 ºC for 16 h and flooded with apramycin (1mL, 25 μg/ml). After 3-5 d incubation at 30 ºC, the surviving spores were picked up and grown in 10 ml YEME medium supplied with apramycin (50 μg/ml) at 28 ºC until the spores were observed. Then the medium was incubated with YEME medium (100 ml) supplied with apramycin (50 μg/ml) at 28 ºC for 60 h and the protein was induced by adding thiostreptin (10 μg/m) for another 24 h. The cells were harvested and subject to sonication. The cell-free extract was partially purified by Ni^{2+} chromatography and subjected to SDS-PAGE and MS-MS analysis. The enzyme activity was monitored by incubation with FAld **40** (1 mM), PLP (20 μM) and L-threonine (1 mM) at 37 ºC for 16 h for ^{19}F NMR analysis.

6.14.1 SFM Medium

The following reagents were added in 1 L ultrapure water:
Mannitol 20g
Soya flour 20g
Sterilization by autoclaving

6.14.2 Yeme Medium

The following reagents were added in 1 L ultrapure water:
Yeast extract 3g
Bacto-peptone 5g
Malt extract 3 g
Glucose 10g
Sucrose 340 g
after autoclaving add MgCl2.6H2O to 5 mM

6.15 SCO3014 from *S. coelicolor* and MTRI-Sca from *S. cattleya*

6.15.1 Gene Amplification

Genomic DNA from *S. coelicolor* and *S. cattleya* was prepared as a template for the amplification of *SCO1844* and the MTRI-Sca ORFs in the presence of the primers from Tables 2.1 and 2.3 respectively. PCR reactions were performed in 20 µl of final volume with 6 % DMSO and *pFu* DNA polymerase (1.5 unit, Promega). The PCR reaction was preheated to 98 ºC for 5 min, followed by 30 cycles of denaturation at 95 ºC for 1 min, annealing at 58ºC for 1 min and extension at 72 ºC for 1-2 min dependent on the size of DNA amplification, with 7 min infilling at 72 ºC. The PCR products were subjected to DNA gel analysis. In a 1 % agarose TAE gel, run in TAE buffer at 100 V for 30 min. Gels were then analysed by UV, DNA bands were purified by the SV Wizard Gel Cleanup Kit to ~100 ng/µl. The excised DNA bands were subjected to 4 h digestion by the EcoRI and XhoI restriction enzymes as was the pHISTev vector according to manufacturer's instructions. All of the DNA preparations were then repurified into nuclease-free water using the SV Wizard Gel Cleanup Kit. The final DNA concentrations

were measured by nanodrop. The SCO1844 and MTRI-Sca preparations were individually incubated with the pHISTev preparation in the presence of T4 DNA ligase in a ratio of 3:1 respectively for 16 h at 4 ºC. The ligation mixture was then used to transform competent *E. coli* Bl21(DE3) Gold cells by heat shock. Recombinant plasmids were purified by the QIAPREP™ spin miniprep kit (Qiagen) according to manufacturer's specifications. DNA sequencing was carried out by Dundee University Sequencing Service, and all DNA was prepared according to their requirements.

6.15.2 Protein Overexpression

The resultant plasmids pHISTev-SCO1844 and pHISTev-MTRI-Sca were introduced into *E. coli* BL21 (DE3) Gold (Stratagene) competent cells and grown in Luria broth containing kanamycin (50 µg/ml) at 37 °C until an absorbance of 0.6 at 600 nm was reached. The proteins were over expressed by adding IPTG (1 mM) and cells were left to grow at 16 °C for 16 h. Cells were then harvested by centrifugation and were subject to sonication for lysis. The cell-free extract with PBS and imidazole (10 mM) was then centrifuged (2x, 20,000g) at 4 °C for 15 min. The supernatant was subjected to Ni-affinity chromatography and the active fractions were eluted by adding PBS buffer with imidazole (100 mM). The eluent was dialysed for 16 h at 25 °C by adding thrombin (0.5 unit; Sigma Aldrich Co. Ltd.,) The dialysate was then subjected to size exclusion chromatography (Column, Amersham Co), anion exchange chromatography, followed by desalting and SDS-PAGE, confirmed by MS-MS of the excised SDS-PAGE gel band.

6.15.3 Assays

6.15.3.1 5-FDRP Generation

5-FDRP **39** was generated from synthetic 5'-FDA **35** (prepared by M. Onega and Dr. M. Winkler, University of St Andrews). 5'-FDA **35** was dissolved into phosphate buffer (10 mM, pH 7.8) to a final concentration of 20 mM. Commercially available 5'-adenylic acid deaminase (0.1 mg) was then incubated with this solution for 2 h at 37 ºC. The reaction was stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). A sample of the supernatant was made up to a volume 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis. Following confirmation of 5-FDI **36** generation, the supernatant was incubated with commercially available PNP (0.1 mg) for 16 h at 37 ºC. The sample was then stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). A sample of the supernatant was made up to a final volume of 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis. The supernatant was then removed and stored at -20 ºC until required for the assay.

6.15.3.2 SCO3014 and MTRI-Sca Assay

The purified SCO3014 and MTRI-Sca proteins were incubated with 25 µl of 5-FDRP **38** solution for 6 h at 37 ºC. MTRI-Sca was also preincubated with 1 mM EDTA for 30 min at 37 ºC before incubation with 5-FDRP **38** in a similar manner. Control experiments were also set up in the absence of SCO3014 or MTRI-Sca. All of the above reactions were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min)

a sample of the supernatant was made up to 800 µl with ultrapure water and 100 µl D_2O for ^{19}F NMR for analysis.

6.15.3.3 Reconstituted MTRI-Sca assay

The fluorinase, PNP (*FlB*) and MTRI-Sca were purified as detailed above into phosphate buffer (10 mM, pH 7.8) to a final concentration of ~1 mg/ml. They were added (0.1 mg) into an eppendorf (1.5 ml) in the presence of 2 mM SAM **34**, and 50 mM KF and incubated for 16 h at 37 ºC. Control experiments were also set up in the absence of the MTRI-Sca protein. All of the above reactions were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). A sample of the supernatant was made up to 800 µl with ultrapure water and 100 µl D_2O for ^{19}F NMR analysis.

6.15.3.4 Isothermal titration calorimetry (ITC)

ITC experiments were carried out using a VP-ITC device (microCal, Northampton, MA). The MTRI-Sca protein was purified as before and dialyzed against 10 mM HEPEs buffer (pH 7.8), and the DHAP and L-G3P ligands were dissolved in the same buffer to a final concentration of 600 µM. All solutions were degassed and the ligand solutions were injected at 25 °C into the sample cell containing ~1.4 ml of MTRI-Sca with the concentration around 20 µM. Each titration consisted of an initial injection (1 µl) followed by 25 subsequent injections (5 µl) of the ligands with 180 s intervals.

Calorimetric data was analysed using MicroCal ORIGIN software using a single binding site model.

6.16 Fuculose aldolase

6.16.1 Degenerate PCR of the fuculose aldolase from *S. cattleya*

DNA fragments amplified using a combination of two degenerate primers (Table 3.2) by PCR reactions, were performed in 20 µl of final volume with 6 % DMSO and GoTaq DNA polymerase (1.5 unit, Promega) in the presence of *S. cattleya* genomic DNA as a template. The samples were preheated in 98 ºC for 5 min, followed by 30 cycles of denaturation at 95 ºC for 1 min, annealing at 55 ºC for 1 min and extension at 72 ºC for 1-2 min depending on the size of DNA amplification, with 7 min infilling at 72 ºC. The PCR products were subjected to DNA gel analysis. In a 1 % agarose TAE gel, run in TAE buffer at 100 V for 30 min. Gels were then analysed by UV, DNA bands were purified and ligated into the pGEM-T easy vector and transfected into JLM109 competent cells and plated on agar plates in the presence of X-Gal for blue-white screening detection. Selected colonies were picked from the agar plate and grown in LB media containing ampicillin at 37 ºC for 16 h. The media was then centrifuged, and the resultant cell pellet was used for plasmid extraction, using the QIAprep Spin Miniprep Kit (Qiagen, 27104). The resultant plasmids were subject to DNA sequencing at Dundee University according to their specifications.

6.16.2 Amplification of *SCO1844* from *S. coelicolor* genomic DNA

Genomic DNA from *S. coelicolor* was prepared as a template for the amplification of *SCO1844*. PCR reactions were performed in a final volume of 20 µl, with 6 % DMSO and *pFu* DNA polymerase (1.5 unit, Promega) in the presence of the primers from Table 3.3. The PCR reaction was preheated to 98 ºC for 5 min, followed by 30 cycles of denaturation at 95 ºC for 1 min, annealing at 58 ºC for 1 min and extension at 72 ºC for 1-2 min dependent on the size of DNA amplification, with 7 min infilling at 72ºC. The PCR products were subjected to DNA gel analysis. In a 1 % agarose TAE gel, run in TAE buffer at 100 V for 30 min. Gels were then analysed by UV, DNA bands were purified by the SV Wizard Gel Cleanup Kit to ~80 ng/µl. The ORF SCO1844 was subjected to 4 h digestion by the EcoRI and HindIII restriction enzymes as was the pHISTev vector according to manufacturers instructions. The DNA was then repurified into nuclease-free water using the SV Wizard™ Gel Cleanup Kit and the final DNA concentration was measured by nanodrop. The SCO1844 and pHISTev preparations were then incubated in the presence of T4 DNA ligase in a ratio of 3:1 respectively for 16 h at 4 ºC. The ligation mixture was then used to transform competent *E. coli* Bl21(DE3) Gold cells by heat shock.

6.16.3 Over-expression of the putative fuculose aldolase from *S. coelicolor* in *E. coli*

The resultant plasmid pHISTev-SCO1844 was introduced into *E. coli* BL21 (DE3) Gold (Stratagene) competent cells and grown in LB medium containing kanamycin (50µg/ml) at 37 °C until an absorbance of 0.6 at 600 nm was reached. The SCO1844 protein was

over expressed by adding IPTG (1 mM) and cells were left to grow at 16 °C for 16 h. Cells were then harvested by centrifugation and were subject to sonication for lysis. The cell-free extract with PBS and imidazole (10 mM) was then centrifuged (2 x, 20,000 g) at 4 °C for 15 min. The supernatant was collected by passing through a Ni-affinity column (Qiagen) and the active fractions were eluted by adding PBS buffer with imidazole (100 mM). The eluent was dialysed for 16 h at 25 °C by adding thrombin (0.5 unit; Sigma Ltd) The dialysate was then subjected to size exclusion chromatography. The active fractions gave a monomeric mass of 25 kDa by SDS-PAGE, confirmed by MS-MS of the excised SDS-PAGE gel band.

6.16.4 Assay of the SCO1844 protein

6.16.4.1 Aldol Reaction

The protein product of *SCO1844* was over expressed and purified as described previously to a final concentration of ~1 mg/ml. DHAP was purchased from Sigma and a stock solution in ultrapure water of 20 mM was made. Synthetic FAld **40** was prepared from fluoroethanol according to previous methods (M. Onega, University of St Andrews), to a final concentration of ~20 mM. 0.1 mg of the SCO3014 protein was incubated with DHAP (1 mM) and FAld **40** (~ 5mM) for either 6 h at 37 °C, or 24 h at RT and at 4 ºC. A control experiment without the SCO1844 protein was also set up alongside this assay. The above reactions were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). The supernatant was made up to 800 µl with ultrapure water and 100 µl D_2O for ^{19}F NMR analysis for confirmation of 5-FDRulP **39** generation.

6.16.4.2 Aldol time course reaction

The SCO1844 protein (0.1 mg) was incubated with DHAP (1 mM) and FAld **40** (~ 5mM) for 6 h at 37 °C. Samples (100 µl) were taken at 10 min, 20 min, 30 min, 40 min, 50 min, 1 h, 2 h and 3 h. The samples were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). The supernatant was made up to 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis.

6.16.4.3 Aldol reaction with EDTA incubation

The SCO1844 protein (0.1 mg) was preincubated with EDTA (1 mM) for 30 m at 37 ºC before being incubated with DHAP (1 mM) and FAld **40** (~ 5mM) for 6 h at 37 °C. A control reaction was also set up using the SCO1844 protein without EDTA incubation and the assay solution incubated in a similar manner. The samples were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min) and the supernatant was made up to 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis.

6.16.4.4 Aldol reaction with Zn^{2+} incubation

The SCO1844 protein (0.1 mg) was preincubated with Zn^{2+} at different concentrations (10, 20 and 100 µM) before being incubated with DHAP (1 mM) and FAld **40** (~ 5mM) for 6 h at 37 °C. A control reaction was also set up using the SCO1844 protein without Zn^{2+} and incubated in a similar manner. The samples were stopped by heat deactivation

(95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). The supernatant was made up to 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis.

6.16.4.5 Reconstituted SCO1844 assay: Retro-aldol reaction

The fluorinase, PNP (*FlB*) and MTRI-Sca and SCO1844 proteins were purified as detailed above in phosphate buffer (10 mM, pH 7.8) to a final concentration of ~1 mg/ml. They were added (0.1 mg) into an eppendorf (1.5 ml) in the presence of 2 mM SAM **34**, and 50 mM KF and incubated for 16 h at 37 ºC. Control experiments were also set up in the absence of the SCO1844 protein. All of the above reactions were stopped by heat deactivation (95 ºC, 5 min) and centrifugation (12,000 rpm, 2 min). A sample of the supernatant was made up to 800 µl with ultrapure water and D_2O (100 µl) for ^{19}F NMR analysis.

6.17 *In vitro* reconstitution of FAc 8 from inorganic fluoride ion

The fluorinase, PNP, isomerase and fuculose aldolase were all over-expressed in *E. coli* and purified to homogeneity by Ni-affinity and size exclusion chromatography to ~1 mg/ml in PBS buffer. The aldehyde dehydrogenase and its cofactor, NAD(P) were purchased and dissolved into PBS buffer. All of the pathway enzymes were added into an eppendorf tube (1.5 ml) to a final concentration of 0.1 mg/ml. They were incubated with SAM (1.4 mM), KF (35 mM) and $NAD(P)^+$ (1 mM, Sigma Ltd) for 24 h at 37 °C. Samples were removed after 0, 1, 2, 3, 4, 5, 6 and 24 h. The samples were stopped by heat inactivation (95 °C, 5 min) followed by centrifugation (2 min, 14,000 g), the

supernatant was then made up to 700 µl and D_2O (100 µl) was added. Control experiments were carried out by removing the aldehyde dehydrogenase from the reaction (enzymes and substrates/cofactors) replacing with an equivalent volume of PBS, whilst subject to the same conditions and analysis. FAc **8** production was confirmed by $^{19}F\{^1H\}$ NMR, by add-mixing with a synthetic reference compound (Sigma Ltd).

6.17.1 PBS Buffer

In 1 L ultrapure H_2O:
8 g NaCl
0.2 g KCl
1.44 g Na_2HPO_4
0.24 g KH_2PO_4

Adjusted pH to 7.5 with HCl
Sterilized by autoclaving.

6.18 *In vitro* reconstitution of 4-FT from inorganic fluoride ion

The fluorinase, PNP, isomerase and fuculose aldolase were all over-expressed in *E. coli* and purified to homogeneity by Ni-affinity and size exclusion chromatography to ~1 mg/ml in PBS buffer. The PLP transaldolase was purified as previously described to approximately 0.25 mg/ml in PBS buffer. All of the pathway enzymes were added into an eppendorf tube (1.5 ml) to a final concentration of 0.1 mg/ml. They were incubated with SAM (1.4 mM), KF (35 mM), PLP (0.7 mM) and L-threonine (35 mM) for 16 h at 37 °C. The reaction was stopped by heat inactivation at 95 °C for 5 min followed by centrifugation of 2 min at 14,000 g, and then the supernatant was then made up to 700 µl

and 100 µl D_2O was added. The resultant mixture was subject to ^{19}F NMR analysis. Control experiments were carried out by removing a single component of the reaction (enzymes and substrates/cofactors) in turn and replacing with an equivalent volume of PBS and subject to the same conditions and analysis. 4-FT **33** production was confirmed by $^{19}F\{^{1}H\}$ NMR, ^{19}F NMR and GC-MS after lyophilisation of the samples.

6.19 Solid Phase Assay

An aliquot (10 µl) of stock solution of *E. coli* (DE3) transformed with pET28(a) containing the fluorinase gene was added to 1 ml LB broth containing kanamycin (0.01%). The aliquot was then shaken at 250 rpm at 37 °C for 2 h, and a sample (5 µl) was added to an LB agar plate containing kanamycin and IPTG (1 mM). Control experiments were carried out without IPTG added to the growth medium. The cells were spread evenly using a plate spreader and then incubated at 37 °C for 16 h. After this period, filter paper discs were placed over the surface of the plates, and pressure was applied from above in order that colonies on the plate were transferred faithfully to the filter paper. The filter paper was then removed and placed in liquid nitrogen for 10-15 s, and then allowed to thaw at RT. A second filter paper disc was then placed in a Petri dish and soaked in 1 ml wash solution (see below), and excess liquid was removed. The filter paper disc containing the cell colonies was then placed on top of the filter paper and wash solution ensuring that no bubbles existed between the two discs. The Petri dish was then covered, and placed in a container with a damp cloth to ensure moisture retention, and then put in an incubator at 37 °C for 90 min. At the same time, the assay mixture minus

diaminobenzidine (DAB) was prepared, and left at room temperature until the next step. After 90 min DAB (the assay mix was also prepared using SIGMA *FAST*™ DAB with $CoCl_2$ (Sigma D-0426)) was added to the assay mixture and a separate filter paper disc was placed in another Petri dish, which was then soaked with 1 ml assay mixture, and the excess removed. The filter paper containing cell colonies was then removed from the wash filter paper, and transferred to the assay mixture-containing filter paper. Again it was ensured that no bubbles existed between the two sheets of filter paper, and that cell colonies were facing up. The assay Petri dishes were again placed in a container with a damp cloth, and incubated at 37 °C and then monitored at regular 15 min intervals for colour change.

6.19.1 Wash Solution (to make 4 mls):

ultra pure water (2 ml)
snake venom l-amino acid oxidase (Sigma A-9253) (1 mg/ml) (1 ml)
horseradish Peroxidase (Fluka 77335) (1 mg/ml) (1 ml)
potassium fluoride (BDH Laboratory Reagents 29613) (500 mM) (1 ml)

6.19.2 Assay Mixture (to make 4 mls):

snake venom l-amino acid oxidase (1 mg/ml) (0.5 ml)
horseradish peroxidase (1 mg/ml) (0.5 ml)
potassium fluoride (500 mM) (1.5 ml)
s-adenosyl-l methionine (Sigma A-7007) (20 mM) (0.5 ml)
SIGMA *FAST*™ 3,3'-diaminobenzidine (Sigma D-4418) (0.5 ml)

6.19.3 Solid phase agar assay with L-methionine controls

All enzymes and DAB stock solutions were prepared as above. A 1 mM L-methionine stock solution was also prepared using ultrapure water. A 2 % agar solution was prepared using ultrapure water and the resulting mixture was heated at full power in a microwave

until the solution was homogenous. The solution was then cooled to 40 ºC in a water bath and 25 ml was removed into a 50 ml falcon tube. To this mixture, LAAO (1 ml, 5 mg/ml), HRP (1 ml, 5 mg/ml) and DAB were added to the solution stirred at 40 ºC. The resulting mixture was then poured into a petri dish and cooled until solid at RT. Filter paper discs were made by a hole punch and soaked into the L-methionine stock solution. They were then dried in open air (10 min). The discs were then placed onto the solidified assay plate, and incubated for 30 min at 37 ºC. DAB oxidation was then analysed visually.

6.20 Liquid phase assay

The fluorinase enzyme was purified as previously described (Section 6.13) and concentrated to 9.93 mg/ ml in phosphate buffer (10 mM, pH 7.8). LAAO (1 mg/ml), HRP (1 mg/ml), DAB, SAM (20 mM) and KF (500 mM) solutions were set up as previously described. Fluorinase was added to a final concentration of ~0.6 mg/ml, LAAO and HRP to 0.1 mg/ml, KF to 20 mM and 40 µl of DAB solution was incubated with varying concentrations of SAM **34** (0, 0.25, 0.5, 1 and 2 mM) and made up to 1 ml using ultrapure water in microcuvettes. The optical density was measured at 480 nm in a spectrophotometer over a period of 90 min, with a reading taken automatically every 10 min. The concentration of L-methionine was determined using a standard curve (see Figure 5.2) by correlating the extinction co-efficient of DAB (5,500 M^{-1}). The steady-state kinetic parameters were obtained by fitting the initial velocity against the substrate concentrations according to the Michaelis-Menten equation. Liquid assay samples (20 µl) were taken for HPLC analysis at 30, 60 and 90 min. Samples were subjected to

denaturing conditions of 95 ºC for 3 min, then spun at 14,000 rpm for 3 min. Aliquots (10 µl) of the supernatant were then resuspended in ultra pure water to make 100 µl, which was then subjected to HPLC analysis. 5'-FDA **35** concentrations were determined by reading from the standard curve, which was created by injecting known concentrations of synthetic 5'-FDA **35** into the HPLC machine.

6.20.1 L-Methionine assays

LAAO (1 mg/ml), HRP (1 mg/ml) and DAB, solutions were prepared as previously described. L-methionine (Sigma) was suspended to a final concentration of 100 mM in ultrapure water. LAAO and HRP were added to a final concentration of 0.1 mg/ml and 40 µl of DAB solution was incubated with varying concentrations of L-methionine (0, 0.125, 0.25, 0.5, 1 and 2 mM) and made up to 1 ml using ultrapure water in microcuvettes. The optical density was measured at 480 nm in a spectrophotometer after 10 and 30 min in a spectrophotometer at 480 nm.

References

[1] F. E. Koehn, G. T. Carter. (2005) *Nat. Rev. Drug. Discov.*, **4**, 206-220.

[2] T. L. Simmons, E. Adrianasolo, K. McPhail, P. Flatt, W. H. Gerwick. (2005) *Mol Cancer Ther.*, **4**, 333-342.

[3] D. J. Newman, G. M. Cragg, K. M. Snader. (2003) *J. Nat. Prod.*, **66**, 1002-1037.

[4] Alan Crozier. (2006) "*Plant secondary metabolites in diet and health*", 1st Ed., Blackwell publishing, pp 1.

[5] A. Rahman, P. L. Quesne, Ed., (1988) "*Natural Products Chemistry*", Vol. 3, Springer Verlag, Berlin, pp 10.

[6] D. Schwarzer, R. Finking, M. A. Marahiel. (2003) *Nat. Prod. Rep.*, **20**, 275–287.

[7] E. Van Geldre, A. Vergauwe, E. Van der Eeckhout. (1997) *Plant Mol. Biol.*, **33**, 199-209.

[8] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggin, A. T. McPhail. (1971) *J. Am. Chem. Soc.*, **93**, 2325-2327.

[9] M. Hayashi. (1977) *Folia Pharmacol. Japon.*, **73**, 205-214.

[10] M. E. Wall, M. C. Wani. (1995) *Cancer Res.*, **55**, 753-760.

[11] G. W. Gribble. (2004) *J. Chem. Educ.*, **81**, 1441-1449.

[12] C. S. Neumann, D. G. Fujimori, C. T. Walsh. (2008) *Chem & Biol.*, **15**, 99-109.

[13] Y. Kan, B. Sakamato, T. Fujita, H. Nagai. (2000) *J. Nat. Prod.*, **63**. 1599-1602.

[14] N. Sitachitta, B. L. Marquez, R. T. Williamson, J. Rossi, M. A. Roberts. W. H. Gerwick, V. A. Nguyen. C. L. Wills. (2000) *Tetrahedron*, **56**, 9103-9113.

[15] C. M. Harris, R. Kannan, H. Kopecka, T. M. Harris. (1985) *J. Am. Chem. Soc.*, **107**, 6652-6658.

[16] E. Rodrigues Pereira, L. Belin, M. Sancelme, M. Prudhomme, M. Ollier. M. Rapp, D. Severe, J. F. Riou, D. Fabbro, T. Meyer. (1996) *J. Med. Chem.*, **39**, 4471-4477.

[17] P. D. Shaw, L. P. Hager. (1959) *J. Biol. Chem.*, **234**, 2565-2569.

[18] L. P. Hager, D. R. Morris, F. S. Brown, H. Eberwein. 1966, *J. Biol. Chem.*, **241**, 1769-1777.

[19] S. L. Niedleman, J. Geigert. (1986) "*Biohalogenation: Principles, basic roles and applications*.", p 46, Chicester: Ellis Horwood.

[20] H. Vilter. 1984, *Phytochemistry*, **23**, 1387-1390.

[21] A. Butler, J. N. Carter-Franklin. (2004) *Nat. Prod. Rep.*, **21**, 180-188.

[22] J. M. Winter, M. C. Moffitt, E. Zazopoulos, J. B. McAlpine, P. C. Dorrestein, B. S. Moore. (2007) *J. Biol. Chem.*, **282**, 16362-16368.

[23] K. –H. van Peé. (1996) *Ann. Rev. Microbiol.*, **50**, 375-399.

[24] M. Picard, J. Gross, E. Lubbert, S. Tolzer, S. Krauss, K. –H. van Peé, A. Berkessel. (1997) *Angew. Chem. Int. Ed.*, **36**, 1196-1199.

[25] B. Hofmann, S. Tolzer, I. Pelletier, J. Altenbuchner. K. –H. van Peé, A. Berkessel. (1998) *J. Mol. Biol.*, **279**, 889-900.

[26] T. Dairi, T. Nakano, K. Aisaka, R. Katsumata, M. Hasegawa. (1995) *Biosci. Biotech. Biochem.*, **59**, 1099-1106.

[27] S. Kirner, P.E. Hammer, D. S. Hill, A. Altmann, I. Weislo, L. J. Lanahan, K. –H. van Peé, J. M. Ligon. (1998) *J. Bacteriol.*, **180**, 1939-1943.

[28] S. Keller, T. Wage, K. Hohaus, M. Holzer, E. Eichhorn, K. –H. van Peé. (2000) *Angew. Chem. Int. Ed.*, **39**, 2300-2302.

[29] D. H. Williams, B. Bardsley. (1999) *Angew. Chem. Int. Ed.*, **38**, 1172-1193.

[30] J. Ahlert, E. Shepard, N. Lomovskaya, E. Zazopoulos, A. Staffa, B. O. Bachmann, K. Huang, L. Fonstein, A. Czisny, R. E. Whitwam, C. M. Farnet, J. S. Thorson. (2002) *Science*. **297**, 1173-1176.

[31] O. Puk, P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R. D. Submuth, K. –H. van Peé, W. Wohlleben, S. Pelzer. (2002) *Chem. Biol.*, **9**, 225-235.

[32] P. C. Dorrestain. E. Yeh, S. Garneau-Tsodikova, N. L. Kelleher, C. T. Walsh. (2005) *Proc. Natl. Acad. Sci. U.S.A*, **102**, 13843-13848.

[33] C. J. Dong, S. Flecks, S. Unversucht, C. Haupt, K-H. van Peé, J. H. Naismith. (2005) *Science*, **309**, 2216-2219.

[34] E. Yeh, S. Garneau, C. T. Walsh. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 3960-3965.

[35] E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan, C.T. Walsh. (2007) *Biochemistry*, **46**, 1284-1292.

[36] S. Zehner, A. Kotzsch, B. Bister, R. D. Sussmuth, C. Mendez, J. A. Salas, K. –H. van Peé. (2005) *Chem. Biol.*, **12**, 445-452.

[37] K. –H. van Peé, S. Zehner. (2003) "*Enzymology and molecular genetics of biological halogenation*." In Handbook of Environmental Chemistry, Vol.3, pp 171-199, Springer: G. Gribble.

[38] J. Orjala, W. H. Gerwick. (1996) *J. Nat. Prod.*, **59**, 427-430.

[39] A. D. Argoudelis. R. R. Herr, D. J. Mason, T. R. Pyke, J. F. Zieserl. (1967) *Biochemistry*, **6**, 165-170.

[40] D. P. Galonic, F. H. Vaillancourt, C. T. Walsh. (2006) *J. Am. Chem. Soc.*, **128**, 3900-3901.

[41] F. H. Vaillancourt, J. Yin, C. T. Walsh. (2005) *Proc. Natl. Acad. Sci USA*, **102**, 10111-10116.

[42] M. Ueki, D. P. Galonic, F. H. Vaillancourt, S. Garneau-Tsodikova, E. Yeh, D. A. Vosburg, F. C. Schroeder, H. Osada, C. T. Walsh. (2006) *Chem. Biol.*, **13**, 1183-1191.

[43] J. M. Bollinger Jr, J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs. (2005) *Eur. J. Inorg. Chem.*, 4245-4254.

[44] L. M. Hoffart, E. W. Barr, R. B. Guyer, J. M. Bollinger Jr, C. Krebs. (2006) *Proc. Natl, Acad. Sci. USA*, 14738-14743.

[45] L. Que Jr. (2000) *Nat. Struct. Biol.*, **7**, 182-184.

[46] M. Costas, M. P. Mehn, M. P. Jensen, L. Que Jr. (2004) *Chem. Rev.*, **104**, 939-986.

[47] R. P. Hausinger. (2004) *Crit. Rev. Biochem. Mol. Biol.*, **39**, 21-68.

[48] L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh, C. L. Drennan. (2006) *Nature*, **440**, 368-371.

[49] D. P. Galonic, E. W. Barr, C. T. Walsh, J. M. Bollinger, C. Krebs. (2007) *Nat. Chem. Biol.*, **3**, 113-116.

[50] H. J. M. Bowen. (1966) "*Trace elements in biochemistry*", Academic Press, London and New York, p10.

[51] K. –H. van Peé. (1996) *Annu. Rev. Microbiol.*, **50**, 375-399.

[52] C. D. Murphy. 2003, *J. Appl. Microbiol.*, **94**, 539-548.

[53] H. Deng. D. O'Hagan, C. Schaffrath. 2004, *Nat. Prod. Rep.*, **21**, 773-784.

[54] B. E. Smart. (1986) *'Molecular structure and energetics'*, Ed., J. L. Leibman and A. Greenberg, VCH Publishers Inc., Deerfield Park, Florida, Vol 3, p. 141.

[55] C. T. Walsh. (1982) *Adv. in Enzymology*, **55**, 197-225.

[56] J. S. C. Marais, J. Onderstepoort. (1944) *Vet. Sci. Anim. Ind*, **20**, 67-73.

[57] B. Vickery. M. L. Vickery. J. T Ashu. (1973) *Phytochemistry*, **12**, 145-147.

[58] D. O' Hagan, R. Perry, J. M. Lock, J. J. M. Meyer, L. Dasaradhi, J. T. G. Hamilton, D. B. Harper. (1993) *Phytochemistry*, **33**, 1043-1045.

[59] B. Vickery. M. L. Vickery. (1972) *Phytochemistry*, **11**, 1905-1909.

[60] H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert., C. D. Stout. (1996) *Proc. Natl. Acad. Sci USA*, **93**, 13699-13703.

[61] R. A. Peters, R. W. Wakelin, P. Buffa, L. C. Thomas. (1953) *Proc. Roy. Soc. B.*, **40**, 497-507.

[62] R. A. Peters, M. Shorthouse. (1967) *Nature*, **216**, 80-81.

[63] R. A. Peters, M. Shorthouse. (1971) *Nature*, **231**, 123-124.

[64] R. A.Peters, M. Shorthouse. (1959) *Biochem. Pharmacol.*, **2**, 25-36.

[65] J. T. G. Hamilton, D. B. Harper. (1997) *Phytochemistry*, **44**, 1129-1132.

[66] X. -H. Xu, G. –M. Yao, Y. –M. Li, J. -H. Lu, C. -J. Lin, X. Wang, C. –H. Kong. (2003) *J. Nat. Prod*, **66**, 285-288.

[67] R. Ignoffo. (1999) *J. Am. Health-Syst. Pharm.*, **56**, 2417-2430.

[68] G. O.Morton, J. E. Lancaster, G. E. Van Lear, W. Fulmor, W. E. Meyer. (1969) *J. Am. Chem. Soc.*, **91**, 1535-1537.

[69] J. M. Williamson, E. Inamine, K. E. Wilson, A. W. Douglas, J. M. Liesch, G. Albers-Schönberg. (1985) *J. Biol. Chem*., **260**, 4637-4647.

[70] M. Sanada, T. Miyano, S. Iwadare, J. M.Williamson, B. H. Arison, J. L. Smith, A.W. Douglas, J. M. Liesch, E. Inamine. (1986) *J. Antibiotics*, **39**, 259-265.

[71] T. C. Chou, P. Talalay. (1972), *Biochem*, **11**, 1065-1073.

[72] K. A. Reid, J. T. G. Hamilton, R. D. Bowden, D. O'Hagan, L. Dasaradhi, M. R. Amin, D. B. Harper. (1995) *Microbiology*, **141**, 1385-1393.

[73] T. C. Chou, P. Talalay. (1972) *Biochem.*, **11**, 1065-1073.

[74] S. L. Cobb, H. Deng, J. T. G. Hamilton, R. P. McGlinchey, D. O'Hagan, C. Schaffrath. (2005) *Bioorganic Chem*., **33**. 393-401.

[75] M.Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A.W. Douglas, J. M. Liesch., E. Inamine. (1986) *J. Antibiotics*, **39**, 304-305.

[76] C. Schaffrath, H. Deng, D. O'Hagan. (2003) *FEBS Letters*, **547**, 111-114.

[77] H. M. Senn, D. O'Hagan, W. Thiel. (2005) *J. Am. Chem. Soc*, **127**, 13643-13655.

[78] P. Kirsch. (2004) 'Modern fluoroorganic chemistry, synthesis, reactivity, applications', Wiley-VCH, Weinheim, ISBN 3-527-30691-9.

[79] R. D. Chambers. (2004), '*Fluorine in organic chemistry*', Blackwell Publishing Ltd, ISBN 1-4051-0787-1.

[80] D. L. Zechel, S. P. Reid, O. Nashiru, C. Mayer, D. Stoll, D. L. Jakeman, R. A. J. Warren, S. G. Withers. (2001) *J. Am Chem. Soc.*, **123**, 4350-4351.

[81] O. Nashiru, D. L. Zechel, D. Stoll, T. Mohammadzadeh, R. A. J. Warren, S. G. Withers. (2001) *Angew. Chem., Int. Ed.*, **40**, 417-420.

[82] C. J. Dong, F. L. Huang., H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan, J. H. Naismith. (2004) *Nature*, **427**, 561-565.

[83] F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T. L. Li, D. O'Hagan, P. F. Leadley, J. B. Spencer. (2006) *Chem & Biol.*, **13**, 475-484.

[84] C. Dong, H. Deng, M. Dorward, C. Schaffrath, D. O'Hagan, J. H. Naismith. (2003) *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **59**, 2292-2293.

[85] M. A. Vincent, I. H. Hillier. (2006) *Chem. Commun.*, 5902-5903.

[86] C. D. Cadicamo, J. Courtieu, H. Deng, A. Meddour, D. O' Hagan. (2004) *ChemBioChem.*, **5**, 685-680.

[87] X. Zhu, D. A. Robinson, A. R. McEwan, D. O'Hagan, J. H. Naismith. (2007) *J. Am. Chem. Soc.*, **129**, 14597-14604.

[88] D. W. Udwary, L. Zeigler, R. N. Asolkar, V. Singan, A. Lapidus, W. Fenical, P. R. Jensen, B. S. Moore. (2007) *Natl Acad Sci USA*, **104**, 10376-10381.

[89] A. Eustáquio, F. Pojer, J. P. Noel, B. S. Moore. (2008) *Nature Chem Biol.*, **4**: 69-74.

[90] H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson, J. B. Spencer. (2006) *Angew Chem Int Ed Engl.*, **45**, 759-762.

[91] H. Deng, C. H. Botting, J. T. G. Hamilton, R. J. M. Russell, D. O'Hagan. (2008) *Angew. Chem. Int. Ed.*, **47**, 5357-5361.

[92] M. J. Pugmire, S. E. Ealick. (2002) *Biochem. J.*, **361**, 1-25.

[93] S. L. Cobb, H. Deng, J. T. G Hamilton., R. P.McGlinchey, D. O'Hagan. (2004) *Chem. Commun.* 592-594.

[94] E. M. Bennett, C. Li, P.W. Allen, W. B. Parker, S. E. Ealick. (2003) *J. Biol. Chem.*, **278**, 47110-47118.

[95] R. P. McGlinchey. (2006) '*Intermediates and enzymes involved in fluorometabolite biosynthesis in Streptomyces cattleya'*. Ph.D thesis, University of St Andrews.

[96] C. D. Murphy, D. O'Hagan and C. Schaffrath. (2001) *Angew. Chem., Int. Ed.*, **40**, 4479-4481.

[97] M. Onega., R. P. McGlinchey, H. Deng, J. T. G. Hamilton, D. O'Hagan. (2007) *Bioorg Chem*. **35**, 375-385.

[98] R. B. Herbert, B. Wilkinson, G. J. Ellames E. K. Kunec. (1993) *J. Chem. Soc., Chem. Commun.*, 205-206.

[99] H. Deng, S. M. Cross, R. P. McGlinchey, J. T. G. Hamilton, D. O'Hagan. (2008) *Chem & Biol*, in press.

[100] Y. Luo, J. Samuel, S.C. Mosimann, J.E. Lee, M.E. Tanner, N.C.J. Strynadka. (2001) *Biochemistry*. **40**, 14763-14771.

[101] L.V. Lee, M.V. Vu, W. W. Cleland. (2000) *Biochemistry*, **39**, 4808-4820.

[102] C. D. Murphy, S. J. Moss, D. O'Hagan. (2001), *Appl. Environ. Microbiol.*, **67**, 4919-4921.

[103] S. Boubekeur, N. Camougrand, O. Bunoust, M. Rigoulet and B. Guerin. (2001) *Eur. J. Biochem.*, **268**, 5057-5065.

[104] J. Czernin, H. R. Schelbert, D. H. S. Silverman, W. P. Melega. (2004) '*PET: Molecular imaging and its biological applications*', Ed. M. E. Phelps, Springer-Verlag, Berlin-Heidelberg-New York, pp. 321-584.

[105] L. Cai, S. Lu, V. W. Pike. (2008) *Eur. J. Org. Chem.*, 2853-2873.

[106] M. R. Kilbourn. (1990) '*Fluorine-18 labelling of radiopharmaceuticals'*, National Academy Press: Washington, DC.

[107] K. Ishiwata, Y. Kimura, E. F. de Vries, P. H. Elsinga. (2007) *Cent. Nerv. Syst. Agents. Med. Chem.*, **7**, 57-77.

[108] T. Kimura, I. K. Ho, I. Yamamoto. (2001) *Sleep*, **24**, 251-260.

[109] C. G. Kim, D. J. Yang, E. E. Kim, A. Chrif, L. R. Kuang, C. Li, W, Tansey, C. W. Lui, S. C. Li, S. Wallace, D. A. Podoloff. (1996) *J. Pharmaceutical Sci.*, **85**, 339-344.

[110] C. Prenant, J. Gillies, J. Bailey, G. Chimon, N. Smith, G. C. Jayson, J. Zweitb. (2008) *J. Label Compd. Radiopharm*., **51**, 262–267.

[111] Sz. Lehel, G. Horvath, I. Boros, T. Marian, L. Tron. (2000) *J. Radioanalyst Nucl, Chem.*, **245**, 399-401.

[112] L. Marterello, C. Schaffrath, H. Deng, A. D. Gee, A. Lockhart, D. O'Hagan. (2003) *J. Label Compd. Radiopharm.*, **46**, 1181-1189.

[113] M. Winkler, J. Domarkas, L. F. Schweiger, D. O'Hagan. (2008) *Angew. Chemie. Int. Ed.*, **47**, 1-4.

[114] I. Jelesarov, H. R. Bosshard. (1999) *J. Mol. Recognit.*, **12**, 2-18.

[115] E. Freire. 2004, *Drug Discovery Today*, **1**(3), 295-299.

[116] E. S Furfine, R. H Abeles. (1988) *J. Biol. Chem.* **263**. 9598-9606

[117] R. W. Myers, J. W. Wray, S. Fish, R. H. Abeles. (1993) *J. Biol. Chem*. **268**, 24785-24791

[118] J. W. Wray, R. H. Abeles. (1995) *J. Biol. Chem.* **270**, 3147-3153.

[119] J. Heilbronn, J. Wilson, B. J. Berger. (1999) *J. Bacteriol.* **181**, 1739-1747.

[120] Y. Dai, T. C. Pochapsky, R. H. Abeles. (2001) *Biochemistry*, **40**, 6379-6387.

[121] A. Sekowska, A. Danchin. (2002) *BMC Microbiol.*, **2**, 1-14.

[122] H. Ashida, Y. Saito, C. Kojima, K. Kobayashi, N. Ogasawara, and A.Yokota. (2003) *Science*, **302,** 286–290.

[123] M. Bumann, S. Djafarzadeh, A. E. Oberholzer, P. Bigler, M. Altmann, H. Trachsel. U. Baumann. (2004) *J. Biol. Chem.*, **279**, 37087-37094.

[124] A. Bateman, L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. L. Sonnhammer, D. J. Studholme, C. Yeats. S. R. Eddy. (2004) *Nucleic Acids Res.*, **32**, 138-141.

[125] H. Tamura, Y. Saito, H. Ashida, T. Inoue, Y. Kai, A. Yokota, H. Matsumura. (2008) *Protein Science*. **17**. 126-135.

[126] V.K.Rastogi, M.E.Girvin. (1999) *Nature*. **402** (263). 263-268.

[127] E. S. Rangarajan, J. Sivaraman, A. Matte, M. Cygler. (2002) *Proteins*, ***48***, 737-740.

[128] G. Avigad. (1975) *Methods Enzymol.*, **41**, 27-29.

[129] S. C. Winans, B. L. Bassler. (2002) *J. Bacteriol*. **184**, 873-883.

[130] I. A. Rose. (1975) *Adv. Enzymol, Relat. Areas. Mol. Biol.*, **43**, 491-517.

[131] Y. Saito, H. Ashida, C. Kojima, H. Tamura, H. Matsumura, Y. Kai, A. Yokota. (2002) *Biosci Biotechnol Biochem*., **71**, 2021-2028.

[132] T. D. Fenn, D. Ringe, G. A. Petsko. (2004) *Biochemistry*, **43**, 6464-6474.

[133] D. M. Blow, C. A. Collyer, J. D. Goldberg, O. S. Smart. (1992) *Faraday dis.*, 67-73.

[134] K. Langsetmo, J.A. Fuchs, C. Woodward. (1991) *Biochemistry*. **30**. 7603-7609.

[135] S. D. Bentley , K. F. Chater, A. M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T.

Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood. (2002) *Nature.*, **9**, 141-147.

[136] G.M. Whitesides and C- H. Wong. (1985) *Angew. Chem. Int. Ed*., **24**, 617-638.

[137] E. J. Toone, E. S. Simon, M. D. Bednarski, G. M. Whitesides. (1989) *Tetrahedron*, **45**, 5365-5422.

[138] P. G. Wang, W. Fitz and C. –H. Wong. (1995) *CHEMTECH*, **25**, 22-32.

[139] H. J. M. Gijsen, L. Qiao, W. Fitz, C. –H. Wong. (1996) *Chem. Rev.*, **96**, 443-474.

[140] W. –D. Fessner, C. Walter. (1992) *Angew. Chem. Int. Ed.*, **31**, 614-616.

[141] R. Schoevaart, PhD thesis. (2000) '*Applications of aldolases in organic synthesis*', Delft University of Technology.

[142] R. Schoevaart, F. van Rantwijk, R. Sheldon. (2000) *J. Org. Chem.*, **65**, 6940-6943.

[143] W. J. Rutter. (1964) *Fed Proc*, **23**, 1248-1257.

[144] T. Gefflaut, C. Blonski, J. Perie, M. Willson. (1995) *Prog. Biophys. Molec. Biol.*, **63**, 301-340.

[145] G. J. Thomson, G. J. Howlett, A. E Ashcroft, A. Berry. (1998) *Biochem. J.*, **331**, 437-445.

[146] E. Lorentzen, B. Siebers, R. Hensel, E. Pohl. (2005) *Biochemistry*, **44**, 4222-4229.

[147] A. C. Joerger, C. Gosse, W. D. Fessner, G. E. Schulz. (2000) *Biochemistry*, **39**, 6033-6041.

[148] W. D. Fessner, A. Schneider, H. Held, G. Sinerius, C. Walter, M. Hixon, J. V. Schloss. (1996) *Angew. Chem. Int. Ed.*, **35**, 2219-2221.

[149] S. Omura, H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki , M. Hattori. (2001) *Proc. Natl. Acad. Sci. U S A*, **98**, 12215-12220.

[150] Y. Nakamura, T. Gojobori, T. Ikemura. (2000) *Nucl. Acids. Res.*, **28**, 292.

[151] X. Cheng, , J. Liu, M. Asuncion-Chin, E. Blaskova, J. P. Bannister, A. M. Dopico, J. H. Jaggar. (2007) *J. Biol. Chem.*, **282**, 29211–29221.

[152] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, A. Bairoch. (2003) *Nucleic Acids Res.*, **31**, 3784-3788.

[153] S. J. Moss., C. D. Murphy., D. O'Hagan, C. Schaffrath., J. T. G. Hamilton, W. C. McRoberts., D. B. Harper. (2000) *Chem. Commun.*, 2281-2282.

[154] W. –D. Fessner, G. Sinerius, A. Schneider, M. Dreyer, J. Schulz, J. Badia, J. Aquilar. (1991) *Angew. Chem.*, **103**, 596-599.

[155] A. Dalby, Z. Dauter, J. A. Littlechild. (1999) *Protein Sci.*, **8**, 291-297.

[156] D. R. Hall, G. A. Leonard, C. D. Reed, C. I. Watt, A. Berry, W. N. Hunter. (1999) *J. Mol. Biol.* **287**. 383-394.

[157] A. R. Plater, S. M. Zgiby, G. J. Thomson, S. Qamar, C. W. Wharton., A. Berry. (1999) *J. Mol. Biol.*, **285**, 843-855.

[158] T. Suau, G. Alvaro, M. Dolors Benaiges, J. Lopez-Santin. (2006) *Biotechnol Bioeng*, **93**, 48-55.

[159] A. Ozaki, E. J. Toone, C. H. von der Osten, A. J. Sinskey, G. M. Whitesides. (1990) *J. Am. Chem. Soc.*, **112**, 4970-4971.

[160] E. Garcia-Junceda, G-J. Shen, T. Sugai, C-H. Wong. (1995) *Bioorg. Med. Chem.*, **3**, 945-953.

[161] I. Ardao, M. Dolors Benaiges, G. Caminal, G. Alvaro. (2006) *Enzyme and Microbial Technology*., **39**, 22-27.

[162] Q. Cheng, L. Xiang, M. Izumikawa, D. Meluzzi, B. S. Moore. (2007) *Nature Chem. Biol.*, **3**, 557-558.

[163] M. C. Cone, X. Yin, L. L. Grochowski, M. R. Parker, T. M. Zabriskie. (2003) *ChemBioChem*., **4**, 821-828.

[164] M. T. Reetz, C. Torre, A. Eipper, R. Lohmer, M. Hermes, B. Brunner, A. Maichele, M. Bocola, M. Arand, A. Cronin, Y. Genzel, A. Archelas, R. Furstoss. (2004) *Org Lett*, **22**, 177-180.

[165] M. T. Reetz. (2005) *Tetrahedron*, **58**, 6595-6602.

[166] A. Faust, K. Niefind, W. Hummel, D. Schomburg. (2007) *J. Mol. Biol.*, **367**, 234-248.

[167] X-. Y. Du, K. J. Clemetson. (2002) *Toxicon*, **40**, 659-665.

[168] P. D. Pawelek, J. Cheah, R. Coulombe, P. Macheroux, S. Ghisla, A. Vrielink. (2000) *EMBO J.*, **19**, 4204-4215.

[169] S. M. Suhr, D. S. Kim. (1996) *Biochem. Biophys. Res. Commun.*, **224**, 134-139.

[170] S. M. Suhr, D. S. Kim. (1999) *J. Biochem.*, **125**, 305-309.

[171] Z. Y. Li, T. F. Yu, E. C. Lian. (1994) *Toxicon*, **32**, 1349-1358.

[172] G. I. Berglund, G. H. Carlsson, A. T. Smith, H. Szöke, A. Henriksen, J. Hajdu. (2002) *Nature*, **417**, 463-468.

[173] A. Lewis, M. Ough, L. Li, M. M. Hinkhouse, J. M. Ritchie, D. R. Spitz, J. J. Cullen. (2004) *Clin. Cancer. Res.*, **10**, 4550-4558.

[174] C. J. van Noorden, G. N. Jonges. (1992) *Cytometry*, **13**, 644-648.

[175] A. DeJong, M. van Kessel-van Vark, A. K. Raap. (1985) *Histochem. J.*, **17**, 1119-1130.

[176] H. Deng, D. O'Hagan, C. Schaffrath. (2004) *Nat. Prod. Rep.*, **21**, 773-784.

[177] M. Alexeeva, A. Enright, M. J. Dawson, M. Mahmoudian, N. J. Turner. (2001) *Angew. Chem. Int. Ed.*, **41**, 3177-3180.

[178] R. Carr, M. Alexeeva, A. Enright, T. S. C. Eve, M. J. Dawson, N. J. Turner. (2003) *Angew. Chem. Int. Ed.*, **42**, 4807-4810.

[179] R. Carr, M. Alexeeva, M. J. Dawson, V. Gotor-Fernandez, C. E. Humphrey, N. J. Turner. (2005) *ChemBioChem*, **6**, 1-3.

[180] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood. (2000) '*Practical Streptomyces Genetics*', John Innes Foundation, Norwich, UK.