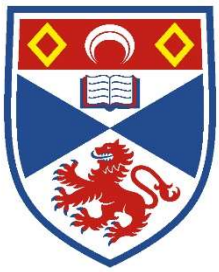
Studies of enzymes relevant to the biotransformation of fluorinated natural products

Nouchali Bandaranayaka



University of
St Andrews

Supervisor: Prof. David O'Hagan

This thesis is submitted in partial fulfilment for the degree of
PhD
at the
University of St Andrews

NOVEMBER 2016

Declarations

1. Candidate's declarations:

I, Nouchali Bandaranayaka hereby certify that this thesis, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in November 2009 and as a candidate for the degree of Doctor of Philosophy in November 2011; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2016.

Date _____ Signature of candidate _____

2. Supervisor's declaration:

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

Date _____ Signature of supervisor _____

3. Permission for publication:

In submitting this thesis to the University of St Andrews I understand that I am 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. I 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. I 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 publication of this thesis: No embargo on printed copy; embargo on all or part of electronic copy for a period of 2 years on the following ground: publication would preclude future publication.

Date _____ Signature of candidate _____

Signature of supervisor _____

To my Parents, Mahinda and Renuka For it is your dream that I have fulfilled.

Acknowledgement

I would like to thank my supervisor Prof David O'Hagan, for encouraging me to pursue a PhD in his research group at University of St Andrews and paving the way for me with constant guidance and support. He understood my strengths and weaknesses and inspired me to focus and develop my skills to reach my goals, and I will forever be indebted for this.

Coming from a Molecular Biology background and starting a PhD in an Organic Chemistry research group was unsettling, however past and present members of DOH research group and other scientists I collaborated with have supported me immensely through this journey. I thank Dr Jason Schmidberger, for the support provided in the lab when I first joined as an inexperienced graduate. I thank Dr Jason Chan and Dr Stephen Thompson for their invaluable advice and chemistry 101 sessions. I thank Dr Andrew Bent, Dr Jesko Kohene, Dr Magnus Alphey and Greg Mann of Prof J. Naismith group for teaching me the skills required for protein crystallography and collaborating on projects. My sincere gratitude goes to Dr Hai Deng, Dr Sergio Dall'Angelo, Dr Mayca Onega, Dr Margit Winkler and Prof Mateo Zanda for their collaborations.

During my time in St Andrews I've met many friends and colleagues who have supported and encouraged me constantly. I thank them all from my heart, Dr Long Ma, Dr Yi Wang, Dr Davide Bello, Dr Qingzhi Zhang, Dr Neil Keddie, Andrea, Tanya, Heidi, Maria, Axel, Carmela, Gosia, Maciek and many others I may have forgotten to mention, without all your support this work will not be possible. A special thank you goes to Dr Phillip Lowe for proof reading my thesis and providing me with helpful advice.

I'm very grateful for the support provided by the dedicated technical staff, Dr Tomas Lebl and Dr Filippo Stella for their help with NMR, Dr Sally Shiran and Dr Catherine Botting for their support with mass spectrometry. I would also like to thank the BMS staff including Margaret, Jean and Val for their support with regards to orders and other BMS related queries.

I would like to thank Lakshmi for being such an understanding, undemanding best friend despite me missing out on so many group events and not being there in need. I would also like to thank Fiona, Blair, Vinodh and all the other friends for being supportive and understanding.

I am especially grateful for the love and support I've received from my family in Sri Lanka. My parents strived to provide me the best education possible and encouraged me to pursue my interests. My father had always believed in me and have supported me through many life changing decisions I've made in the past. My mother taught me to be an independent, versatile woman, she always was and I hope you both are proud of my achievements. My brother has always been supportive of my choices and I cherish the siblinghood and closeness we share despite living far from each other, thanks very much for being there for me. I'm lucky to have such great parents-in law who were supportive and encouraging and I thank them whole heartedly.

A special thank you goes to my loving husband, Shyam who has been supportive, encouraging, patient and caring constantly though the best and worst times over the past 6 years we've been married. Thank you for putting up with my moods, tantrums and complains. It is your love that had made me strong, fearless and kept me sane through this journey.

Abstract

This thesis is focused on enzymes related to the biosynthetic pathway of fluorometabolite synthesis in *S.cattleya*. The first native fluorinating enzyme, fluorinase was isolated from a soil bacterium, *Streptomyces cattleya* in 2002. Fluorinase catalyses the reaction between *S*-adenosyl-L-methionine and inorganic fluoride to produce 5′-fluoro-5-deoxyadenosine (5′-FDA) and L-methionine as the first step of the fluorometabolite biosynthetic pathway. Fluorinase has been an attractive tool for incorporating ^{18}F into a limited number of substrates for applications in positron emission tomography (PET). This thesis describes a preparation of the fluorinase for PET, and then the production of [^{18}F]-5-fluoro-5-deoxy-D-ribose ([^{18}F]-FDR) *via* fluorinase mediated enzymatic synthesis. *S. cattleya* fluorinase has been the only fluorinase identified until recently when four more fluorinases have been identified by gene mining. These new fluorinase isolations are presented in the thesis. In addition this thesis describes the crystallisation of 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (FDRPi), an aldose-ketose isomerase involved in the biosynthetic pathway of fluorometabolite biosynthesis in *S.cattleya*.

Chapter 1 presents the background of this research focusing on the enzymes involved in the biosynthesis of the two fluorometabolites; fluoroacetate and 4-fluoro-L-threonine, produced by *S.cattleya*.

Chapter 2 describes the development of a practical, 'off the shelf' method of producing [^{18}F]-FDR in remote radiochemistry labs. Enzymes, fluorinase and nucleoside hydrolase, isolated from *Trypanosoma vivax* (*Tv*NH) were freeze-dried in their buffers to produce a shelf stable, potentially portable kit, where rehydration on site, would then provide catalysts on demand for radiochemical synthesis of [^{18}F]-FDR. This kit was practical enough to conduct a successful tumour imaging using a mouse model at Vrije University in Amsterdam.

Chapter 3 presents the over-expression and purification of fluorinase gene product (FLA1) from a newly isolated soil bacterium *Streptomyces sp. MA37*. The gene was identified by sequence mining of the *Streptomyces sp. MA37* genome. This fluorinase shared high homology to *S.cattleya* fluorinase and the *flA1* was cloned into *E.coli*, over-expressed, purified, assayed and shown to be a fluorinase. The FlA1 was also crystallized and the structure solved.

Chapter 4 describes the successful crystallisation of FDRPi, an enzyme involved in the fluorometabolite synthesis of *S. cattleya*. The FDRPi was over expressed, purified, crystallised and the structure was solved.

Abbreviations

[^{18}F]FAZA	[^{18}F]fluoroazomycin arabinose
[^{18}F]-FDU	2-amino-5'deoxy-5'-fluoroadenosine
[^{18}F]FLT	3'-deoxy-3'-[^{18}F]fluorothymidine
[^{18}F]FMISO	[^{18}F]Fluoromisonidazole
4-FT	4-fluoro-L-threonine
5-FDA	5′-fluoro-5′-deoxyadenosine
5-FDR	5-fluoro-5-deoxy-D-ribose
5-FDRP	5-fluoro-5-deoxy-D-ribose-1-phosphate
5-FDRulP	5-fluoro-5-deoxy-D-ribulose-1-phosphate
aa	amino acids
BLAST	basic Local Alignment Tool
CCD	charge coupled device
CFE	cell-free extracts
ClDA	5′-chloro-5′-deoxy-adenosine
coA	coenzyme A
COOT	crystallographic object-oriented toolkit
CT	computed tomography
DHAP	dihydroxyacetone phosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
duf-62	domains of unknown fuction-62
EDTA	ethylenediaminetetraacetic acid
FAc	fluoroacetate
FAd	fluoroacetaldehyde
FDEA	5'-fluoro-5'-deoxy-2-ethynyladenosine
FDI	5'-fluoro-5'deoxyinosine
FHPA	(2*R*,3*S*,4*S*)-5-fluoro-2, 3, 4-trihydroxypentanoic acid
FITR	fourier transform infrared spectroscopy
G3P	glyceraldehyde 3-phosphate
GC	guanine-cytosine
h	hours
HEPES	(4-2-hydroxyethyl)-1-piperazineethanesulfonicacid

HPLC	high performance liquid chromatography
IAG-NH	inosine/adenosine/guanosine nucleoside hydrolase
IF2	initiation factor 2
ISP	International Streptomyces Project
ITC	isothermal titration calorimetry
IU-NHs	inosine and uridine
L-AAO	L-amino acid oxidase
LB	Luria Broth
L-Se-met	L-selenomethionine
MALDI-TOF	matrix-assisted laser desorption/ionisation time-of-flight mass spectrophotometry
MRI	magnetic resonance imaging
min	minute
β-ME	2-mercaptoethanol
MTA	5′-deoxy-5′-(methylthio)adenosine
MTRP	*S*-methyl-5-thioribose-1-phosphate
MTRP	5-methylthioribose 1-phosphate
MTRPi	5-methylthioribose 1-phosphate isomerase
MTRuP	methylthioribulose 1-phosphate
NAD	nicotinamide adenine dinucleotide
NH	nucleoside hydrolase
NI-NTA	ni-nitrilotriacetic acid
NMR	nuclear magnetic resonance
NP40	nonidet P40
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PDA	photodiode array
PDB	protein data bank
PEG	polyethylene glycol
PET	positron emission tomography
PLP	pyridoxal phosphate
PNP	purine nucleoside phosphorylase
QM/MM	quantum-mechanical/molecular-mechanical

RCY	radiochemical yield
RGD	arginylglycylaspartic acid
rpm	rate per minute
rRNA	ribosomal RNA
s	seconds
SAH	*S*-adenosyl homocysteine
SAM	*S*-adenosyl-L-methionine
SCX	strong cation exchange cartridge
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
SHMT	serine hydroxymethyltransferase
SPE	solid phase extraction
SPECT	single-photon emission computed tomography
SUV	standardized uptake value
TCEP	tris(2-carboxyethyl)phosphine hydrochloride
TEG	tetraethylene glycol
TEV protease	tobacco Etch Virus endopeptidase
TFA	trifluoroacetic acid
Tris-HCl	2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride
TvNH	*Trypanasoma vivax* nucleoside hydrolase
UV	ultraviolet
VOIs	volumes of interest
w/v	weight/volume percent
WAX	weak anion exchanger

Table of Contents

Declarations i

Dedication ii

Acknowledgement iii

Abstract v

Abbreviations vi

Contents ix

1. Introduction 1

1.1 Cell metabolism 1

1.2 Fluorinated natural products in nature 3

1.2.1 Natural products in soil bacteria 3

1.2.2 Halogenated natural products 4

1.2.3 Fluorinated natural products 6

1.2.3.1 Fluoroacetate and (2*R*, 3*R*)-fluorocitrate 7

1.2.3.2 Fluorofatty acids 8

1.2.3.3 Nucleocidin 9

1.2.3.4 4-fluoro-L-threonine 9

1.2.3.5 (2*R*,3*S*,4*S*)-5-fluoro-2, 3, 4-trihydroxypentanoic acid 10

1.2.3.6 Misidentified fluorometabolites 10

1.3 Fluorometabolite biosynthetic pathway of *S. Cattleya* 11

1.3.1 The fluorinase 13

1.3.1.1 Mechanism of the fluorinase 17

1.3.1.2 Reversibility of the fluorinase 20

1.3.2 Purine nucleoside phosphorylase 22

1.3.3 Isomerase 23

1.3.4 Aldolase 23

1.3.5 PLP-dependent transaldolase 24

1.3.6 Aldehyde dehydrogenase 26

1.4 Chapter 1 References 27

2. Production of radiotracer [^{18}F]-5-FDR using fluorinase 33

2.1 Positron emission tomography (PET) 33

2.1.1 PET as a molecular imaging tool 33

2.1.2 [^{18}F]FDG as PET radiotracer 36

2.2 Application of the fluorinase for producing [^{18}F]-labelled PET radiotracers 40
2.2.1 Shortened enzymatic route to produce [^{18}F]-5-FDR 42
2.2.2 Nucleoside hydrolase 43
2.2.3 Previous studies conducted on production of [^{18}F]-5-FDR using the fluorinase 45
2.3 Project aims 46
2.4 Results and discussion 47
2.4.1 Over-expression of the fluorinase 47
2.4.2 Over-expression of the *Tv*NH 50
2.4.3 Enzyme immobilisation 52
2.4.4 Lyophilisation of the enzymes 53
2.4.5 Heat mediated SAM-Cl decomposition 54
2.4.6 Use of a sucrose-glycine excipient system for enzyme lyophilisation 56
2.4.7 Evaluating long term enzyme stability during storage 60
2.4.8 Non-labelled enzyme assays " cold reactions" 63
2.4.9 [^{18}F]-5-FDR production using freeze-dried fluorinase and *Tv*NH 64
2.4.10 PET studies on tumour mice using enzymatically produced [^{18}F]-5-FDR as a radiotracer 65
2.5 Conclusions 68
2.6 Experimental 68
2.6.1 Purification of Fluorinase enzyme in *E.coli* 68
2.6.2 Purification of Nucleoside hydrolase in *E.coli* 69
2.6.3 Freeze drying enzymes 70
2.6.4 Freeze-drying with sucrose excipient system 70
1.3.4 Aldolase 70
2.6.5 Non-labelled enzyme assays " cold reactions" 70
2.6.6 Assaying both enzymes for stability 70
2.6.7 Enzymatic preparation of [^{18}F]-5-FDR 71
2.6.8 Modified method for enzymatic preparation of [^{18}F]-5-FDR [^{18}F]-37 for tumour mice imaging 71
2.7.9 PET imaging of tumour induced mice using [^{18}F]-5-FDR 72
2.7 Chapter 2 References 74
3. Identification of other fluorinases 79
3.1 *S.cattleya* fluorometabolite gene cluster 79
3.1.1 Enzymes with homology to fluorinase 81

3.2 Characterisation of a fluorinase isolated from *Streptomyces sp.* MA37 84
3.2.1 Isolation and growth of *Streptomyces sp.* MA37 84
3.2.2 Genome analysis of *Streptomyces sp. MA37* 85
3.2.3 Gene mining for the *flA* gene 87
3.3 Project aims 90
3.4 Results and discussion 90
3.4.1 Isolation and Cloning of *flA1* 90
3.4.2 Over-expression and purification of MA37 fluorinase 91
3.4.3 Crystallization of MA37 fluorinase 98
3.4.3 Enzymatic assays- evaluation of kinetic parameters 100
3.4.5 Comparison of all fluorinases 104
3.5 Conclusions 105
3.6 Experimental 105
3.6.1 Genomic DNA extraction from *Streptomyces sp.* MA-37 105
3.6.2 Cloning of *flA1* gene into pEHISTEV plasmid 105
3.6.3 Over-expression and purification of *Streptomyces sp.* MA37 fluorinase 106
3.6.4Crystallization of *Streptomyces sp.* MA37 fluorinase 108
3.6.5 K_m determination of *Streptomyces sp.* MA37 fluorinase 109
3.7 Chapter 3 References 110
4. Crystallisation of 5-FDRPi Isomerase 113
4.1 Methionine salvage pathway 113
4.2 Aldose ketose isomerase 114
4.3 Ribose 1-phosphate isomerases 116
4.3.1 Identification of 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (5-FDRPi) from *S. cattleya* 116
4.3.2 Enzymes related to 5-FDRPi by sequence 117
4.3.3 The MTRPi from *Bacillus subtilis* 118
4.4 Previous studies on *Streptomyces cattleya* 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (5-FDRPi) 121
4.4.1 Site directed mutagenesis of 5-FDRPi 121
4.4.2 Truncation mutants of the 5-FDRPi based on a secondary structure prediction 121
4.4.5 A phosphonate analogue of the 5-FDRPi substrate 122
4.5 Project aims 123
4.6 Results and discussion 123

4.6.1 Over-expression and purification of 5-FDRPi 123
4.6.2 NMR assaying of 5-FDRPi and small scale enzymatic preparation of 5-FDRPi 126
4.6.3 Crystallisation of 5-FDRPi 131
4.6.4 Binding assays by Isothermal titration calorimetry (ITC) 136
4.6.5 Structural comparison of *B. subtilis* MTRPi and *S. cattleya* 5-FDRPi 137
4.7 Conclusions 139
4.8 Experimental 140
4.8.1 Over-expression and purification of 5-FDRPi 140
4.8.2 NMR assays to test for activity of 5-FDRPi 141
4.8.3 Small scale enzymatic preparation and purification of 5-FDRP 142
4.8.4 Isothermal titration calorimetry (ITC) mediated binding assays 142
4.8.6 Crystallisation of 5-FDRPi 142
4.9 Chapter 4 References 144
5. Thesis conclusion 146
5.1 Thesis conclusion and future work 146
5.2 Chapter 5 References 150
Appendix 1 152
Appendix 2 153
Appendix 3 154
Appendix 4 155

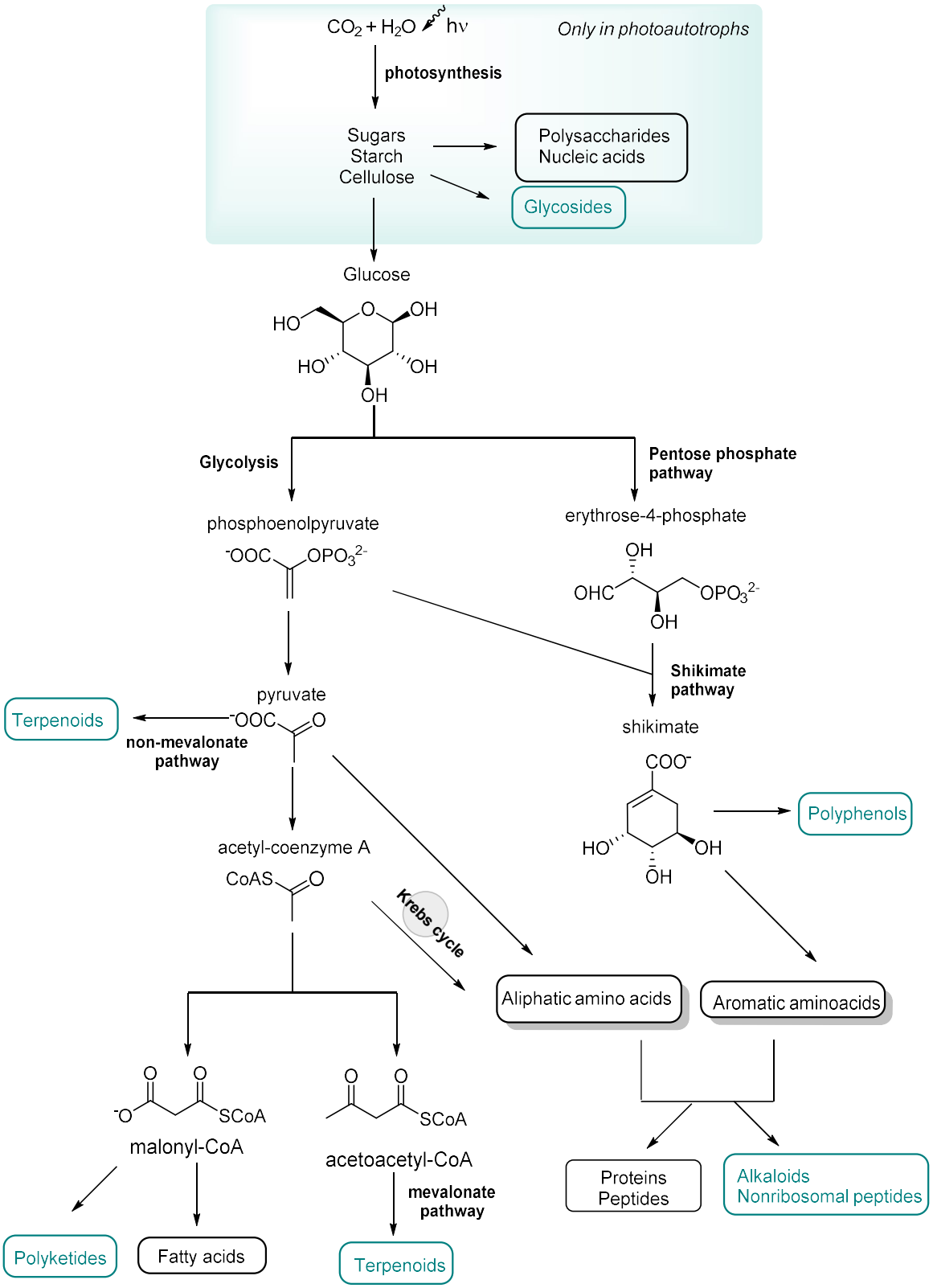
1. Introduction

1.1 Cell metabolism

A cell's routine operations are accomplished through biochemical reactions that take place within the cell itself. These reactions are switched on/off and their rates increased/decreased in order to cater to the cell's immediate needs and overall survival. There are a vast number of essential enzyme-mediated pathways, involved in the building up and breaking down of important cellular components, and these pathways are often highly regulated through diverse mechanisms.

Metabolism is divided into two main components:
Primary metabolism consists of the common metabolic pathways which are essential for cell survival and function. It involves catabolic and anabolic reactions which leads to energy production, cellular reproduction, growth and development, as well as the synthesis of building blocks for proteins, nucleic acids and lipids. Plants and algae produce organic compounds from inorganic materials which are then utilised by other animals and microorganisms. Biomolecules produced by primary metabolism are known as primary metabolites and are universally found in Prokaryotes and Eukaryotes.[1,2]

Secondary metabolism includes chemical processes dependant on primary metabolism. This involves the biosynthesis of metabolites that have no apparent basic function to the cell, but instead serve as signalling molecules, antibiotics, toxins and pigments which contribute to the management of limited resources and interspecies defence. Unlike primary metabolism, hindrance to secondary metabolism does not incur cell death, but may affect in the organisms survival and reproducibility. Many secondary metabolites are produced by bacteria, varying in their chemical nature and the biosynthetic pathways in which they are produced. Secondary metabolites are not necessarily produced under all conditions and have limited distribution in nature.[3] Mostly, primary and secondary metabolism are competing metabolically as they share common precursors (Scheme 1).[1] The main building blocks for the biosynthesis of secondary metabolites are acetate, shikimate and amino acids which are used to generate highly diverse and complex molecules. The main classes of secondary metabolites include: polyketides and polyphenols[4], terpenoids, non-ribosomal peptides, alkaloids and glycosides.[5]



Scheme 1: Interconnection between primary and secondary metabolism. Secondary metabolites are highlighted in green.[1,2,5]

1.2 Fluorinated natural products in nature

1.2.1 Natural products in soil bacteria

Plants and plant preparations have been used for thousands of years, to treat human and animal diseases.[6] Natural products have been the largest contributors to drugs in the history of medicine. Estimates suggest that over 49% of all human drugs have originated from natural sources.[7] The total number of species on Earth is estimated to be over 8 million,[8] however less than 2 million organisms have been identified and characterised, of which relatively few have been explored for their chemical content.[9] The majority of the antimicrobial products currently used in the medical field have been isolated from soil bacteria. There is a large genetic diversity in soil microorganisms that carry out the biosynthesis of complex molecules. These soil microbes produce antibiotics as secondary metabolites mainly in order to out-compete other bacteria for the limited resources available in their environment and in interspecific interactions. [10,11]

More than 80% of antibiotics currently in use were originally discovered from actinomycetes, most importantly from the *Streptomyces* genus (Figure 1).[12,13] The history of antibiotics derived from *Streptomyces* began with the discovery of streptothricin in 1942[14] following which streptomycin[15,16] was discovered soon after. Streptomycin was one of the first effective treatments for tuberculosis at that time.[17] Following the discovery of streptomycin, screening for antibiotics within the *Streptomyces* genus increased and over time *Streptomyces* have become the leading antibiotic producing genera of bacteria.[18]

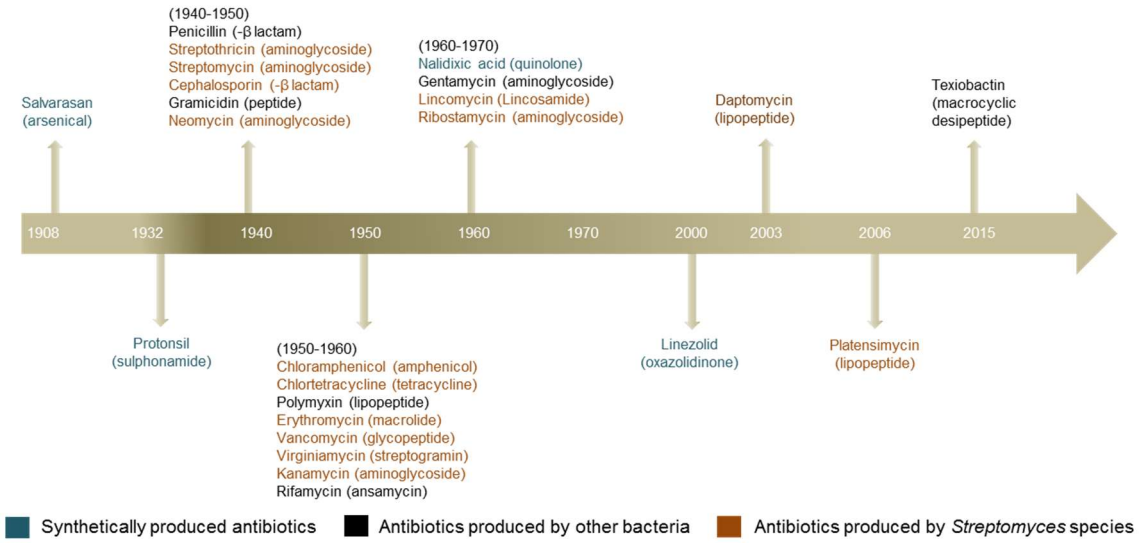


Figure 1: A timeline of antibiotics discovered from 1940s to 2006, most of which were discovered from the *Streptomyces* genus (highlighted in orange) and the class of each antibiotic is enclosed in brackets **.**[13,19,20]

Streptomyces are aerobic, Gram positive bacteria with particularly high guanine and cytosine (GC) content in their DNA (deoxyribonucleic acid). Over 70% of the *Streptomyces* DNA is GC rich as oppose to an *E.coli* bacterium, which is around 50% GC rich. [21] They have a complex life cycle consisting of various morphological and physiological changes. Their development initiates with spores germinating to form germ tubes as they encounter favourable conditions and nutrients. The germ tubes grow by tip extension and branch to form a vegetative mycelium, which grows until the surrounding nutrients are exhausted. In response to nutrient depletion both the production of secondary metabolites and morphological differentiation are initiated. Some of these secondary metabolites are antimicrobials, which are excreted to the surrounding environment in order to out-compete other microbes and safeguard their limited resources. As the vegetative mycelium matures, it starts to form aerial hyphae, which disrupt the surface tension of the air-water interface and grow in to the air. The aerial hyphae further transform into coiled threads of spore particles which mature and disperse as the life cycle concludes (Figure 2). This process of hyphae differentiating into spores is unique among gram-positive bacteria and requires a coordinated metabolism.[18,19,22]

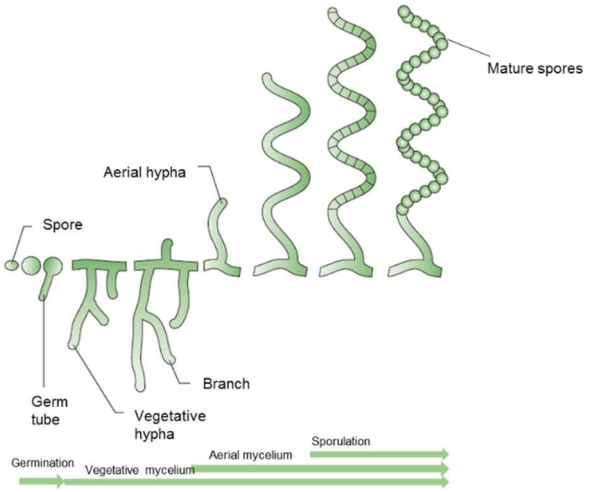


Figure 2 : Developmental life cycle of *Streptomyces coelicolor*. Image reproduced with permission from Buttner *et al*.[22]

1.2.2 Halogenated natural products

Biological halogenation occurs on a variety of organic molecules ranging from simple methyl halides[23,24] to terpenes,[25] complex polyketides[26] and large nonribosomal peptides.[27] Within these chemical structures, halogen atoms are incorporated on to aliphatic carbons, olefinic centers and a wide variety of aromatic and heterocyclic rings.[28]

Over 4500 halogen containing natural products have been discovered and characterised to date.[29] The most common class of halogenated natural products are the organochlorines. There have been over 2300 organochlorines identified to date, among which vancomycin **1**,[30] chlortetracycline **2** [31] and chloramphenicol **3** [32] (all isolated from *Streptomyces*) are clinically used antibiotics (Figure 3). Rebeccamycin **4** and salinosporamide A **5** are organochlorines isolated from actinomycetes, and are potent anticancer drugs (Figure 3).[33]

Around 2100 organobromines including thyrsiferyl acetate **6** [34], halomon **7** [35] and brominated acetogenins such as laurenidificin **8** [36] (isolated from marine red algae) have been identified. Brominated compounds are predominantly produced by marine organisms. Nearly 120 organoiodides such as calicheamicin **9** (isolated from *actinomycetes*),[37,38] and few thyroid hormones produced by thyroid gland in humans, have been identified. However only around 6 organofluorine compounds have been identified so far, demonstrating that fluorinated natural products are a rare occurrence in nature.[28]

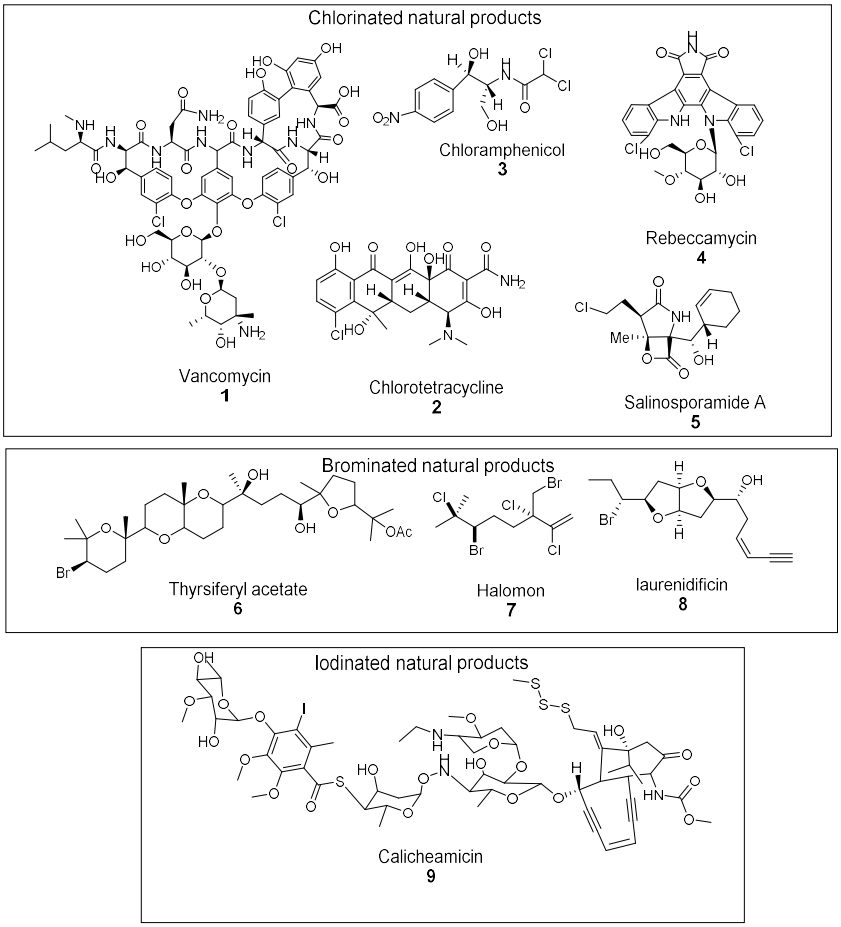


Figure 3: Examples of halogenated natural products that have been discovered and identified to possess cytotoxicity or antibiotic activity.[9,39]

1.2.3 Fluorinated natural products

Fluorine is the thirteenth most abundant element in the Earth's crust and the most abundant halogen on earth (chlorine 20th, bromine 62nd and iodine 64th).[40] However, most of the fluorine found in nature is combined within insoluble salts in the form of cryolite (Na_3AlF_6) and fluorspar (CaF_2).[41] Fluorinated natural products are rare as many of the biosynthetic mechanisms applicable to the other halogens are not mechanistically plausible for fluorine (i.e. oxidation to F^+). This is due to fluorine being the most electronegative element with a very high oxidation potential. However, a major contributing factor in reduced participation of fluorine in biochemistry is due to its high heat of hydration. The high hydration energy of fluoride ion, due to its high charge density lowers its nucleophilicity further reducing its reactivity.[42] Therefore fluorinated natural products are a rare occurrence in nature. Figure 4 shows the structures of the 6 known fluorinated natural products.

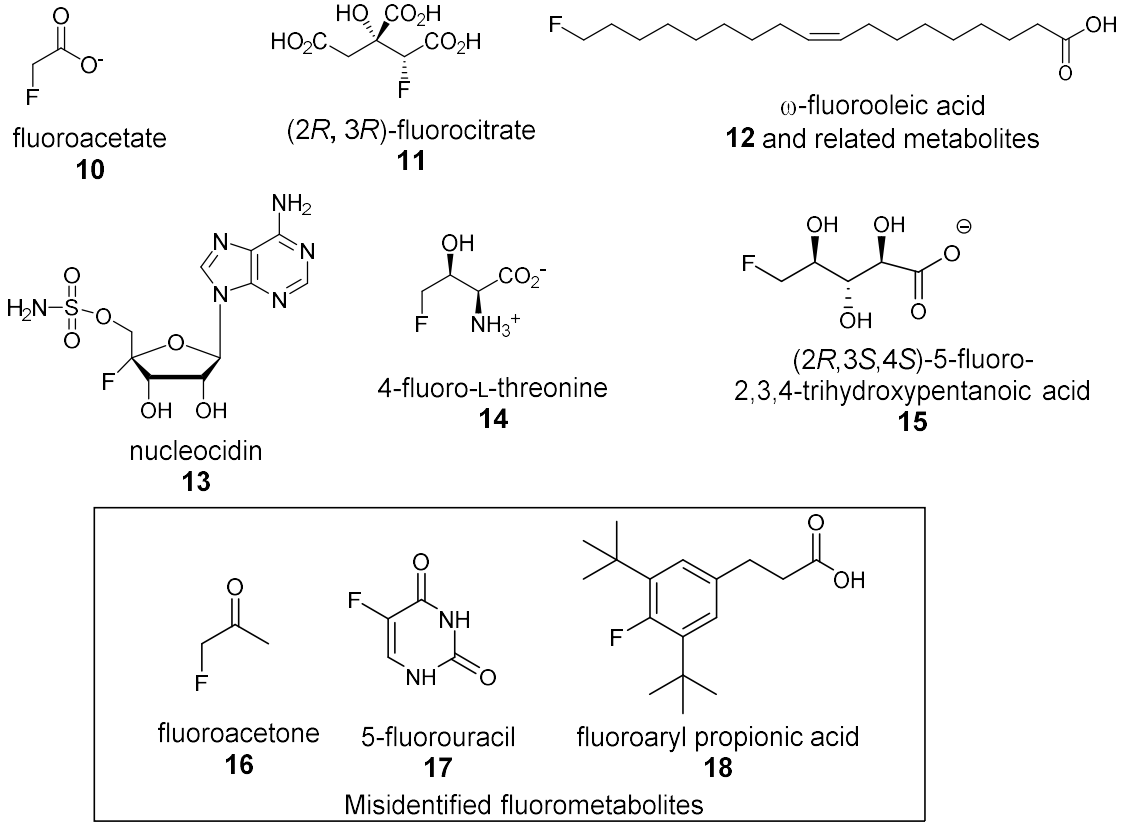


Figure 4: The known fluorinated natural products along with some fluorinated compounds thought to be misidentified as natural products.[42–46]

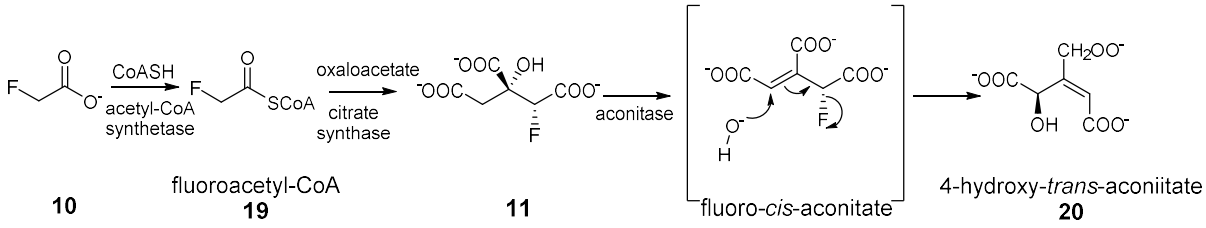
1.2.3.1 Fluoroacetate and (2*R*, 3*R*)-fluorocitrate 11

Fluoroacetate (FAc) **10** is the most common organofluorine compound found in nature and was the first fluorinated natural product to be identified.[42] This metabolite was first isolated from the toxic South African plant, *Dichapetalum cymosum* in 1943.[47] FAc **10** was found to accumulate at around 250 ppm in its leaves, which are highly toxic to livestock.[46] Very high concentrations of FAc **10** were observed in young leaves and seeds, and as they mature the FAc **10** levels seem to drop. The FAc **10** levels expressed in various plant species seem to vary widely depending on the season, age of the plant, parts of the plant and geographical locations. Since this discovery a variety of plants, mainly of the *Dichapetalum* genus along with other poisonous plants native to Brazil, Africa and Australia were also identified to produce FAc **10**.[42,46] Although the exact mechanism of FAc **10** biosynthesis in plants is unknown, its toxicity and the mechanism of **10** poisoning in mammals is well studied.

The extreme toxicity of FAc **10** arises from its similar metabolism to acetate. Its toxicity is dependent on a metabolic activation or 'lethal synthesis' to fluoroacetyl coenzyme A (CoA). FAc **10** taken up by cells subsequently enters the Kreb's cycle as fluoroacetyl-CoA **19**, by attaching to CoA in a reaction catalysed by acetyl-CoA synthetase. Usually acetyl-CoA synthetase catalyses the reaction between acetate and CoA in the presence of adenosine triphosphate (ATP) to form acetyl-CoA, a key molecule involved in cellular respiration. When fluoroacetyl-CoA **19** is substituted for acetyl-CoA in the Kreb's cycle, it is processed by the first enzyme, citrate synthase, to generate (2*R*,3*R*)-fluorocitrate **11** in a condensation reaction with oxaloacetate.[48] (2*R*,3*R*)-fluorocitrate **11**, specifically inhibits acotinase,[49] whereas the other diastereomer, (2*S*,3*R*)-fluorocitrate **11** is non-inhibitory. Aconitase is the next enzyme of the Kreb's cycle and catalyses the stereospecific isomerization of citrate to isocitrate *via* a *cis*-aconitate intermediate. When citrate is substituted with (2*R*, 3*R*)-fluorocitrate **11** it is metabolised to fluoro-*cis*-aconitate. This is followed by hydration along with concomitant defluorination to form 4-hydroxy-*trans*-aconitate **20**. This intermediate binds tightly to the active site and is a potent competitive inhibitor of aconitase[50] resulting in a total shut down of the metabolic cycle halting ATP production (Scheme 2).[51]

Plants and bacteria, which produce FAc **10** for self-defence, require themselves to be resistant to the fluoroacetyl-CoA **19** mediated inhibition. The process by which plants achieve resistance to fluoroacetyl-CoA **19** has been widely investigated however little is known as plants are complex organisms, and difficult to subject to gene manipulation. Studies in *S. cattleya*, an FAc **10** producing soil bacteria presented an agreeable

biosynthetic system, in which to study the biosynthesis of FAc **10**. The enzyme fluoroacetyl-CoA **19** thioesterase (FIK) was identified[52] in *S. cattleya* which could selectively hydrolyse the thioester bond of fluoroacetyl-CoA **19** to form FAc **10** and CoA with a ~10^6-fold selectivity for the fluoroacetyl-CoA **19** over the intracellularly abundant acetyl-CoA.[53] This selectivity towards fluoroacetyl-CoA **19** by the enzyme, despite the subtle change in chemical structure, assures that acetyl CoA remains unhydrolysed and available for cellular energy production. Crystal structure analysis of FIK demonstrated specific hydrogen bonding of catalytic site residues with the fluorine of the fluoroacetyl-CoA **19** which grants high selectivity towards fluoroacetyl-CoA **19**.[54,55] In addition the molecular recognition of fluorine in both formation and breakdown of the acyl-enzyme intermediate was shown to control the active site reactivity in FIK.[56,57] A similar hydrolase was also identified from mitochondrial extracts of the plant *D. cymosum,* which could also catalyse the specific hydrolysis of fluoroacetyl-CoA **19**.[58] Therefore plants and bacteria may have similar resistance mechanisms.

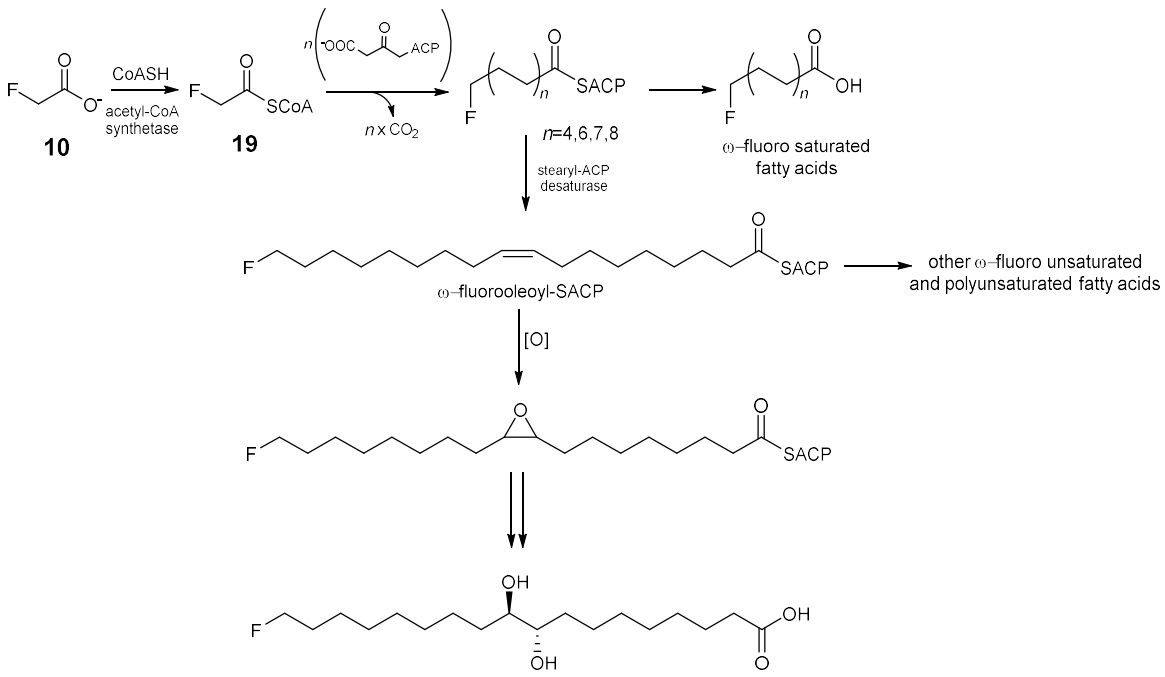


Scheme 2: The 'lethal synthesis' of *(2R, 3R)*-fluorocitrate **11**, which inhibits aconitase by metabolising into 4-hydroxy-*trans*-aconitase **20** and binding as a potent inhibitor.[48,59]

1.2.3.2 Fluorofatty acids

In 1964 a West African shrub *Dichapetalum toxicarium* was found to produce FAc **10** in its leaves, and when the plant products were analysed for fluorometabolites, ω-fluorofatty acids were isolated from the seeds.[42,60] The seeds mostly contained ω-fluoro-oleic acid **12** with small quantities of ω-fluoropalmitic acid and trace amounts of other fluorinated fatty acids.[61] These fatty acids are produced when **19** is substituted for acetyl-CoA as a starter unit in the fatty acid biosynthesis. Fatty acid synthases are multi-enzyme complexes that are involved in fatty acid assembly. At the initial stage they perform a condensation reaction of acetyl-CoA (starter unit) with malonyl acyl carrier protein (extender unit). This enzyme has broad substrate specificity which allows incorporation of various starter units other than acetyl-CoA. In the case of fluorofatty acids, the starter unit is a fluoroacetyl-CoA **19** which is consistent with the localisation of fluorine to the terminus of the fatty acid such as in **21** which was isolated from the saponified seed oil of *D. toxicarium*. The lack of

fluorine along the chain suggests enzyme constraints in the process or an inability of the fatty acid synthase complex to utilise fluoromalonyl-CoA as a substrate (Scheme 3).



Scheme 3: Biosynthetic pathway of ω-fluorofatty acids in *D. toxicarium*. Image reproduced and modified with permission from O'Hagan *et al*.[46]

1.2.3.3 Nucleocidin 13

Streptomyces calvus, a soil bacterium isolated from an Indian soil sample, was found to produce the adenosine derived antibiotic, nucleocidin **13**. The chemical structure of nucleocidin **13** is significantly different to any other fluorometabolite identified to date, as the fluorine atom is attached to the C-4 of the ribose ring. This does not show any obvious biosynthetic origin from FAc **10** and suggests the presence of a novel fluorination enzyme. Nucleocidin **13** was found to be an effective anti-trypanosomal agent, however due to its toxicity clinical use had to be ceased.[62] Further studies on the biosynthesis of this antibiotic have been challenging, as the strains available in public collections are unable to produce nucleocidin **14**.[63,64]

1.2.3.4 4-Fluoro-L-threonine 14

In 1986, Sanada and co-workers carried out experiments optimising the productivity of the β-lactam antibiotic, thienamycin in *S. cattleya.* This led to the discovery and production of the fluorinated antibiotic 4-fluoro-L-threonine **14** (4-FT). The antibiotic was produced only when a specific soya protein was used in the culture medium. The soya proteins were found to contain significant amounts of fluoride, which led to fluorometabolite production.

In addition to 4-FT **14** the organism also produced FAc **10** and the enzymes which catalyse the production of these fluorometabolites were subsequently identified and purified.[59]

1.2.3.5 (2R,3S,4S)-5-fluoro-2, 3, 4-trihydroxypentanoic acid (5-FHPA) 15

A novel fluorometabolite, (2*R*,3*S*,4*S*)-5-fluoro-2, 3, 4-trihydroxypentanoic acid (5-FHPA) **15** was identified from *Streptomyces* sp. MA37 (Figure 4). The production of **15** was distinctive to the production of **10** and **14** and a 5-fluoro-5-deoxy-D-ribose (5-FDR) **37** mediated biosynthesis was suggested.[43] Details of the enzymes involved in the production of **15** will be discussed in detail in Chapter 2.

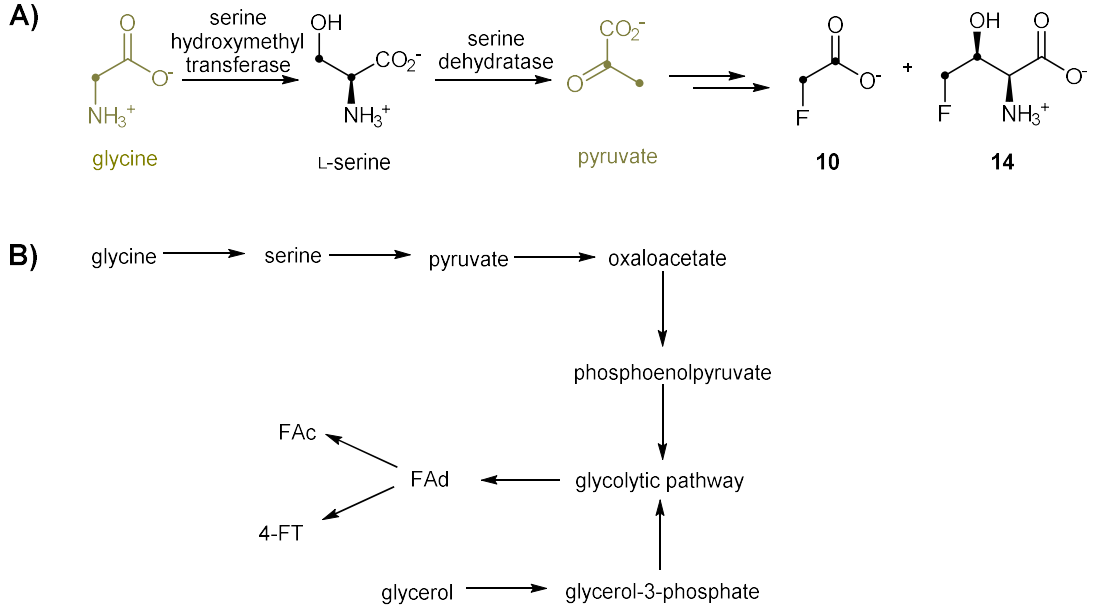
1.2.3.6 Misidentified fluorometabolites

There have been some misidentifications of fluorometabolites reported in the literature. Fluoroacetone **16** was reported from volatiles of the plant extracts of *Acacia georginae* in 1970 as a dinitrophenylhydrazone derivative. However, the identity of volatile fluoroacetone **16** was reported to be uncertain.[65,66] 5-Fluorouracil **17** alkaloids and similar derivatives were isolated from the marine sponge, *Phakellia fusca* collected from the South China sea. However the origin of these metabolites in the sponge remains unclear due to the structural similarity to 5-fluorouracil **17,** a well-known anti-cancer pharmaceutical, as it is possible that they are acquired from the ocean as a by-product of industrial contamination.[45,67] In 2014 Marimuthu *et al.* reported the identification of fluoroarylpropionic acid **18** from *Streptomyces* sp. TC1 (Figure 4). This discovery suggested that the bacterium contains an enzyme capable of enzymatic aryl fluorination, which was not known before. The structure assignment was largely supported by X-ray crystallography where the presence of a hydroxyl group and a fluorine is not accurately distinguishable due to their similar electron density and the spectroscopic support was unclear.[68] O'Hagan *et al.*[44] synthesised the fluoroarylpropionic acid **18** and performed comparative spectroscopic analysis on the proposed structure to confirm its identity. The spectroscopic data obtained from the synthesised compound was inconsistent with the data of the reported structure and no supporting evidence was obtained to show the proposed structure contained any fluorine.[44]

1.3 Fluorometabolite biosynthetic pathway of *S. cattleya*

S. cattleya is a known producer of **10** and **14**. Since it was the first bacterium to be discovered able to produce fluorometabolites, *S. cattleya* has become a model organism in which the biosynthesis of fluorometabolites are studied.

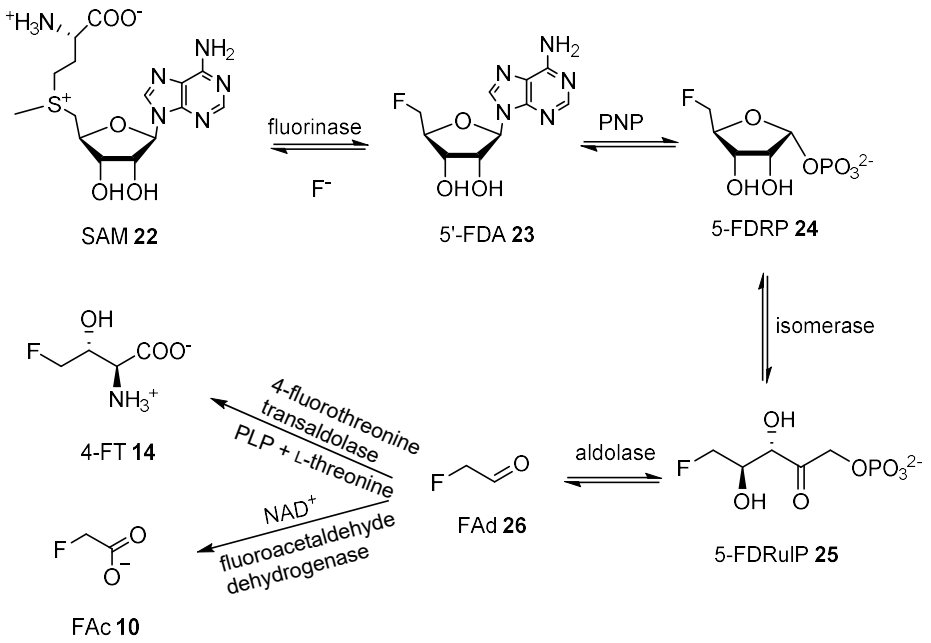
Isotope labelling studies were carried out to probe the biosynthetic pathway of fluorometabolite biosynthesis by *S. cattleya*. Various ^{13}C and ^{2}H labelled precursors were used to analyse the incorporation of these labels into the fluorometabolites by ^{19}F {H} NMR and GC/MS. Carbon-^{13}C labelled glycine and pyruvates was observed to give high incorporations into the two fluorometabolites. The [2- ^{13}C]-glycine and the [2- ^{13}C] and [3- ^{13}C] pyruvate were shown to provide both carbon atoms of FAc **10** as well as C-3 and C-4 of 4-FT **14**, suggesting a common origin for both molecules (Scheme 4A).[69] The labelling patterns indicated that glycine was converted first to serine, and then to pyruvate. Pyruvate was further metabolised to phosphenolpyruvate before incorporation into the fluorometabolites (Scheme 4B).[70] Furthermore [2-^{13}C]-glycerol was shown to be incorporated into C-1 of FAc **10** by *S. cattleya.*[71,72] As both pyruvate and glycerol feed into the glycolytic pathway, the common intermediate for the fluorometabolite production was thought to be a product of the glycolytic pathway. Further studies confirmed that the phosphenolpyruvate formed from glycine is converted to fluoroacetaldehyde (FAd) **26** which is the common intermediate for the two fluorometabolite production (Scheme 5).[73]



Scheme 4: A) Proposed pathway resulting in labelled fluorometabolites, **10** and **14** *via* [2-^{13}C]-glycine and ^{13}C labelled pyruvates. Image reproduced with permission from O'Hagan *et al*.[59] **B)** Relationship between primary and secondary metabolites in fluorometabolite synthesis of *S. cattleya*. Image reproduced with permission from Harper *et al*.[59]

When resting cell cultures of *S. cattleya* were incubated with synthetically prepared [2,2-$^{2}H_2$] FAd **26**, an efficient conversion to deuterated [4-$^{2}H_2$]-4-FT and [2-$^{3}H_2$]-FAc was observed. These studies confirmed that FAd **26** is a common precursor to both FAc **10** and 4-fluorothreonine **14**.[73] Intermediates involved in the biosynthetic pathway of the fluorometabolites in *S. cattleya* are shown in Scheme 5.

The first step in fluorometabolite biosynthesis involves the fluorinase. This enzyme catalyses C-F bond formation. The fluorination process involves a nucleophilic reaction between *S*-adenosyl-L-methionine (SAM) **22** and a fluoride ion to generate 5′-fluoro-5′-deoxyadenosine (5′-FDA) **23**. 5′-FDA **23** is the substrate for a purine nucleoside phosphorylase (PNP), which generates 5-fluoro-5-deoxy-D-ribose-1-phosphate (5-FDRP) **24**.[74] An isomerase then catalyses the conversion of 5-FDRP **24** into 5-fluoro-5-deoxy-D-ribulose-1-phosphate (5-FDRulP) **25**.[75] The next step, involves an aldolase, which converts 5-FDRulP **25** in a retro-aldol reaction, to **26** and dihydroxyacetone phosphate (DHAP) **31**. So far the aldolase responsible for cleavage of 5-FDRulP **25** to **26**, is not identified.[76] FAd **26** is a key branch point in the biosynthetic pathway. An NAD^+-dependent aldehyde dehydrogenase was isolated from cell-free extracts (CFE) of *S. cattleya* which oxidises **26** to **10**. This enzyme has a high affinity for **26** and glycoaldehyde but not acetaldehyde,[77] possibly due to the enzyme evolving from an aldehyde dehydrogenase which utilised either glycoaldehyde or a similar compound as the natural substrate. A pyridoxal phosphate (PLP)-dependent enzyme, 4-FT **14** transaldolase was isolated from CFE of *S. cattleya*. This enzyme catalyse a transaldol reaction between the amino acid L-threonine and FAd **26** to produce **14** and acetaldehyde. This transaldolase does not utilise glycine as the amino acid donor, and for every molecule of 4-FT **14** produced, it sacrifices a molecule of L-threonine.[42] The next section discusses the details of each of these enzymes.



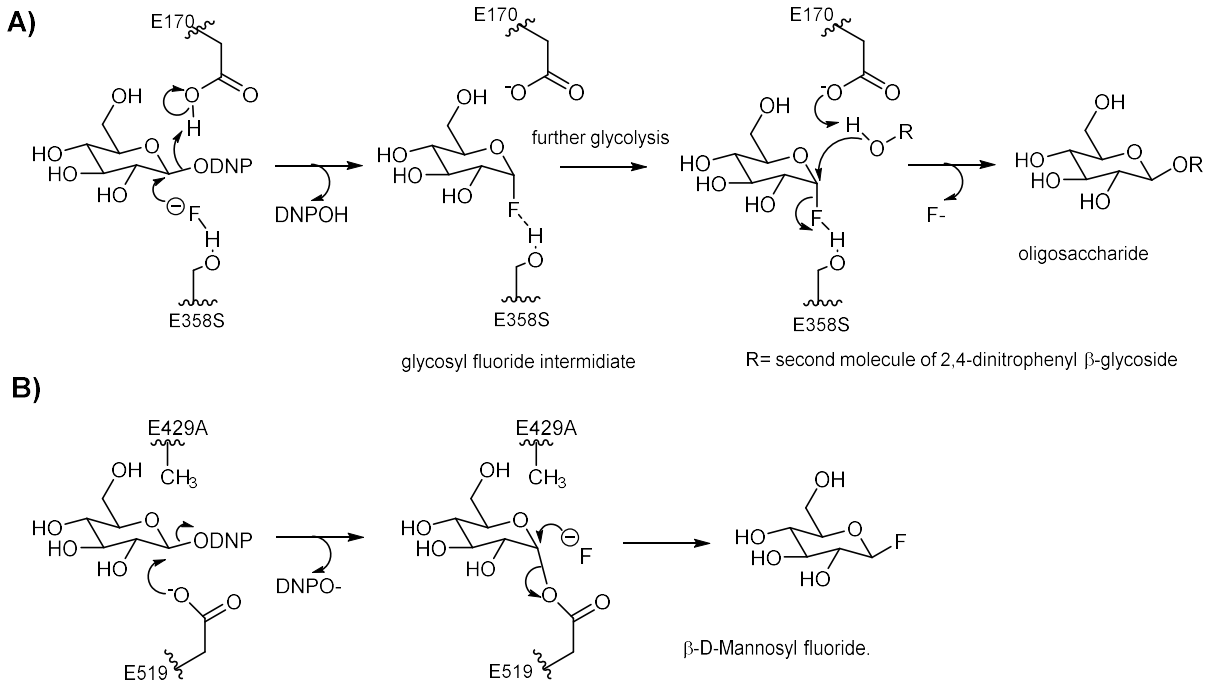
Scheme 5: Biosynthetic pathway of fluorometabolites in *S. cattleya*, leading to the production of **10** and **14**. Image reproduced and modified with permission from O'Hagan *et al*.[75]

1.3.1 The fluorinase

The key step in the biosynthesis of **10** and **14** in *S. cattleya* is C-F bond formation. Until the discovery of the fluorinase from *S. cattleya* in 2002,[78] a native enzyme which could catalyse C-F bond formation from fluoride ion was unknown.

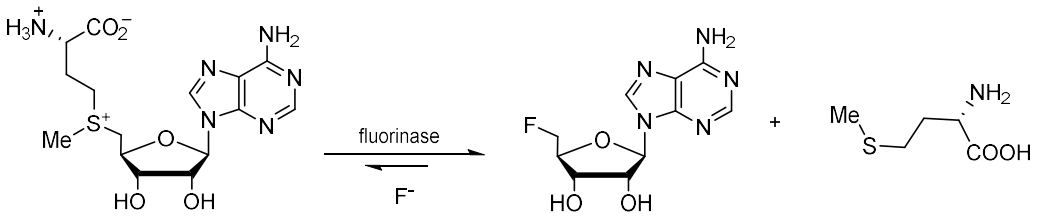
An enzyme mediated synthesis of C-F bond formation was reported in two mutant glycosidases isolated from *Agrobacterium* sp. and *Cellulomonas fimi*. These mutants were capable of incorporating a fluoride ion into the C1 position of gluco- and mannosaccharides *via* nucleophilic displacement of 2, 4-dinitrophenol. Catalytic glutamate residues Glu 358 in *Agrobacterium* sp. and Glu-429 in *C. fimi* were replaced with either a serine or alanine respectively, which inhibited the glycosidase activity. However, in the presence of high fluoride concentrations (2 M) the glycosidic activity of these mutants was restored. Mutation of the glutamate to serine in *Agrobacterium* sp. provides an opportunity for the solvated fluoride to hydrogen bond to serine resulting in partial desolvation of the fluoride. This allows the fluoride ion to perform a nucleophilic attack on the highly electrophilic glycosyl donors to produce the glycosyl fluoride intermediate. Since these mutants are glycosynthases, catalysing trans-glycosidation the glycosyl fluoride intermediate acts as the activated substrate for glycosylation forming a new glycosidic bond (Scheme 6A). Mutation of glutamate to alanine in *C. fimi* creates more space which

can be occupied by fluoride for nucleophilic attack of β-manopyranosyl to form β-D-mannosyl fluoride (Scheme 6B). These mutant glycosidases are fluoride specific and are unable to use other halides, however replacing glutamic acid with glycine produced mutants capable of utilising fluoride, chloride and bromide but not iodide. [79–81]



Scheme 6: A) *Agrobacterium* sp. E358S mutant catalyses the nucleophilic fluorination on 2, 4-dinitrophenyl *β*-glycoside, followed by intermediate α-glucosyl fluoride becoming a substrate for glycosylation. **B)** *C. fimi* E429A β-mannosidase catalyses nucleophilic fluorination of enzyme bound mannosyl intermediate to form β-D-mannosyl fluoride. Image reproduced and modified with permission from Withers *et al*.[79]

Fluorometabolite production was observed when CFE of *S. cattleya* was incubated with ATP and fluoride. Further experiments with the fluorinase revealed, that added SAM **22** (in the absence of ATP) was able to support fluorometabolite production in CFEs. Since SAM **22** synthetase can mediate the reaction between ATP and L-methionine **27** to form SAM **22**, this suggested SAM **22** as the immediate substrate for fluorination by the enzyme (Scheme 7).[82]



Scheme 7: Fluorinase catalyses the conversion of **22** and a fluoride ion to **23** and **27**.

When SAM **22** was recognised as the immediate substrate it was postulated that 5′-FDA **23** was possibly the product of the fluorinase reaction. This was confirmed by preparing a synthetic sample of 5′-FDA **23** and performing comparative analytical analysis with the product of the enzyme reaction.[83]

Isolating and purifying the fluorinase, was important to investigate the characteristics of the enzyme. Partial purification of the wild type enzyme from CFE was carried out using a combination of ammonium sulphate precipitation and several chromatography steps. *N*-Terminal amino acid analysis and a trypsin digestion was performed on the purified protein fraction which provided a partial amino acid sequence for the fluorinase gene. Using the partial sequence, PCR primers were designed for the fluorinase gene and the gene was PCR amplified using genomic DNA extracted from *S. cattleya* as a template. This gene was cloned in to pET28A (+) vector and the recombinant protein was over expressed in *E. coli* on a mg scale. Gene sequencing revealed the *flA* gene to be 897 base pairs long encoding for a protein of 299 amino acids. For ease of purification, the protein was overexpressed with a polyhistidine affinity tag. Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry confirmed the molecular weight of the fluorinase monomer to be 32 kDa. The intact mass of the fluorinase was approximately 196 kDa confirmed by size exclusion chromatography, suggesting a hexameric tertiary structure.[84]

The fluorinase crystal structure was resolved to a 1.8 Å resolution. This proved the hexameric arrangement of the enzyme constructed as a dimer of trimers with three active sites (Figure 5). The secondary structure of the fluorinase monomer did not show any similarity to other previously characterised proteins. Figure 4 shows the hexameric structure of fluorinase and a single trimer. The monomer consists of two distinctive domains (N and C terminal) linked with a short polypeptide chain as shown in Figure 5. The monomer contained a unique 21 amino acid loop which spans close to the C-terminal domain and consists of a part of the active site.[85] This amino acid loop had no relationship to any other protein superfamily identified at the time of discovery, possibly suggesting a

conserved region of the protein essential for function. Each trimer has three N-terminal domains, closely arranged around a three-fold axis while the three C-terminal domains are in contact with the adjacent N-terminal domain, but not with each other.[85]

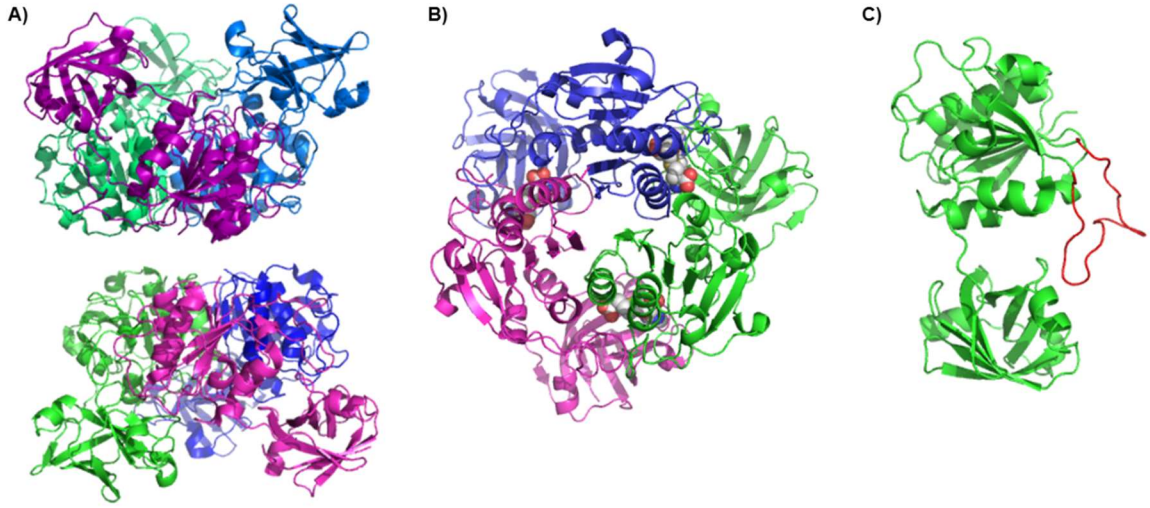


Figure 5: Cartoon depictions of the crystal structure of the fluorinase **A)** The hexameric fluorinase made up of a dimer of trimers. **B)** A single trimer made of up of three monomers showing 3 SAM **22** molecules bound to each catalytic site. **C)** A monomer unit showing the N-terminal (top) and C terminal (bottom) linked by a single polypeptide chain. The unique 21 amino acid loop is highlighted in red.[85]

The crystal structure of fluorinase was resolved with SAM **22** bound at the active site showing specific interactions surrounding the components of SAM **22**. The adenine ring, ribose ring and the L-methionine **27** chain of SAM **22** is recognised by an arrangement of hydrogen bonds and van der Waals interactions. The SAM **22** molecule was observed to be completely concealed within the protein upon binding, suggesting the presence of an open conformation prior to substrate binding. The ribose ring of SAM **22** was held in an unusually planar conformation by hydrogen bonds between the 2′ and 3′-hydroxyl groups and the carboxylate side chain of Asp-16 as shown in Figure 6. Due to these hydrogen bonds the ribose ring appears to be held in a strained manner. Fluorinase was co-crystallised with 2′-deoxy-FDA to better understand the importance of the strained planar conformation of the ribose ring in driving the substitution reaction. However this conformation was shown to be un-important in driving the reaction. In the SAM **22** bound fluorinase structure all heteroatoms of the adenine ring are involved in hydrogen bonding however the ether oxygen of the ribose ring is not. Studies aimed at understanding the mechanism of the enzyme reaction are discussed in the next section.

1.3.1.1 Mechanism of the fluorinase

A study of the mechanism of the fluorinase enzyme has been supported by crystallography,[85] site-directed mutagenesis[86] and quantum-mechanical/molecular-mechanical (QM/MM) theory.[87]

Co-crystallisation attempts of the fluorinase with SAM **22** in the presence of fluoride gave rise to a crystal structure of the fluorinase with 5′-FDA **23** and L-methionine **27** bound to the active site. Comparative analysis of the active sites of the two crystal structures provided insight into the mechanism of the reaction. The sulphur was seen to be positionally displaced and the C-F bond was shown to be approximately antiperiplaner to the C-S bond in SAM **22** indicating a substitution reaction (S_N2) with an inverse configuration. The 5′-FDA **23** bound crystal structure showed that the fluoromethyl group of **23** makes contact with residues Phe-156, Tyr-77, Thr-80 and Ser-158 defining a hydrophobic pocket (Figure 6). Fluoride was observed to make contact with Ser-158 and Thr-80.[85] The side chain OH of the Ser-158 also forms a divergent hydrogen bond to 3′-OH of the ribose ring in the product 5′-FDA **23**. In addition, the ether oxygen does not interact with any residues and this may be beneficial to the reaction. Any hydrogen bonding to the ether oxygen from surface residues would increase the electron withdrawing power of the oxygen which would strengthen the C–S^+R_2 bond. This would increase the activation energy required for the reaction thus reducing the reaction rate. Therefore the geometry and disposition of SAM **22** binding appears to be optimised to accommodate this reaction.[83]

An S_N2 reaction mechanism was also supported by isotope labelling studies, probing stereochemistry. [$^2H_1$-5′]-SAM was prepared from (*R*)-[$^2H_1$-5′]-ATP by the action of SAM synthetase. This generated (*R*)-[$^2H_1$-5′]-FDA when incubated with fluorinase, with an overall retention of configuration. However since the SAM synthetase is known to mediate an inversion in configuration at C-5′ of ATP, in order to retain the overall configuration in the resultant [$^2H_1$-5′]-FDA, the fluorinase must also be carrying out substitution with an inversion of configuration.[88] This is consistent with a S_N2 substitution process. S_N2 reactions are uncommon in enzymology and are usually related with low-activation energy barriers, requiring substitution reactions such as methyl transfer reactions.[64] QM/MM studies have indicated that the fluorinase is capable of lowering the activation energy required for the C-F bond formation relative to solution. [87]

The proposed molecular mechanism for the fluorinase suggests the hydrated fluoride ion diffuses into the hydrophobic active site with low affinity and interacts with the Ser-158, Phe-156 and Asp-16 residues *via* hydrogen bonds. The fluoride ion is desolvated within

the active site as no water molecules are observed in the X-ray structure of fluorinase co-crystallised with 5′-FDA **23** (Figure 5). In addition, examination of the SAM **22** bound crystal structure shows a 1.4-1.6 Å pocket, ideal for binding fluoride (1.3 Å ionic radius) rather than a halide with larger radius. This is consistent with the inability to process other halide ions.[85] Computational studies of the surface of fluorinase does not reveal any channels suitable for transporting fluoride to active site indicating that fluoride may bind first, followed by SAM **22**.[85,87]

These findings are further confirmed through QM/MM theory studies, where conformation of reactants (SAM **22** and F^-) and products (5′-FDA **23** and L-methionine **27**) in the pre and post formation of reactive complexes and transition states were analysed.[87]

Several mutants of the fluorinase were produced by site directed mutagenesis and the mutant proteins were expressed in *E.coli*. These mutant proteins were assayed for their activity relative to the wild type fluorinase. Residues chosen for mutations were assumed to be important for catalysis of the enzyme. The F156A mutant showed significant loss of activity which was mostly restored in F156V mutant suggesting that the hydrophobic pocket is involved in increasing the reactivity of fluoride ion. D16 mutants, D16A, D16S and D16N all showed reduced activity suggesting the aspartate residue interacting with the ribose ring plays a critical role in organising substrate for catalysis. The T80S mutant (OH group retained) was observed to retain its activity whereas T80A mutant (without the OH group) was observed to have a 10-fold reduction in activity suggesting the importance of the Thr-80 residue. In addition the S158A mutant showed a clear reduction in the enzyme activity due to the loss of hydrogen bond donors for fluoride in the active site suggesting the importance of Ser158 in F^- binding. Binding is also assisted by the electrostatic interactions provided by the sulphonium center of SAM **22** and the negatively charged fluoride. QM/MM calculations also supported the importance of hydrogen bonding F^- to Thr-80.[86]

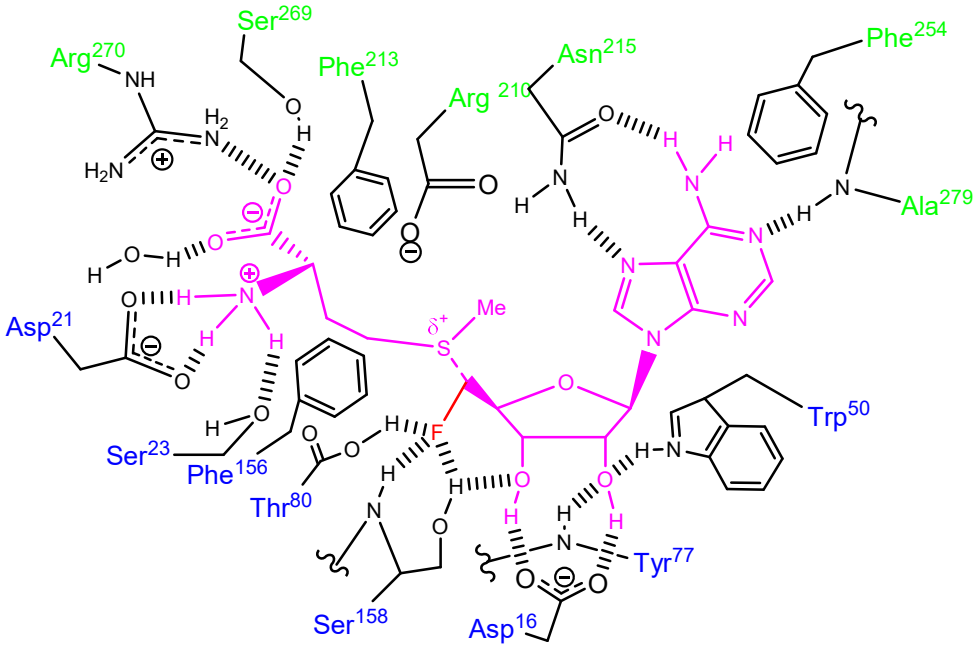


Figure 6: Interactions within the fluorinase active site with bound 5′-FDA **23** and L-methionine **27** (magenta). Residues that make up the active site belong to two monomers which are represented in blue and green. Image reproduced and modified with permission from Naismith *et al.* [85]

Isothermal calorimetry analysis was used to study the steady state kinetics of the fluorinase. The data suggest that during the forward reaction the fluoride ion binds prior to SAM **22** for catalytic turnover. Fluorinase has relatively high binding affinity for SAM **22** (K_m between 6.5 μM and 29μM) however, fluoride was shown to have low affinity for fluorinase (K_m > 5mM). Concentrations of SAM **22** higher than 20μM were shown to inhibit the enzyme despite it being a substrate. Fluorinase has a turn over k_{cat} of around 0.07 min^{-1} [85,86] which suggests it is a very slow enzyme. This low turnover is possibly related to the formation of the C-F bond under aqueous conditions.

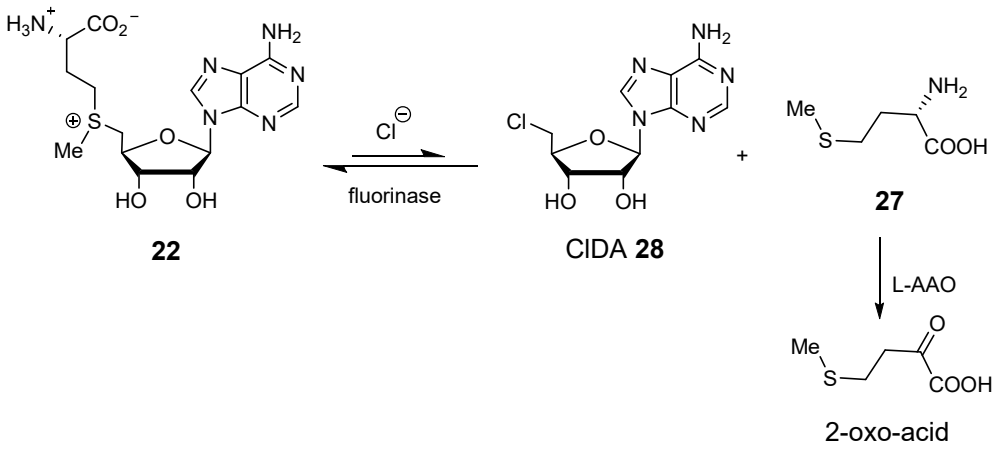
These data collectively suggest that the hydrated fluoride binds to an unoccupied hydrophobic pocket with low affinity. It exchanges water molecules for hydrogen bonds with Ser-158. The partial dehydration of fluoride drives the reaction forward and as SAM **22** binds with high affinity it promotes further dehydration of the fluoride ion trapping it for nucleophilic attack, while a hydrogen bond to Thr-80 further stabilises the fluoride ion. The positive charge on SAM **22** also stabilises the transition state. Nucleophilic F^- then attaches C-5′ of SAM **22** displacing L-methionine **27**. The L-Methionine **27** is released by the enzyme first, and then following a conformational change, releases the product, 5′-FDA **23**.[86]

Although various synthetic fluorinating procedures are available, biological fluorination is rare. Therefore the discovery of the fluorinase opens up a new avenue in biocatalysis. Biotransformations offer some advantages over synthetic methods as enzymes provide a handle in sustainable chemistry. One area in which C-F bond formation is important is in positron emission tomography (PET). The fluorinase was exploited for incorporating ^{18}F into compounds, which can then be used in PET imaging. This will be further discussed in Chapter 2.

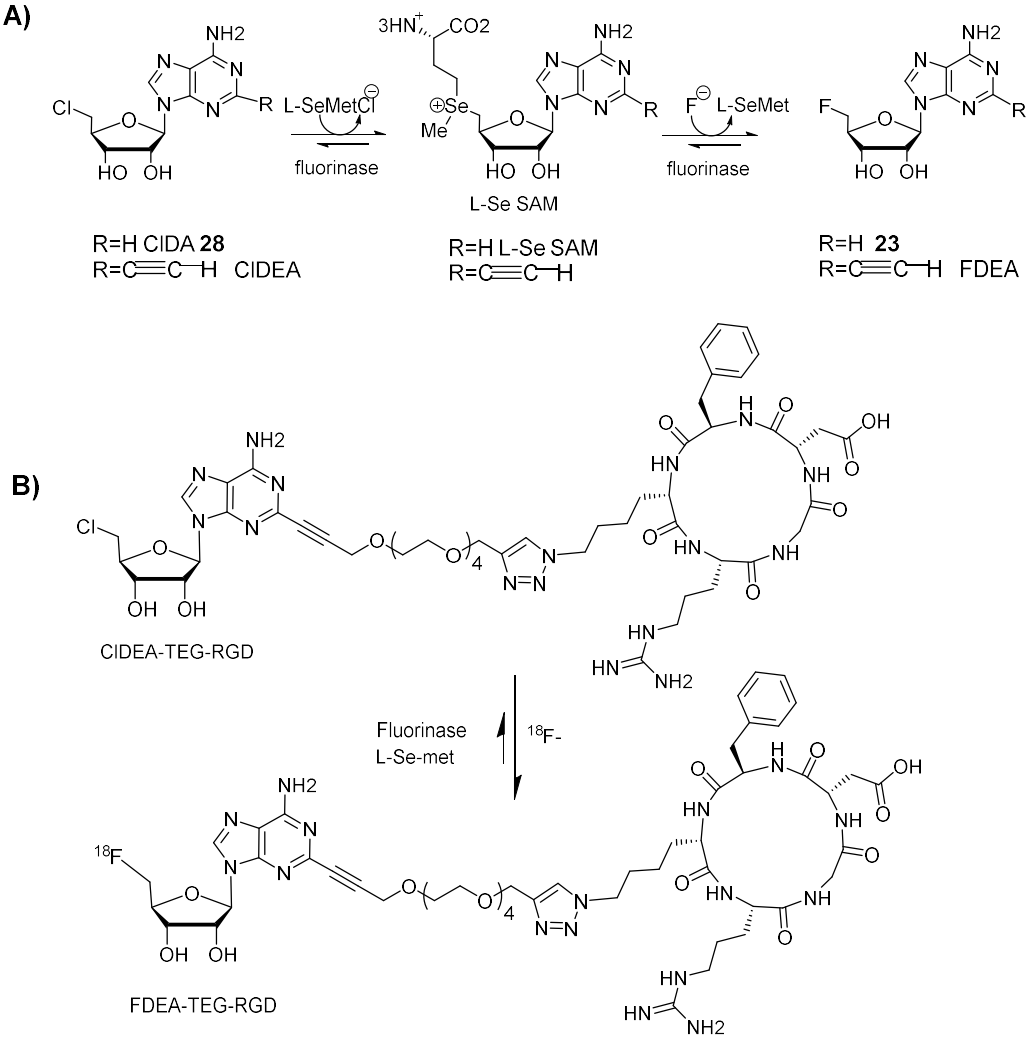
1.3.1.2 Reversibility of the fluorinase

The fluorinase was observed to have high substrate specificity and able to use SAM **22** to produce 5′-FDA **23**. When fluorinase was incubated with chloride and SAM **22** no conversion to the chlorinated product, 5′-chloro-5′-deoxy-adenosine (ClDA) **28** was observed.[89] Through coupled enzyme assays it was observed that the failure to produce ClDA **28** was due to the equilibrium of the chloride reaction significantly favouring the substrates over products. This reverse reaction was exploited in performing transhalogenation reactions using fluorinase to catalyse the progressive conversion of ClDA **28** to 5′-FDA **23** *via* SAM **22** intermediate.[89] Despite the high dissociation energy of C-F bond the fluorinase was able to cleave the C-F bond when incubated with 5′-FDA **23** and L-methionine **27** to form SAM **22** and fluoride. Addition of L-selenomethionine (L-Se-met), instead of L-methionine **27** exhibited a six-fold increase in the rate of the reaction in generating L-Se-SAM. ClDA **28** was shown to be a more efficient substrate for the fluorinase than 5′-FDA **23** for the generation of SAM **22**. This is because chloride is a better leaving group than fluoride, and fluoride is a better nucleophile (substrate) in the forward reaction.[90] Enzyme assays revealed that the apparent lack of chloride activity was due to the equilibrium of the reaction lying in considerable favour of substrates over products.[91]

The reversibility of the fluorinase reaction was investigated *via* several enzyme-coupled assays. The reaction was coupled to L-amino acid oxidase (L-AAO) enzyme which oxidises L-methionine **27** to 2-oxo-acid in an irreversible reaction. Incubation of these enzymes with SAM **22** and chloride ions generated detectable amounts of ClDA **28** as the backward reaction was supressed (Scheme 8). Production of ClDA **28** from SAM **22** and Cl^- was only detected in coupled enzyme assays which prevent the reverse reaction. Reactions using Br^- and I^- did not produce any halogenated products.[91]



Scheme 8: Fluorinase generates **28** when coupled to an L-AAO to remove **27** which drives the reaction forward.

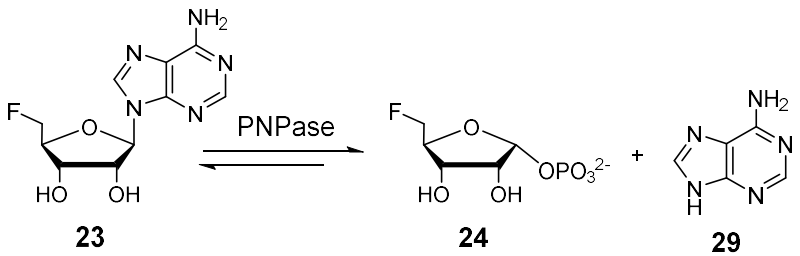


Scheme 9: A) Trans-halogenation reaction catalysed by fluorinase which converts ClDA **28** to 5′-FDA **23** through an L-Se-SAM intermediate. **B)** ClDEA-TEG-RGD converted to FDEA-TEG-RGD by trans-halogenation reaction catalysed by fluorinase. Image reproduced and modified with permission from O'Hagan *et al.*[92]

This reversibility of the fluorinase reaction was exploited in a trans-halogenation reaction where ClDA **28** was incubated with fluoride ions, L-Se-Met and fluorinase. The fluorinase combines L-Se-Met and ClDA **28** to form L-Se-SAM. This intermediate acts as a substrate for the fluorination reaction to generate 5′-FDA **23** (Scheme 9A).[91] Despite the high substrate specificity of fluorinase it was observed that incorporation of acetylene moiety at the C-2 position of the adenosine ring did not hinder catalysis. Therefore substrate analogues such as 5'-chloro-5'-deoxy-2-ethynyladenosine (ClDEA) were explored in trans-halogenation reactions which has led to production of novel fluorinated products such as 5'-fluoro-5'-deoxy-2-ethynyladenosine (FDEA) (Scheme 9A). FDEA was further explored for its ability to retain a polyethylene glycol (PEG) chain as an extension of the acetylene terminus which can be used to tether peptides which could be radiolabelled with fluorine-18. A ClDEA with a tetraethylene glycol (TEG) linker at the C-2 position of the acetylene substituent was designed that could be linked to an arginyl-glycyl-aspartic acid (RGD) peptide. This was incubated with $^{18}F^-$ and L-Se-met to generate the ^{18}F labelled FDEA-TEG-RGD in reasonable radiochemical yield (Scheme 9B) and this was further explored in its ability to bind a subset of integrins [92] for studying diseases related to cell adhesion such as cancer, osteoporosis and thrombosis.[93]

1.3.2 Purine nucleoside phosphorylase

The second enzyme in the fluorometabolite pathway of *S. cattleya* is a PNP.[94,95] It catalyses the phosphate displacement of the adenine base at the anomeric carbon of 5′-FDA **23** to form 5-FDRP **24** and adenine **29**. Studies on the PNP catalytic mechanism have shown that the PNP cleaves the glycosidic bond with inversion of configuration to produce the α-anomer of the phosphorylated sugar. The PNP found in *S. cattleya* is similar to other PNPs which carry out the normal degradation of purine nucleosides. Studies carried out in CFE's of *S. cattleya* and partially purified PNP from the CFE demonstrated the production of 5-FDRP **24** from 5′-FDA **23** (Scheme 10).[74]

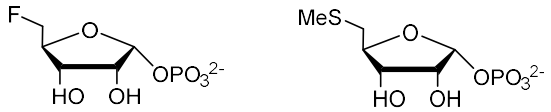


23 **24** **29**

Scheme 10: PNP enzyme from *S. catttleya* catalyses the cleavage of C-N glycosidic bond of 5′-FDA **23** to produce 5-FDRP **24** and adenine **29**.

1.3.3 Isomerase

The product of the PNP reaction, 5-FDRP **24** structurally resembles *S*-methyl-5-thioribose-1-phosphate (MTRP) **30**[96] (Figure 7) which is a well-known intermediate in the L-methionine **27** salvage pathway of yeast [97] and bacteria.[98] The L-methionine **27** salvage pathway is responsible for regenerating L-methionine **27** from the waste by-products of SAM **22** metabolism in order to recycle sulphur within the cell.[99] MTRP **30** is the substrate for a 5-deoxy-5-methylthioribose 1-phosphate isomerase (MTRPi) which opens the hemiacetal group of the ribose ring into an acyclic keto-sugar.[96] This offered a model for the metabolism of 5-FDRP **24** as a substrate and the expected product was proposed to be 5-FDRulP **25**.



5-FDRP **24** MTRP **30**

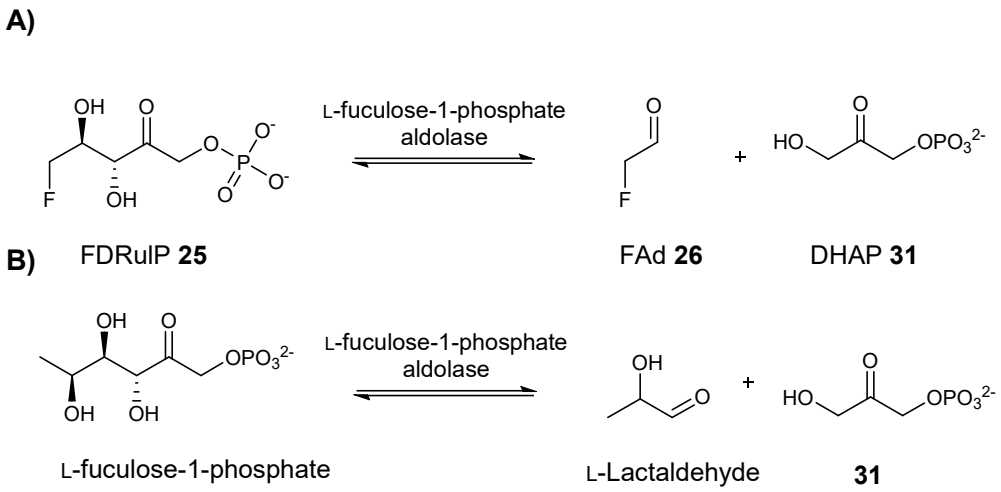
Figure 7: 5-FDRP **24**, an intermediate in the fluorometabolite pathway structurally resembles MTRP **30** (intermediate of the methionine salvage pathway).

A CFE of *S. cattleya* incubated with 5-FDRP **24** was monitored by ^{19}F NMR, and this showed the production of 5-FDRulP **25** which was unambiguously identified by GC-MS, following derivatisation. These experiments proved that 5-FDRulP **25** was a biosynthetic intermediate in fluorometabolite biosynthesis. When 5-FDRP **24** was incubated with the *S.coelicolor* MTRPi, the production of 5-FDRulP **25** was observed.This supports the hypothesis for the presence of a similar homologue in *S. cattleya*.[75]

An isomerase was purified from the CFE of *S.cattleya* was indeed able to convert 5-FDRP **24** to 5-FDRulP **25**. The gene was identified, over-expressed in *E.coli* and purified to obtain a 42 kDa isomerase [75,76] Chapter 4 of this thesis focus on solving the crystal structure of this *S. cattleya* isomerase (5-FDRPi).

1.3.4 Aldolase

Production of 5-FDRulP **25** by the isomerase in *S. cattleya* suggested a biosynthetic route to **26**, as ribulose phosphates are known to be common products/substrates of DHAP **31** dependent aldolase enzymes.[100] Thus a ribulose aldolase in *S. cattleya* would be sufficient to convert 5-FDRulP **25** to **26** and **31**. [75,76]



Scheme 11: **A)** A surrogate L-fuculose-1-phosphate aldolase from *S. coelicolor* was capable of catalysing the retro-aldol reaction to produce **26** and **31**. **B)** Natural substrate for L-fuculose-1-phosphate aldolase is L-fuculose-1-phosphate [75]

When **26** and **31** were incubated with CFE of *S. cattleya*, two diastereoisomeric aldolase products were isolated suggesting the presence of two aldolase activities. Fractionation of the CFE resulted in the isolation of a Type II Zn^{2+}-dependent fructose aldolase which catalyses the conversion of **26** and **31** to an alternative diastereoisomer, 5-fluorodeoxyxylulose phosphate in the reverse reaction. In addition a fuculose-1-phosphate aldolase was also identified which is important in fructose and mannose metabolism.

In order to perform an *in vitro* reconstruction of fluorometabolite biosynthesis, an alternative fuculose aldolase from *Streptomyces coelicolor* was isolated and over-expressed in *E.coli*. This aldolase was able to catalyse the efficient cleavage of **25** to **26**, (Scheme 11A) although the natural substrate of this enzyme is L-fuculose-1-phosphate (Scheme 11B). To this date the aldolase involved in the fluorometabolite pathway in *S.cattleya* has not been identified or characterised. A putative fuculose aldolase gene is observed in the genome of *S.cattleya*, however this putative aldolase has not been over-expressed as of yet. [101]

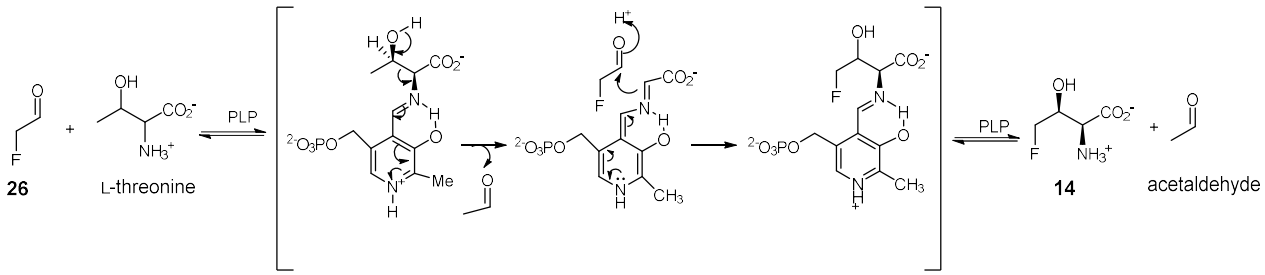
1.3.5 PLP-dependent transaldolase

FAd **26** is a branch point metabolite proceeded by two enzymes to form the two fluorometabolites, FAc **10** and 4-FT **14**. A transladolase generates 4-FT **14** whereas an aldehyde dehydrogenase generates FAc **10**. The transaldolase is a PLP dependent enzyme, which uses L-threonine and **26**. In 1986 Sanada *et al*., suggested that **14** is possibly formed through a direct condensation between **26** and PLP activated glycine

following common L-threonine aldolase reactions.[47] Surprisingly however glycine is not the substrate for this transaldolase enzyme. The data suggest that 4-FT **14** transaldolase has a mechanism according to Scheme 12 where PLP activates L-threonine and then undergoes a retro-aldol reaction, forming a glycyl-PLP aldimine followed by the aldimine attacking FAd **26**. Thus for every molecule of 4-FT **14** formed a molecule of L-threonine is sacrificed. *S.cattleya* produces 4-FT **14** at the expense of depleting the L-threonine pool suggesting that this interplay is crucial in increasing its fitness within its ecological niche.[64,76,102]

Subsequently the transaldolase enzyme was purified and was shown to exist as a dimer of around 120 kDa. Investigations into the gene encoding for PLP- transaldolase (*flFT*) showed that it has 40% homology to L-serine hydroxymethyl transferase (SHMT) from other bacteria. The transaldolase was found to consist of two domains, a PLP-transaldolase and a sugar phosphate epimerase. The larger domain (440 aa) was shown to be homologous to a PLP-binding domain of SHMT enzymes (~35% identity). The smaller domain (145 aa) was found to have homology with a phosphate binding domain of bacterial ribulose-1-phosphate-4-epimerases or L-fuculose-aldolases. These two domains are connected by a 35 amino acid chain, which does not have any obvious function.[76]

The transaldolase gene was introduced, induced and over expressed in *Streptomyces lividans* and these cultures were incubated with L-threonine, PLP and **26**. The 4-FT **14** production of these cultures was monitored by ^{19}F NMR and this confirmed 4-FT **14** transaldolase as the last enzyme in the 4-FT **14** biosynthetic pathway of *S. cattleya*.[76] 4-FT **14** biosynthesis was reconstructed *in vitro* by over-expressing the five enzymes identified in the biosynthetic pathway and incubating these enzymes together with fluoride and SAM **22**. The reaction progress was followed by ^{19}F NMR and confirmed the biosynthetic pathway as shown in Scheme 5.[76]



Scheme 12: Mechanism of PLP mediated 4FT-transaldolase which catalyses a retro aldol reaction on the PLP activated L-threonine forming glycyl-PLP aldimine which in turn attacks **26** to form **14** and acetaldehyde.[76]

1.3.6 Aldehyde dehydrogenase

Aldehyde dehydrogenase was the first enzyme to be identified on the fluorometabolite pathway of *S. cattleya*. It is an NAD^+ dependent aldehyde dehydrogenase which can catalyse the oxidation of **26** to **10**. This aldehyde dehydrogenase was able to oxidise a variety of aldehydes however the most efficient transformations were reported from **26** and glycolaldehyde. This indicates that the electronic factors maybe more important for reactivity than steric properties. When chloroacetaldehyde was tested as a substrate for the aldehyde dehydrogenase the rate of the reaction decreased significantly. Incubation time with chloroacetaldehyde was increased and this reduced activity, suggesting chloroacetaldehyde might also be a time-dependent inhibitor. This is possibly due to alkylation of the active site thiol following nucleophilic attack by the sulphur to the chloromethyl group. Since fluoride is a poor leaving group this type of inhibition is not observed with **26**. As the aldehyde dehydrogenase is capable of accepting glycolaldehyde, it is possible that this enzyme evolved from an enzyme that used glycolaldehyde or a similar compound as its substrate.[77]

Chapter 2 attempts to develop a practical, user friendly, shelf stable enzyme kit using freeze-dried fluorinase to incorporate ^{18}F to generate [^{18}F]-FDR ([^{18}F]-**37**) for PET imaging of tumour induced mice.

1.4 Chapter 1 References

1 P M. Dewick, *Medicinal Natural Products*, John Wiley & Sons, Ltd, Chichester, UK, Second Edi., 2001, 7–34.

2 A. Aharoni and G. Galili, *Curr. Opin. Biotechnol.*, 2011, **22**, 239–244.

3 J. Davies and D. Davies, *Microbiol. Mol. Biol. Rev.*, 2010, **74**, 417–433.

4 C. Hertweck, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 4688–4716.

5 M. Bonfill, S. Malik, M. H. Mirjalili, M. Goleniowski, R. Cusido and J. Palazón, in *Natural Products*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, 2761–2796.

6 N. Porter and F. M. Fox, *Pestic. Sci.*, 1993, **39**, 161–168.

7 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.

8 C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson and B. Worm, *PLoS Biol.*, 2011, **9**, 1–8.

9 G. W. Gribble, *J. Chem. Educ.*, 2004, **81**, 1441–1449.

10 O. Tyc, C. Song, J. S. Dickschat, M. Vos and P. Garbeva, *Trends Microbiol.*, 2016, 1–13.

11 J. O'Brien and G. D. Wright, *Curr. Opin. Biotechnol.*, 2011, **22**, 552–558.

12 J. Bérdy, *J. Antibiot. (Tokyo).*, 2005, **58**, 1–26.

13 M. Watve, R. Tickoo, M. Jog and B. Bhole, *Arch. Microbiol.*, 2001, **176**, 386–390.

14 S. A. Waksman and H. B. Woodruff, *Exp. Biol. Med.*, 1942, **49**, 207–210.

15 A. Schatz, E. Bugle and S. A. Waksman, *Exp. Biol. Med.*, 1944, **55**, 66–69.

16 J. H. Comroe, *Am. Rev. Respir. Dis.*, 1978, **117**, 773–81.

17 J. D. Wassersug, *N. Engl. J. Med.*, 1946, **235**, 220–229.

18 D. A. Hopwood, *Streptomyces in Nature and Medicine: The Antibiotic Makers*, Oxford University Press, 2007.

19 R. E. de Lima Procópio, I. R. da Silva, M. K. Martins, J. L. de Azevedo and J. M. de Araújo, *Brazilian J. Infect. Dis.*, 2012, **16**, 466–471.

20 L. L. Ling, T. Schneider, A. J. Peoples, A. L. Spoering, I. Engels, B. P. Conlon, A. Mueller, T. F. Schäberle, D. E. Hughes, S. Epstein, M. Jones, L. Lazarides, V. A. Steadman, D. R. Cohen, C. R. Felix, K. A. Fetterman, W. P. Millett, A. G. Nitti, A. M. Zullo, C. Chen and K. Lewis, *Nature*, 2015, **517**, 455–459.

21 M. Ventura, C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater and D. van Sinderen, *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 495–548.

22 K. Flärdh and M. J. Buttner, *Nat. Rev. Microbiol.*, 2009, **7**, 36–49.

23 M. I. Cowan, A. T. Glen, S. A. Hutchinson, M. E. MacCartney, J. M. Mackintosh and A. M. Moss, *Trans. Br. Mycol. Soc.*, 1973, **60**, 347–351.

24 D. B. Harper, *Nature*, 1985, **315**, 55–57.

25 N. Ji, X. Li and B. Wang, *Molecules*, 2008, **13**, 2894–2899.

26 A. S. Eustáquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. a Reynolds and B. S. Moore, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 12295–12300.

27 A. Tooming-Klunderud, T. Rohrlack, K. Shalchian-Tabrizi, T. Kristensen and K. S. Jakobsen, *Microbiology*, 2007, **153**, 1382–1393.

28 F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3364–3378.

29 G. Laus, *Stud. Nat. Prod. Chem.*, 2001, **25**, 757–809.

30 R. S. Griffith, *Rev. Infect. Dis.*, 1981, **3**, 200–204.

31 T. Dairi, T. Nakano, K. Aisaka, R. Katsumata and M. Hasegawa, *Biosci. Biotechnol. Biochem.*, 1995, **59**, 1099–106.

32 J. Controulis, M. C. Rebstock and H. M. Crooks, *J. Am. Chem. Soc.*, 1949, **71**, 2463–2468.

33 J. A. Bush, B. H. Long, J. J. Catino, W. T. Bradner and K. Tomita, *J. Antibiot. (Tokyo).*, 1987, **40**, 668–678.

34 T. Suzuki, M. Suzuki, A. Furusaki, T. Matsumoto, A. Kato, Y. Imanaka and E. Kurosawa, *Tetrahedron Lett.*, 1985, **26**, 1329–1332.

35 R. W. Fuller, J. H. Cardellina, Y. Kato, L. S. Brinen, J. Clardy, K. M. Snader and M. R. Boyd, *J. Med. Chem.*, 1992, **35**, 3007–3011.

36 X. Liu, X. M. Li, C. S. Li, N. Y. Ji and B. G. Wang, *Chinese Chem. Lett.*, 2010, **21**, 1213–1215.

37 W. Maiese, M. P. Lechevalier, H. A. Lechevalier, J. Korshalla, N. Kuck, A. Fantini, M. J. Wildey, J. Thomas and M. Greenstein, *J. Antibiot. (Tokyo).*, 1989, **42**, 558–563.

38 S. F. Love, W. M. Maiese and D. M. Rothstein, *Appl. Environ. Microbiol.*, 1992, **58**,

1376–1378.

39 C. S. Neumann, D. G. Fujimori and C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99–109.

40 A. Earnshaw and N. N. Greenwood, in *Chemistry of the Elements*, Elsevier, 1997, 1294–1301.

41 A. Earnshaw and N. N. Greenwood, in *Chemistry of the Elements*, Elsevier, 1997, 789–887.

42 C. D. Murphy, C. Schaffrath and D. O'Hagan, *Chemosphere*, 2003, **52**, 455–461.

43 L. Ma, A. Bartholome, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng and D. O'Hagan, *Chem. Sci.*, 2015, **6**, 1414–1419.

44 M. S. Ayoup, D. B. Cordes, A. M. Z. Slawin and D. O'Hagan, *J. Nat. Prod.*, 2014, **77**, 1249–1251.

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

46 D. B. Harper and D. O'Hagan, *Nat. Prod. Rep.*, 1994, **11**, 123–133.

47 M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, *J. Antibiot. (Tokyo).*, 1986, **39**, 259–265.

48 M. W. van der Kamp, J. D. McGeagh and A. J. Mulholland, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 10349–10351.

49 H. L. Carrell, J. P. Glusker, J. J. Villafranca, A. S. Mildvan, R. J. Dummel and E. Kun, *Science*, 1970, **170**, 1412–1414.

50 H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13699–13703.

51 R. A. Peters, *Adv. Enzymol.*, 1957, **18**, 113–159.

52 F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T. L. Li, D. O'Hagan, P. F. Leadlay and J. B. Spencer, *Chem. Biol.*, 2006, **13**, 475–484.

53 M. C. Walker and M. C. Y. Chang, *Chem. Soc. Rev.*, 2014, **43**, 6527–6536.

54 M. V. B. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. V. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, *J. Biol. Chem.*, 2010, **285**, 22495–22504.

55 A. M. Weeks, S. M. Coyle, M. Jinek, J. A. Doudna and M. C. Y. Chang, *Biochemistry*, 2010, **49**, 9269–9279.

56 A. M. Weeks, N. S. Keddie, R. D. P. Wadoux, D. O'Hagan and M. C. Y. Chang, *Biochemistry*, 2014, **53**, 2053–2063.

57 A. M. Weeks and M. C. Y. Chang, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 19667–19672.

58 J. J. Marion Meyer, N. Grobbelaar, R. Vleggaar and A. I. Louw, *J. Plant Physiol.*, 1992, **139**, 369–372.

59 D. O'Hagan and D. B. Harper, *J. Fluor. Chem.*, 1999, **100**, 127–133.

60 P. F. V. Ward, R. J. Hall and R. A. Peters, *Nature*, 1964, **201**, 611–612.

61 J. T. G. Hamilton and D. B. Harper, *Phytochemistry*, 1997, **44**, 1129–1132.

62 C. W. Waller, J. B. Patrick, W. Fulmor and W. E. Meyer, *J. Am. Chem. Soc.*, 1957, **79**, 1011–1012.

63 S. O. Thomas, V. L. Singleton, J. A. Lowery, R. W. Sharpe, L. M. Pruess, J. N. Porter, J. H. Mowat and N. Bohonos, *Antibiot. Annu.*, 1957, 716–721.

64 K. K. J. Chan and D. O'Hagan, in *Methods in Enzymology*, 2012, 219–235.

65 R. A. Peters and M. Shorthouse, *Nature*, 1967, **216**, 80–81.

66 R. A. Peters and M. Shorthouse, *Nature*, 1971, **231**, 123–124.

67 S. Ozaki, Y. Watanabe, T. Hoshiko, H. Mizuno, I. K and M. Hori, *Chem. Pharm.Bull.*, 1984, **32**, 733–738.

68 N. Jaivel, C. Uvarani, R. Rajesh, D. Velmurugan and P. Marimuthu, *J. Nat. Prod.*, 2014, **77**, 2–8.

69 J. T. G. Hamilton, C. D. Murphy, M. R. Amin, D. O'Hagan and D. B. Harper, *J. Chem. Soc. Perkin Trans. 1*, 1998, 759–768.

70 J. T. G. Hamilton, M. R. Amin, D. B. Harper and D. O'Hagan, *Chem. Commun.*, 1997, 797–798.

71 J. Nieschalk, J. T. G. Hamilton, C. D. Murphy, D. B. Harper and D. O'Hagan, *Chem. Commun.*, 1997, 799–800.

72 T. Tamura, M. Wada, N. Esaki and K. Soda, *J. Bacteriol.*, 1995, **177**, 2265–2269.

73 S. J. Moss, C. D. Murphy, D. O'Hagan, C. Schaffrath, J. T. G. Hamilton, W. C. McRoberts and D. B. Harper, *Chem. Commun.*, 2000, 2281–2282.

74 S. L. Cobb, H. Deng, J. T. G. Hamilton, R. P. McGlinchey and D. O'Hagan, *Chem. Commun.*, 2004, 592–593.

75 M. Onega, R. P. McGlinchey, H. Deng, J. T. G. Hamilton and D. O'Hagan, *Bioorg. Chem.*, 2007, **35**, 375–385.

76 H. Deng, S. M. Cross, R. P. McGlinchey, J. T. G. Hamilton and D. O'Hagan, *Chem.*

Biol., 2008, **15**, 1268–1276.

77 C. D. Murphy, S. J. Moss and D. O'Hagan, *Appl. Environ. Microbiol.*, 2001, **67**, 4919–4921.

78 D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton and C. D. Murphy, *Nature*, 2002, **416**, 279.

79 D. L. Zechel, S. P. Reid, O. Nashiru, C. Mayer, D. Stoll, D. L. Jakeman, R. A. J. Warren and S. G. Withers, *J. Am. Chem. Soc.*, 2001, **123**, 4350–4351.

80 D. L. Zechel and S. G. Withers, *Curr. Opin. Chem. Biol.*, 2001, **5**, 643–649.

81 M. Sugimura, M. Nishimoto and M. Kitaoka, *Biosci. Biotechnol. Biochem.*, 2006, **70**, 1210–1217.

82 T. C. Chou and P. Talalay, *Biochemistry*, 1972, **11**, 1065–1073.

83 H. Deng, D. O'Hagan and C. Schaffrath, *Nat. Prod. Rep.*, 2004, **21**, 773–784.

84 C. Schaffrath, H. Deng and D. O'Hagan, *FEBS Lett.*, 2003, **547**, 111–114.

85 C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer and J. H. Naismith, *Nature*, 2004, **427**, 561–565.

86 X. Zhu, D. Robinson, A. R. McEwan, D. O'Hagan and J. H. Naismith, *J. Am. Chem. Soc.*, 2007, **129**, 14597–14604.

87 H. M. Senn, D. O'Hagan and W. Thiel, *J. Am. Chem. Soc.*, 2005, **127**, 13643–13655.

88 C. D. Cadicamo, J. Courtieu, H. Deng, A. Meddour and D. O'Hagan, *ChemBioChem*, 2004, **5**, 685–690.

89 H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, *Angew. Chemie Int. Ed.*, 2006, **45**, 759–762.

90 S. L. Cobb, H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan and D. a Robinson, *Org. Biomol. Chem.*, 2006, **4**, 1458–1460.

91 H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. Robinson and J. B. Spencer, *Angew. Chemie - Int. Ed.*, 2006, **45**, 759–762.

92 S. Thompson, Q. Zhang, M. Onega, S. McMahon, I. Fleming, S. Ashworth, J. H. Naismith, J. Passchier and D. O'Hagan, *Angew. Chemie Int. Ed.*, 2014, **53**, 8913–8918.

93 E. Ruoslahti, *Annu. Rev. Cell Dev. Biol.*, 1996, **12**, 697–715.

94 M. D. Erion, K. Takabayashi, H. B. Smith, J. Kessi, S. Wagner, S. Hönger, S. L. Shames and S. E. Ealick, *Biochemistry*, 1997, **36**, 11725–11734.

95 M. J. Pugmire and S. E. Ealick, *Biochem. J.*, 2002, **361**, 1–25.

96 F. R. Tabita, T. E. Hanson, H. Li, S. Satagopan, J. Singh and S. Chan, *Microbiol. Mol. Biol. Rev.*, 2007, **71**, 576–599.

97 I. Pirkov, J. Norbeck, L. Gustafsson and E. Albers, *FEBS J.*, 2008, **275**, 4111–4120.

98 Y. Dai, T. C. Pochapsky and R. H. Abeles, *Biochemistry*, 2001, **40**, 6379–6387.

99 E. Albers, *IUBMB Life*, 2009, **61**, 1132–1142.

100 C. Y. Lai and B. L. Horecker, *Essays Biochem.*, 1972, **8**, 149–178.

101 C. Zhao, P. Li, Z. Deng, H. Y. Ou, R. P. McGlinchey and D. O'Hagan, *Bioorg. Chem.*, 2012, **44**, 1–7.

102 C. D. Murphy, D. O'Hagan and C. Schaffrath, *Angew. Chem. Int. Ed. Engl.*, 2001, **40**, 4479–4481.

2. Production of radiotracer [^{18}F]-5-FDR [^{18}F]-37 using fluorinase

2.1 Positron emission tomography (PET)

Positron emission tomography (PET) is a molecular imaging tool used widely to image tumours in the clinical environment. There is a considerable demand for novel PET radiopharmaceuticals, which can be rapidly generated using clean synthetic procedures. Enzymatic methods are attractive due to their chemo-specificity, ability to perform reactions under aqueous conditions at a neutral pH with low generation of side-products. Identification of the fluorinase enzyme has opened up an enzymatic approach to incorporating ^{18}F into compounds for PET, as the enzyme utilises fluoride ions as the fluorine source.

2.1.1 PET as a molecular imaging tool

Molecular imaging is an emerging technology by which cellular processes and functions can be visualised *in vivo*. Anatomical techniques like computerised tomography (CT) and magnetic resonance imaging (MRI) are used in conjunction with other molecular imaging techniques to provide topographical information. Appropriate radioisotopes are used to label biological molecules and these are injected into the subject for imaging. Two of the most widely used molecular imaging techniques, are single photon emission tomography (SPECT) and PET[1]. SPECT uses gamma-emitting radioisotope such as ^{99m}Tc, ^{123}I and ^{111}In and the direct emission of gamma rays produced by these radiotracers[1] are observed, whereas PET uses positron-emitting radioisotopes, where the positron annihilates to form gamma photons. PET is at least tenfold more sensitive than SPECT and the positron-emitting isotopes (^{18}F, ^{13}N, ^{15}O, ^{11}C) can be substituted directly for atoms in bioactive molecules without affecting the chemical characteristics of the molecule. Therefore, PET is known to be a more versatile technique for imaging most molecular events and has become a major imaging tool in oncology, cardiology and neurology. [2]

The PET technique has good resolution, high sensitivity, accurate quantification[2,3] and provides quantitative information on: physiological, biochemical and pharmacological processes of living subjects. The sensitivity of PET is relatively high in the range of 10^{-11}-10^{-12} moles/L[4] therefore picomolar (10^{-12}) concentrations of radiotracers are administered to the subject. These radiotracers decay through the emission of a positron and as this encounters an electron from surrounding tissue it annihilates to produce two gamma 511keV photons that travel in opposite directions to each other (Figure 1A). These

gamma photons are detected by a circular array of gamma detectors (Figure 1B) and the registered events are reconstructed into a three dimensional image representing the spatial distribution of the radioactive source in the subject.[2,5] The spatial resolution of common PET scanners are ~ $(6\text{-}8)^3$ mm^3, however there are higher resolution brain scanners with ~3^3 mm^3 resolution.[6] For a PET scanner to visualize a clear signal against the background it requires a few hundred million cells in relatively close proximity to accumulate the radiotracer. PET images are usually displayed in colour reflecting gamma-ray events with the same energy, the colour scale reflects the concentration of isotope.[5]

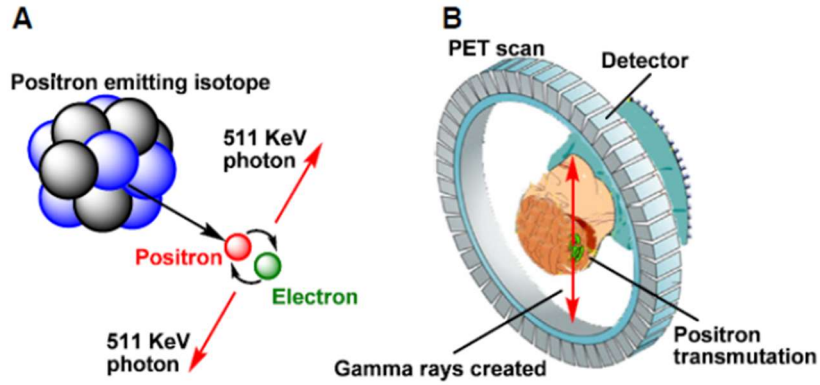


Figure 1: **A)** Two gamma rays travelling 180° from each other are produced as a positron (emitting from the radioisotope) and an electron annihilate. **B)** The 511KeV photons are registered by the circular gamma ray detector array in the PET camera. Image reproduced with permission from Conti *et al*.[2]

In PET the most widely used radioactive isotopes are ^{11}C, ^{13}N, ^{15}O and ^{18}F, as they have relatively short half-lives (20.4 min, 9.9 min, 2.0 min and 109 min respectively) and as they are elements which can be substituted for their stable counterparts in biological molecules and drugs. Due to the short half-life of the ^{11}C isotope multistep synthesis is not generally used for ^{11}C containing radiotracers. ^{13}N, ^{15}O are useful labelling isotopes as they are universal in biological molecules; however both have very short half-lives making it very difficult to perform more than one synthetic step and therefore are rarely used in synthesis of complex molecules. They are used to prepare simple radiolabelled molecules such as [^{15}O] CO_2, [^{15}O] H_2O, and [^{13}N] NH_3 which can be obtained directly from the cyclotron. The availability of such isotopes offers a wide range of possibilities for the investigation of various *in vivo* biological processes.[2]

^{18}F is the most widely used radioisotope for PET as it has the relatively long half-life of 109 min which allows time for synthesis, biological incorporation and imaging. The half-life of ^{18}F is long enough to allow for transportation of doses to sites a few hours away. In addition

^{18}F-labelled radiopharmaceuticals, can be produced in sufficient quantities for multiple injections from a single synthesis.[7]

^{18}F labelling for PET has some disadvantages; not all bioactives contain fluorine and thus fluorine tagging can alter the physiological response relative to the molecule of interest, as well as the synthetic methods to introduce ^{18}F into molecules are limited. The two main strategies for incorporating ^{18}F are either *via* direct or indirect fluorination. For direct fluorination the ^{18}F isotope is introduced into the target molecule in one step whereas indirect fluorination involves labelling a prosthetic groups such as ^{18}F-labelled alkyl or aryl moiety of a reactive molecule. Indirect methods are used when incorporating ^{18}F into complex biological molecules which might not withstand direct fluorination conditions.[7]

Direct methods involve electrophilic or nucleophilic fluorination. Electrophilic fluorination uses elemental fluorine [^{18}F]F_2 produced by deuteron irradiation of neon gas supplemented with F_2. [^{18}F]F_2 is highly reactive resulting in nonspecific reactions generating a mixture of products. This method also produces labelled products with low specific activity as the [^{18}F]F_2 is produced by a carrier-added method. Therefore currently electrophilic fluorination is an unfavoured method. However there are still some ^{18}F radiotracers such as [^{18}F]fluoro-L-DOPA and 2-L-[^{18}F]fluorotyrosine which depend on electrophilic fluorination methods for their preparation.[7,8]

Nucleophilic fluorination methods are much more widely used as they provide greater selectivity and higher specific activities. Radiotracers widely used in cancer detection, such as [^{18}F]-2-deoxy-2-fluoroglucose ([^{18}F]FDG) [^{18}F]-**32** and 3'-deoxy-3'-[^{18}F]fluorothymidine ([^{18}F]FLT)[9,10]; radiotracers used in dopamine studies, [^{18}F]fallypride,[11] [^{18}F]haloperidol[12] and [^{18}F]spiperone[13]; popular hypoxia imaging radiotracers such as [^{18}F]fluoroazomycin arabinose ([^{18}F]FAZA) and [^{18}F]Fluoromisonidazole ([^{18}F]FMISO) [14] are all prepared *via* nucleophilic fluorination (Table 1).

Nucleophilic $^{18}F^-$ is produced in a cyclotron by bombarding ^{18}O-enriched water ([^{18}O]H_2O) with high energy protons. This reaction generates a dilute solution of [^{18}F]F^- along with some unreacted [^{18}O]H_2O. In order to separate the [^{18}F]F^- from its aqueous environment the reaction is passed through an ion exchange column. $^{18}F^-$ produced is then trapped on the column and unreacted [^{18}O]H_2O is recovered.[7] The trapped $^{18}F^-$ is eluted in a water/acetonitrile solution containing potassium carbonate. The aqueous fluoride obtained is a poor nucleophile due to its high degree of solvation, therefore the production is followed by the removal of water and the phase-transfer reagent kryptofix-222 (K_{222}) is added to bind the potassium (K+) counter ion, improving the reactivity of [^{18}F] fluoride ion for nucleophilic substitution reactions.[7,8]

In order for radiotracers to be successful imaging probes for PET, they should fulfil the following requirements; the positron-emitting isotope should be chemically linked to the tracer so as not easily dissociate chemically or removed metabolically. Isotope labelling should not alter the behaviour or chemical properties of the parent molecule. The radiotracer should efflux from non-target sites and blood to achieve a good image contrast between the background and the target tissue which would increase the spatial resolution.[5,15]

2.1.2 [^{18}F]FDG [^{18}F]-**32** as PET radiotracer

The most common radiopharmaceutical, used in over 90% of PET imaging is [^{18}F]-2-deoxy-2-fluoroglucose ([^{18}F]-FDG) [^{18}F]-**32**, which is the fluorinated analogue of D-glucose where the 2-hydroxyl group has been replaced with ^{18}F.[8,16] [^{18}F]FDG [^{18}F]-**32** is used to analyse glucose utilization and metabolism of tumours, as they require a higher rate of glucose uptake, which leads to higher accumulation of [^{18}F]-**32**. When [^{18}F]-**32** is injected into a subject it is transported through the blood stream to cells, where specific glucose transporters (GLUT 1) recognise these molecules and transport [^{18}F]-**32** across the cell membrane.[17] Usually glucose molecules taken up this way are processed through several enzymes in the glycolysis pathway to form pyruvate which enters Krebs's cycle and plays a key role in cellular respiration.

[^{18}F]FDG [^{18}F]-**32** also enters the glycolysis pathway and is phosphorylated at C-6, by hexokinase to generate FDG-6-phosphate **33**. Unlike glucose however, [^{18}F]-**32** lacks the 2′-hydroxyl group and therefore, cannot act as a substrate for the second enzyme in the pathway, glucose-6-phosphate isomerase. Efflux of the phosphorylated [^{18}F]FDG-6-P **33** from the cell is low, thus [^{18}F]FDG [^{18}F]-**32** accumulation can be measured by PET (Figure 2). In addition [^{18}F]-**32** is removed through the renal system and unlike glucose, is not reabsorbed well in renal tubules, which leads to reduced levels of [^{18}F]-**32** in the blood. This provides a high contrast between the background and the target of interest being imaged. Tumours are associated with high energy demands, a decrease in glucose-6-phosphate and upregulation of glucose transporters (GLUT-1) and hexokinase which promotes most cancer cells to actively accumulate [^{18}F]FDG [^{18}F]-**32**.[18]

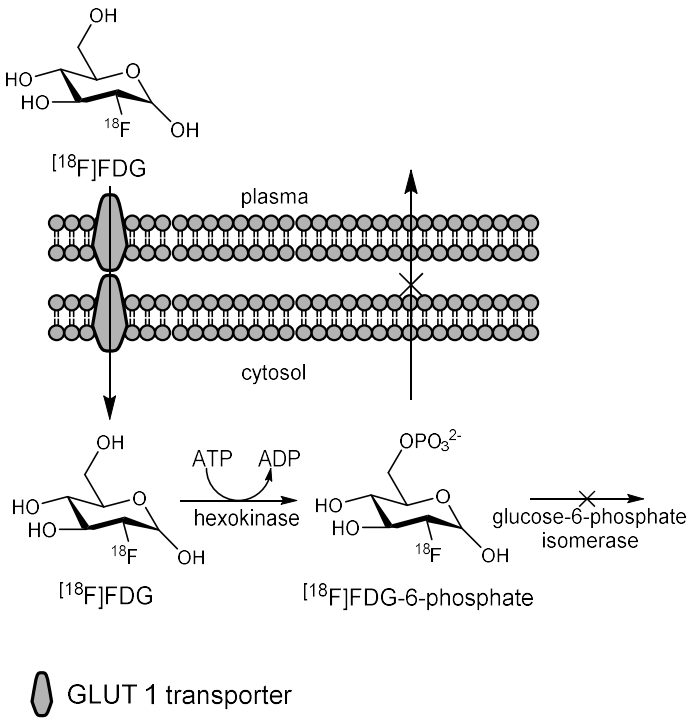
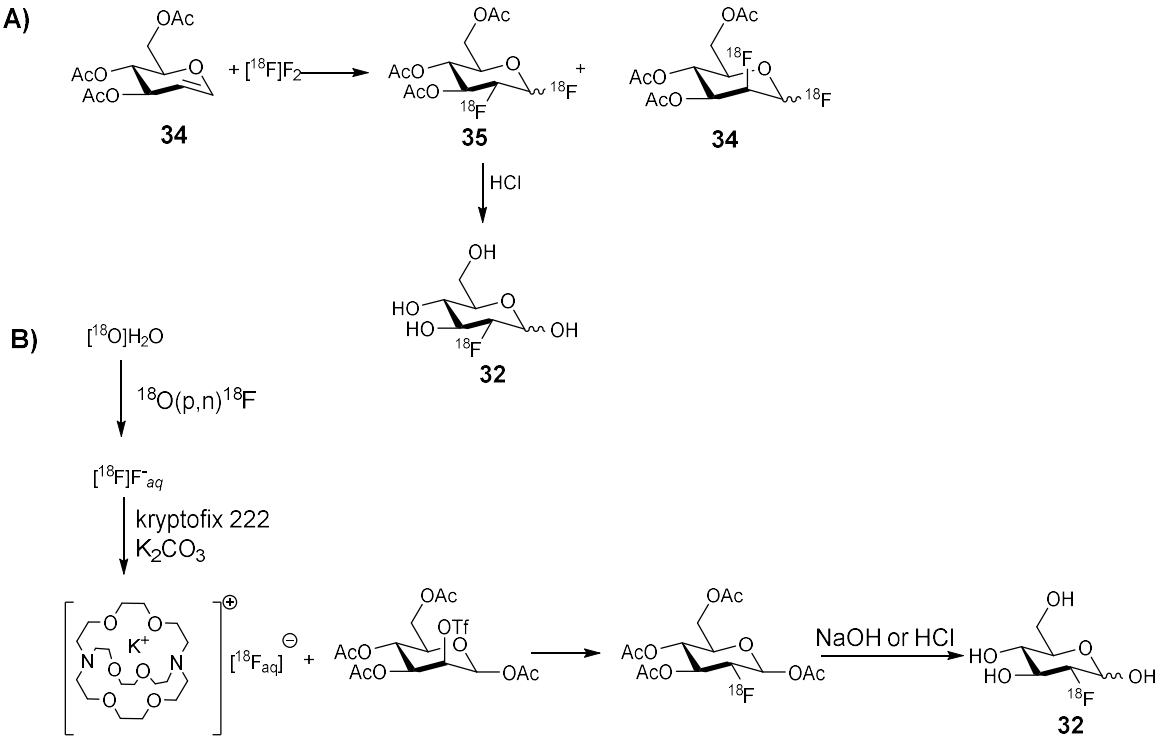


Figure 2: GLUT1 transporters transport [^{18}F]FDG **32** into the cell cytosol which is phosphorylated by hexokinase forming [^{18}F]FDG-6-phosphate **33**. This gets trapped inside the cell and is exploited in PET studies.

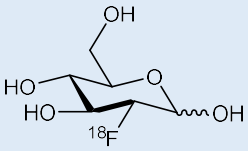
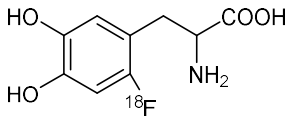
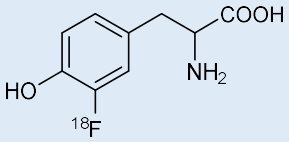
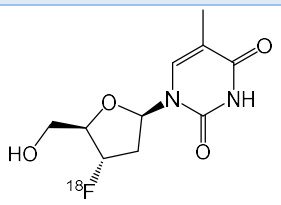
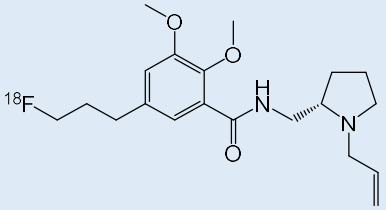
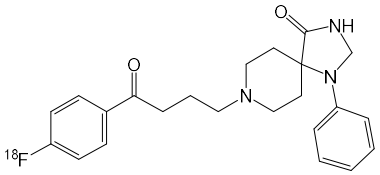
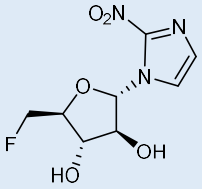
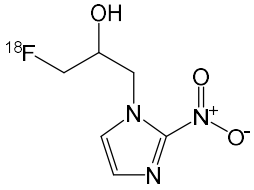
[^{18}F]FDG [^{18}F]-**32** can be synthesised by either electrophilic fluorination or nucleophilic fluorination. The first synthesis of [^{18}F]-**32** was based on direct electrophilic substitution. 3,4,6-Tri-O-acetyl-D-glucal **34** was used as a precursor and treated with [^{18}F]F_2 to produce ^{18}F-labelled difluoro-glucose **35** with mannose derivatives. The ^{18}F-labelled difluoro-glucose **35** was then hydrolysed with HCl to form [^{18}F]-**32** (Scheme 1A) with a yield of ~8% in a 2 h synthesis. In order to increase the yield and reduce the synthesis time, attempts were made to develop a nucleophilic approach for [^{18}F]FDG [^{18}F]-**32** synthesis.[19] This was achieved in 1986 when Hamacher *et al*. reported a method by which Kryptofix 222 was used to increase the reaction yield to over 50% and in under 50 min reaction time. The method involves fluoride displacement of triflates (Scheme 1B).[20] To date [^{18}F]FDG [^{18}F]-**32** is synthesised using the method described by Hamacher *et.al.*[20,21]



Scheme 1: A) [^{18}F]FDG [^{18}F]-**32** synthesis by electrophilic fluorination using [^{18}F]fluoride ion. **B)** [^{18}F]FDG [^{18}F]-**32** synthesis by nucleophilic fluorination using kryptofix222 as a catalyst.

The success of [^{18}F]FDG [^{18}F]-**32** in PET has encouraged further research on the incorporation of ^{18}F into novel tracers. Some of the currently known fluorinated PET radiotracers are listed in Table 1.[21]

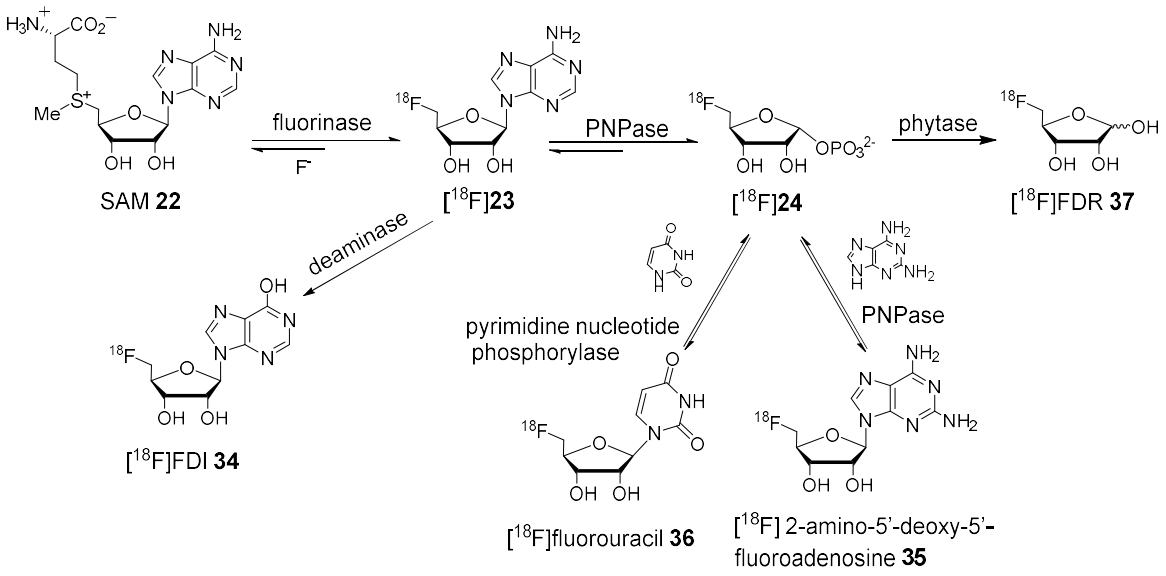
Table 1: Some [^{18}F]-labelled PET radiotracers produced from [^{18}F]F^{-}. [21]

Radiotracer	Structure	PET indicator for:
[^{18}F]-2-deoxy-2-fluoroglucose ([^{18}F]-FDG)	 **32**	Glucose utilisation
[^{18}F]fluoro-L-DOPA		Dopamine metabolism in the brain
2-L-[^{18}F] fluorotyrosine		Cerebral protein synthesis
3'-deoxy-3'-[^{18}F] fluorothymidine ([^{18}F]FLT)		Cell proliferation in tumours
[^{18}F] fallypride		Brain D_2/D_3 receptor
[^{18}F] spiperone		Brain D_2 dopamine and 5-hydroxytryptamine (serotonin) receptors
[^{18}F] fluoroazomycin arabinose ([^{18}F]FAZA)		Hypoxia
[^{18}F] Fluoromisonidazole ([^{18}F]FMISO)		Hypoxia

2.2 Application of the fluorinase for producing [^{18}F]-labelled PET radiotracers

A current challenge in the ^{18}F-labelled PET radiopharmaceuticals industry is to develop efficient chemistry for the rapid and clean synthesis, along with straightforward purification protocols to produce PET radiotracers in an appropriate time scale. Enzymatic methods are in principle attractive due to their chemoselectivity and low generation of side-products. Enzymes enable fast chemical conversions yielding highly pure products under mild aqueous conditions. Until now, there are only a few examples in PET radiotracer synthesis where enzymes have been used to introduce isotopically-labelled atoms, mostly because enzymes are substrate specific and finding suitable enzymes for a particular reaction maybe difficult in addition to having stability issues.[22–24] There are few examples where enzymes have been used for synthesis of different ^{11}C [25,26] and ^{13}N [27] labelled radiotracers. The identification of the fluorinase enzyme has opened up an enzymatic approach to incorporating ^{18}F into compounds.

It has been reported that ^{18}F-labelled fluorodeoxyadenosine analogues have potential as tumour imaging radiotracers[28] therefore, the use of [^{18}F]-FDA ([^{18}F]-**23**) as a potential PET radiopharmaceutical was considered.[29] In 2003 the first radiolabelling study with the fluorinase was carried out for the production of [^{18}F]-**23**.[29] A partially purified fluorinase fraction from *S.cattleya* was incubated at 40 °C with [^{18}F]F^- and SAM-Cl **22** to form [^{18}F]-**23**. The reaction time was 5 h and the radiochemical yield (RCY) was ~1%, clearly too inefficient for a PET radiotracer synthesis. Subsequently, realisation of the revers ibility of the fluorinase reaction lead to optimisation where various enzymes were used in coupled reactions to obtain novel [^{18}F]-labelled purine nucleosides. Over-expressed fluorinase was coupled to an L-AAO which actively removes L-methionine **27** and supresses the reverse reaction to produce [^{18}F]-**23**. This protocol increased the rate of [^{18}F]-**23** synthesis and conversion up to a 95% RCY (Scheme 2).



Scheme 2: Fluorinase mediated ^{18}F labelling, coupled to various enzymes to produce ^{18}F-labelled compounds.[30,31]

Furthermore various [^{18}F]-nucleosides were synthesised by coupling the fluorinase to appropriate nucleoside phosphorylases (mediating base swap reactions). In the biosynthetic pathway of *S. cattleya* (Scheme 5 Chapter 1), 5′-FDA **23** is converted to 5-FDRP **24** by a PNP in a reversible reaction. The reversibility of this reaction was exploited for its ability to 'swap' adenine **29** with another purine base (Scheme 2). 5-FDRP **24** produced by PNP was incubated with 2, 6-diaminopurine to produce the fluorinated nucleoside, [^{18}F] labelled 2-amino-5′deoxy-5′-fluoroadenosine **35**. 5′-Fluorinated uridine derivatives such as fluorouracil **36** were produced by coupling the fluorinase reaction to a pyrimidine nucleoside phosphorylase isolated from *Bacillus stearothermophilus*. This enzyme catalyses the reversible phosphorolysis of uridine and thymidine by displacing the pyrimidine base. These experiments illustrate the use of the fluorinase and the versatility of these base swap bio-transformations to generate a variety of 5'-deoxy-5'-fluoronucleosides [32,33].

The fluorinase reaction was also coupled to an adenylate deaminase isolated from *Aspergillus* sp. to produce [^{18}F]-5′-fluoro-5′deoxyinosine ([^{18}F]-5′-FDI) **34** as shown in (Scheme 2). This provided access to another labelled purine nucleoside. In addition coupling of the fluorinase to a PNP and then a phytase, which catalyse a depurination and hydrolysis of phosphate, respectively generated [^{18}F]-5-fluoro-5-deoxy-D-ribose ([^{18}F]5-FDR [^{18}F]-**37**). All of these reactions produced around 30-40% RCY within 1-4 h time scale. [31,32,34]

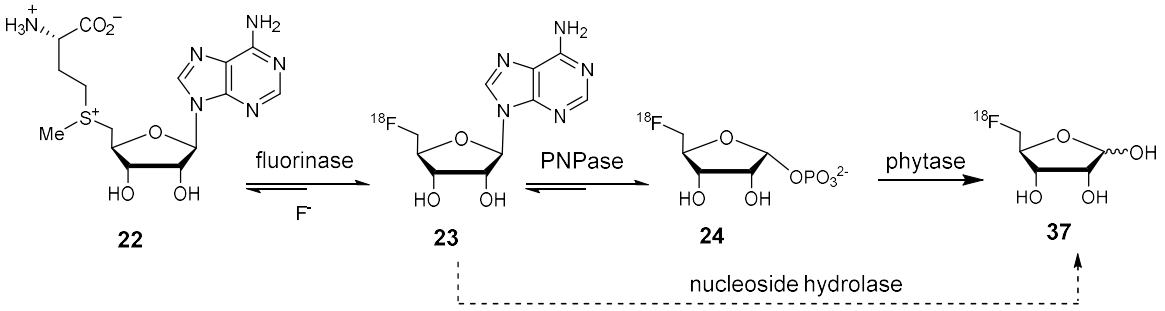
[^{18}F]-FDA [^{18}F]-**23** was explored for its role as a potential PET radiotracer by incubating with human breast cancer cells (MDA-MB-453). [^{18}F]-FDA [^{18}F]-**23** showed active uptake by cancer cells, and addition of cold adenosine **47** decreased [^{18}F]-**23** uptake significantly, suggesting that [^{18}F]-**23** uptake is probably facilitated by adenosine transporters.[34] Comparison of the influx and efflux data showed that only about 60% [^{18}F]-**23** remained inside the cancer cells.[29,30,34] However small animal PET studies demonstrated that [^{18}F]-**23** experiences rapid defluorination to release [^{18}F]F$^-$ *in vivo* which binds avidly to bone. Accumulation of radioactivity in bone disrupts the spatial resolution of PET which would make imaging inaccurate. Therefore [^{18}F]-**23** was ruled out as an unsuitable radiotracer for PET.[35]

In addition to the above mentioned ^{18}F-labelled compounds (Scheme 2), [^{18}F]FAc [^{18}F]-**10** was explored as a radiotracer with selective uptake and metabolism in glial neurones in the brain.[36] [^{18}F]FAc could be used as an acetate analogue in place of [^{11}C] acetate, which has much shorter half-life, to observe acetate metabolism in prostate tumours.[37,38] A chemo-enzymatic synthesis was designed for the preparation of [^{18}F]-**23**. Fluorinase was incubated with SAM-Cl **22** and [^{18}F]F$^-$ to produce [^{18}F]-**23**, a product, which was then chemically oxidised with a solution of chromium trioxide and sulfuric acid to form [^{18}F]-**10**. The tracer was obtained as the sodium salt in high RCY.[38]

Cell uptake studies in human breast cancer cells (MDA-MB-453) done previously have demonstrated almost 10 fold higher uptake of [^{18}F]5-FDR ([^{18}F]-**37**) in comparison to [^{18}F]-FDG ([^{18}F]-**32**). [31] Therefore the prospects of producing [^{18}F]-**37** in sufficient quantities and purity for PET were explored.

2.2.1 Shortened enzymatic route to produce [^{18}F]-5-FDR [^{18}F]-37

The three-step enzyme route coupling the fluorinase to a purine nucleoside phosphorylase (PNP) and then a phytase produced [^{18}F]5-FDR [^{18}F]-**37** with a ~45% RCY in 4 h. This was an inefficient route to produce [^{18}F]-**37**, and an alternative shorter route was considered. An alternative approach to synthesise [^{18}F]-**37** was proposed, by coupling the fluorinase reaction to an enzyme which acts upon [^{18}F]-**23** directly by hydrolysing the N-glycosidic bond to generate [^{18}F]-**37** and adenine **29** in a two-step route as shown in Scheme 3.[30,31]



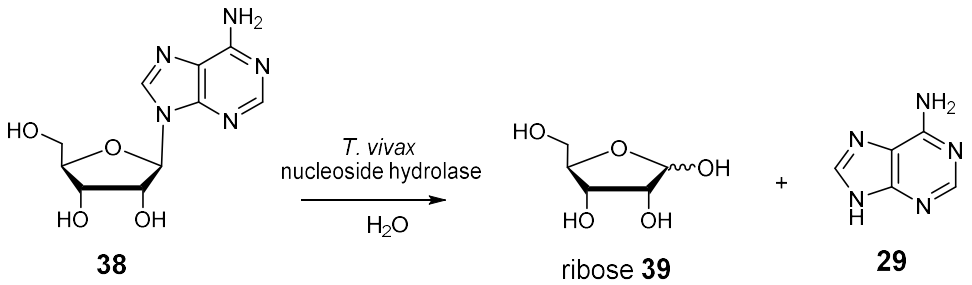
Scheme 3: Three enzyme route reduced to two enzyme route as *Tv*NH catalyses the hydrolysis of adenine **29** from [^{18}F]-**23** to form [^{18}F]-**37**.[31]

The two enzyme routes to produce [^{18}F]-**37** required the fluorinase to be coupled to a nucleoside hydrolase, isolated from the parasitic protozoa, *Trypanosoma vivax* (*Tv*NH).[31] *Tv*NH hydrolyses the N-glycosidic bond of inosine, adenosine **47** and guanosine most efficiently and therefore was considered to catalyse the direct conversion of **23** to **37**, substituting the PNP and phytase enzymes of the three step route.[31]

2.2.2. Nucleoside hydrolase

Nucleoside hydrolases (NHs) are a group of enzymes that catalyse the hydrolysis of the N-glycosidic bond between the anomeric carbon of ribose **39** and the purine or pyrimidine base of β-ribonucleosides forming the free nucleic acid base and ribose **39** (Scheme 4). These enzymes are widely distributed across many organisms such as: bacteria [39–41], protozoa[42,43], yeast[44], insects[45], plants[46,47] and archaea [48,49] but never in mammals. Therefore, these enzymes are attractive targets for drug design against parasitic protozoa[50] and several parasitic protozoan NHs have been studied intensively.

All parasitic protozoa are unable to synthesise purine nucleotides *de novo*, therefore salvage enzymes are used to obtain purine bases and nucleosides from their hosts. NHs have high specificity for the ribose **39** moiety, however they show varying preferences to different nucleoside bases. Nucleoside Hydrolases are divided to four classes according to their substrate specificity. Base non-specific NHs which hydrolyse inosine and uridine (IU-NHs) [42,43]; purine-specific inosine/adenosine **47**/guanosine NHs (IAG-NH) [51]; the 6-oxopurine-specific guanosine/inosine NHs[52]; and the pyrimidine nucleoside-specific cytidine/uridine NHs[53].



Scheme 4: *TvNH* catalyses the irreversible hydrolysis of the N-glycosidic bond of β-ribonucleoside **38** to form adenine **29** and ribose **39**.[54]

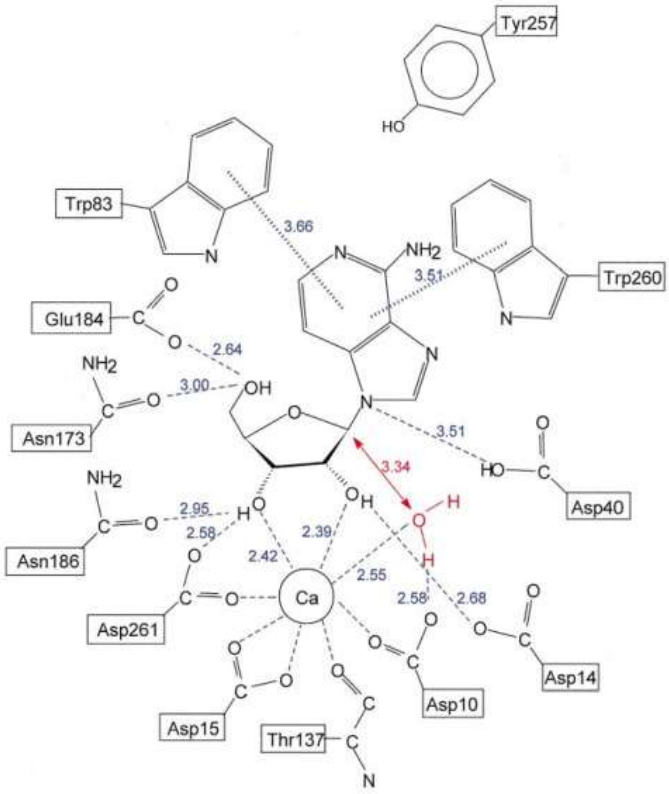


Figure 3: Schematic representation of the active site of *Tv*NH bound to 3-deaza-adenosine. The catalytic water is indicated in red. Image reproduced and modified with permission from Steyaert et al. [55]

Trypanasoma vivax nucleoside hydrolase (*TvNH*) is a key enzyme involved in the purine salvage pathway of *Trypanosoma vivax*.[55] It belongs to the IAG-NH class of enzymes and thus prefer inosine, adenosine **47** and guanosine nucleosides. *TvNH* has K_m of 5μM for inosine, 8.5 μM for adenosine **47**, and 3.8 μM for guanosine whereas K_m for pyrimidines such as cytidine and uridine are above 500μM demonstrating the high affinity for purines.[55]

The crystal structure of *TvNH* demonstrate that it exists as a homodimer with a hydrophobic interface and has two active sites which function independently as demonstrated by kinetic assays which showed no co-cooperativity in binding. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis confirmed the molecular mass of *TvNH* monomer to be ~37000 Da. The co-crystallised structure of *TvNH* with 3-deaza-adenosine revealed an octa-coordinated calcium ion bound to the lower part of the active

site which interacts with the 2′OH and 3′OH groups of the ribose ring (Figure 3). *TvNH* is believed to catalyse the N-glycosidic bond cleavage by a direct displacement mechanism. The Asp-10 from a conserved motif acts as a general base and abstracts a proton from a catalytic water molecule (shown in red in Figure 3), which then attacks the C1′ atom of the ribose moiety of the nucleoside. Ca^{2+} is thought to be involved in lowering the pKa of the water molecule prior to proton transfer to Asp-10. Subsequently Asp-40 donates a proton to N-9 of the purine base. All hydroxyl groups of **39** are shown to interact with residues in the active site, and interactions with 5′-OH are shown to be involved in substrate binding whereas interactions with 2′OH and 3′OH are required for stabilising the intermediate, oxocarbenium ion. The purine leaving group is shown to be involved in extensive aromatic stacking interactions with Trp-83 and Trp-260.[55]

5′-FDA **23** became a potential substrate of interest for *Tv*NH as it contains a ribose ring with the 5′-OH group substituted to fluorine, which has not been studied with this enzyme before. Previous experiments performed by Onega [34] had established that purified *Tv*NH can effectively catalyse the conversion of **23** to **37** when incubated with synthetically prepared **23** in buffer supplemented with $CaCl_2$ at 37 °C for 2-4 h.[34] Efficient synthesis of [^{18}F]-**37** with considerable RCY was necessary in order to explore [^{18}F]-**37** as a PET radiotracer and thus optimisation of the coupled enzyme reaction was attempted.

2.2.3 Previous studies conducted on production of [^{18}F]5-FDR [^{18}F]-37 using the fluorinase

The *Tv*NH enzyme and fluorinase were over-expressed in *E.coli* and purified. The purified enzymes were then used in their aqueous form in a one-pot enzyme reaction to generate [^{18}F]5-FDR [^{18}F]-**37** which ultimately proved to be inefficient and slow. Through optimisation, a sequential one-pot reaction proved to be considerably more efficient. Fluorinase and L-AAO were incubated with SAM-Cl **22** in the presence of [^{18}F]F^- at 37 °C, pH 7.0. The L-methionine **27** produced during the reaction is removed by L-AAO promoting the forward reaction. The reaction was analysed by radio-HPLC and a RCY of 97% to [^{18}F]-FDA [^{18}F]-**23** was observed after 1 h. Subsequently the reaction was heated to 95 °C to denature the fluorinase and L-AAO and the precipitated protein was removed by centrifugation. To the supernatant, *Tv*NH was added at a final concentration of 20 mg/ml in HEPES buffer (pH 8.0) and incubated for further 2 h at 37 °C. Post incubation, the [^{18}F]-**37** production was analysed by radio-HPLC to be ~ 80% RCY. [31] The three-enzyme route produced 45% RCY in 4 h [30]compared to which this two enzyme approach had increased the RCY considerably, and somewhat managed to reduce the reaction time.[31]

Despite reducing the incubation time of the two-enzyme route to 3 h, the actual processing time including: heat denaturing of the enzymes, centrifugation, removing any unreacted [^{18}F]F^{-}and [^{18}F]-**23** by HPLC, increased the total preparation time to 4 h.[31] Therefore further optimisation of this reaction was encouraged. Preliminary experiments revealed that increasing both fluorinase and *Tv*NH enzyme concentrations increased the rate of the reaction, however due to the reduced reaction volumes (600-700 μl), it was necessary to implement a method by which the two enzyme concentrations could be increased evading the dilution of the reaction. In addition previous studies used HPLC to separate and remove any unreacted [^{18}F]F^{-} and [^{18}F]-**23** which was found to be time consuming and added to the total preparation time [^{18}F]-**37** with an increased risk of radioactive exposure. Therefore, an alternative approach to removing any unreacted [^{18}F]F^{-} and [^{18}F]-**23** was explored. The focus of this chapter was to develop a safe, practical and user-friendly method of producing [^{18}F]-**37** in a clinical environment.[56]

2.3 Project aims

The current trend is to develop enzymatic processes of organic synthesis, as the enzymes are attractive catalysts that match the fundamental principles of sustainable development and green chemistry.[57,58] Enzymes are considered costly with limited availability and may be unstable with a narrow substrate specificity. These issues are being addressed by ongoing advances in genetics and protein engineering.[59] Gene mining is increasing the number of genes/enzymes available and developments in enzyme immobilisation techniques have rendered enzymes much more stable for process development.[58,60] Natural enzymes or genetically modified enzymes are available which can be screened for broader specificity acting on non-natural substrates. Despite their high potential, enzymes are not immediately attractive as catalysts in synthesis of PET radiotracers as the industry is not sufficiently acquainted to deal with biological material.

Studies conducted by Onega *et.al* demonstrated the use of fluorinase and *Tv*NH in a sequential one pot reaction to form [^{18}F]-**37** as a prospective PET radiotracer.[31] A practical method for producing [^{18}F]-**37** in the clinical environment in a ready to use, 'off the shelf' formulation using two enzymes was designed. And thus, preparation of a two enzyme kit for use in radiochemical laboratories, would offer a user-friendly technology for [^{18}F]-**37** production. It was an objective to increase the RCY for [^{18}F]-**37** production and further reduce the reaction time. The enzymes (fluorinase and *Tv*NH) were freeze-dried in their

buffers to produce a potentially portable kit, where rehydration on site would then provide ready to use enzymes for the radiochemical synthesis of [^{18}F]-**37** (Figure 4).

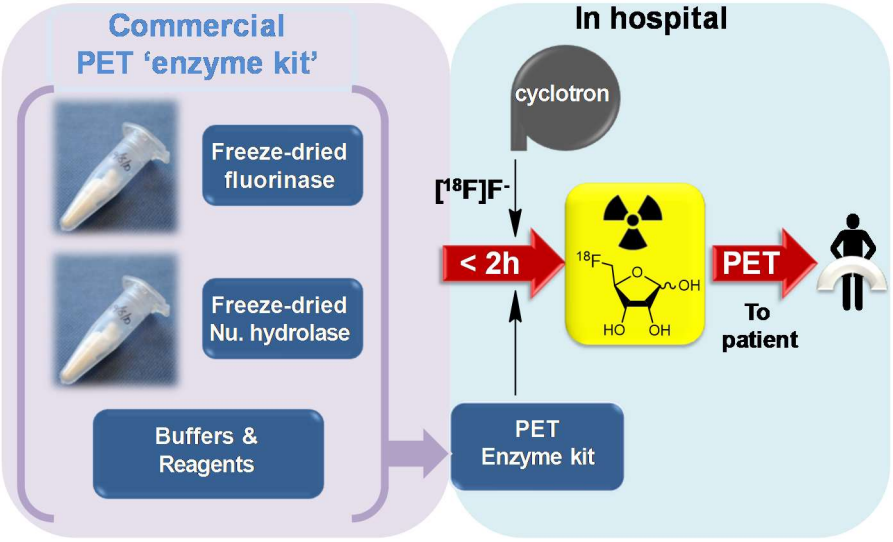


Figure 4: The project aim was to produce a PET 'enzyme kit' using freeze-dried fluorinase and *Tv*NH which could be used by radiochemical laboratories to produce [^{18}F]5-FDR [^{18}F]-**37** in good RCY.

2.4 Results and discussion

2.4.1 Over-expression of the fluorinase

The first attempt at purifying the fluorinase was reported in 2003,[61] where an ammonium sulphate precipitation was performed on a cell free extract (CFE) of *S.cattleya.* The proteins were separated according to their solubility in the presence of a high concentration of ammonium sulphate.[62] Subsequently the precipitate was applied to a gel filtration column and the fractions obtained were evaluated for fluorinase activity. Partial amino acid sequence of the fluorinase was obtained from these fractions and used to design degenerate DNA primers, which in turn were used to amplify the corresponding gene fragments by PCR (polymerase chain reaction). Amplified gene products were cloned into a suitable vector and sequenced. The sequencing data was used to design primers to perform gene walking to reveal the complete *flA* gene.[63] Subsequently the *flA* gene was cloned into a pET28(+) vector and transformed into BL21(DE3) *E.coli* cells for over-expression.[63]

The protocol for over-expression using *E.coli* containing pET28-flA in Luria Broth (LB) media has been reported previously and is described in detail (Experimental 2.6.1).[63] Cloning of the *flA* gene into the pET28 (+) vector offers the advantage of conferring kanamycin resistance to the *E.coli* carrying the pET28-flA recombinant plasmid. Furthermore an N-terminal poly-histidine tag (His_6-tag) is incorporated to the translated *flA* gene product which is useful during the purification of the final protein. Small scale cultures of *E.coli* carrying pET28-flA in LB were used to inoculate larger cultures supplemented with kanamycin (50 μg/ml). These cultures were shaken at 37 °C until the optical density (OD_{600}) reached 0.6, following which protein expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG). The cultures were then incubated at 25 °C for 16 h before harvesting the cells by centrifugation.

The harvested cell pellet was re-suspended in lysis buffer supplemented with Deoxyribonuclease I from bovine pancreas and protease inhibitor cocktail.[64] The cell suspension was lysed by sonication at 4 ºC and centrifuged to remove the cell debris. The lysate was purified by affinity chromatography using Ni-NTA (Ni-Nitrilotriacetic acid) agarose. The fluorinase was eluted from the Ni-NTA column by applying an elution buffer containing high concentration of imidazole which displaces the His_6-tagged fluorinase. The fractions obtained from the affinity purification were analysed by SDS-PAGE (Figure 5). Eluted fractions are shown to contain fluorinase monomer at 32 kDa and dimer at ~64 kDa (confirmed by matrix-assisted laser desorption/ionisation time of flight mass spectrophotometry (MALDI-TOF)). The presence of the dimer is possibly due to incomplete denaturing during sample preparation for SDS-PAGE analysis. Some fluorinase was also present in the insoluble fraction which accounts for misfolded proteins during over-expression. However sufficient fluorinase was obtained in the soluble fraction. The elution fractions containing fluorinase were pooled and dialysed in 20 mM Tris-HCl pH 7, 150 mM NaCl. Further purification of the fluorinase fraction was carried out by size exclusion chromatography (SEC) (Figure 6) which differentiates proteins according to their molecular weight. In order to obtain fluorinase free from impurities the cleanest fractions were pooled and run on SDS-PAGE (Figure 7). The identity of the fluorinase was confirmed by MALDI-TOF. The fractions obtained from SEC containing protein were pooled, concentrated and the protein concentration was measured by Nanodrop. This method produced high yield of fluorinase (~35 mg/L), which was used for freeze-drying trials.

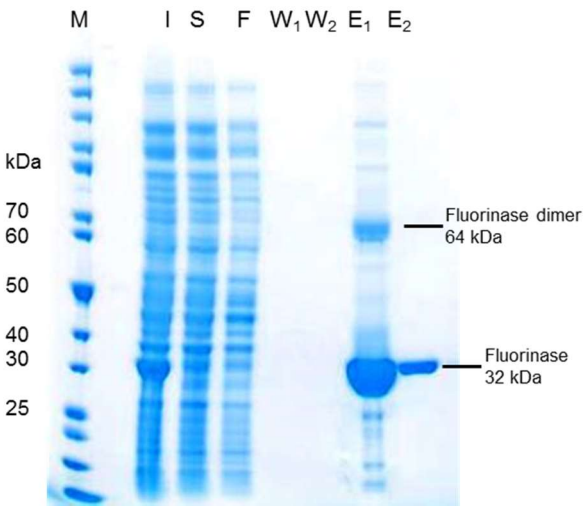


Figure 5: The SDS-PAGE gel electrophoresis of fractions obtained from Ni^{2+} affinity purification of fluorinase M: un-stained PageRuler protein ladder, I- insoluble fraction, S- soluble fraction, F- flow-through, W_1- wash 1, W_2- wash 2, E_1- elution 1, E_2-elution 2

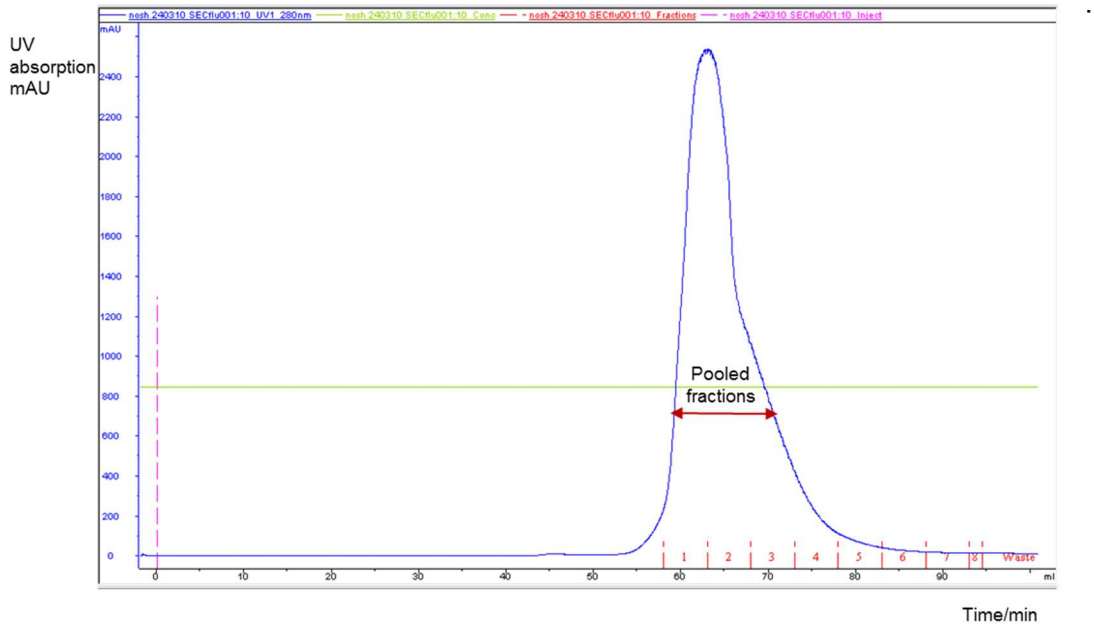


Figure 6: Size-exclusion column chromatography of *S. cattleya fluorinase* showing the UV absorption (280 nm) in blue, flow rate in green and fractions in red.

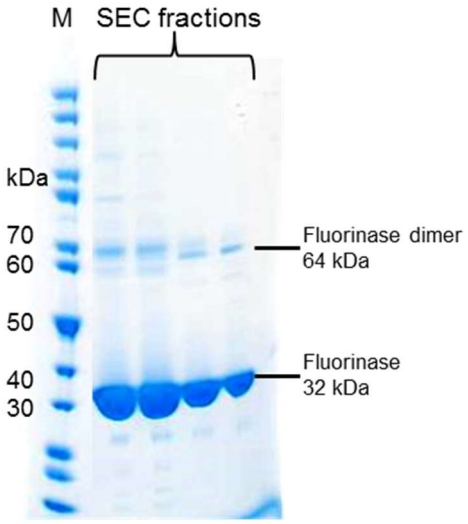


Figure 7: SDS-PAGE of the fractions obtained from size exclusion column chromatography: M- un-stained PageRuler protein ladder, SEC fractions- fractions containing fluorinase.

2.4.2 Over-expression of the *Tv*NH

The over-expression and purification of *Tv*NH was carried out by modification of previously established protocol [55,65] and is described in detail in the Experimental 2.6.2WK6 *E.coli* cells were transformed[66] with recombinant pQE-30-*Tv*NH. The pQE-30 plasmid carries an a*mpR* gene which confers ampicillin resistance. In addition the plasmid was designed to incorporate an N-terminal His_6-tag into *Tv*NH. The recombinant plasmid, pQE30-IAGNH was kindly donated by Prof J. Steyaert and Dr J. Barlow (Vrije Universtiteit Brussel, Belgium).

WK6 *E.coli* with pQE-30-*Tv*NH was used to inoculate large shake flasks containing LB (1 L) supplemented with ampicillin. The flasks were incubated at 37 °C shaking until the OD_{600} reached 0.6 and the over-expression was induced by adding IPTG (0.5 mM). The cells were grown for 16 h at 28 °C and harvested by centrifugation. The cell pellet was re-suspended in lysis buffer supplemented with Deoxyribonuclease I from bovine pancreas and protease inhibitor cocktail. The cells were lysed by sonication at 4 ºC and centrifuged to remove cell debris. The lysate was purified by affinity chromatography by applying the cell lysate on to Ni-NTA agarose. *Tv*NH was eluted with elution buffer containing high imidazole concentration and analysed by SDS-PAGE confirming the presence of *Tv*NH ~ 35 kDa in the elution fractions (Figure 8A). The SDS-PAGE also showed the presence of a ~70 kDa protein which was confirmed by MALDI-TOF to be the dimer of *Tv*NH due to partial denaturation prior to SDS-PAGE. The fractions containing *Tv*NH were pooled and dialysed against 20 mM HEPES pH 7, 150 mM NaCl supplemented with $CaCl_2$ (1 mM) to remove

imidazole **40** and permitting $CaCl_2$ to bind the active site and stabilise the *Tv*NH protein. Subsequently *Tv*NH was further purified by SEC (Figure 8B) and the fractions obtained were run on SDS-PAGE (Figure 8A). The fractions containing protein was pooled, concentrated and the protein concentration was measured by Nanodrop. This method produced high yield of *Tv*NH (~40 mg/L), which was used for freeze-drying trials.

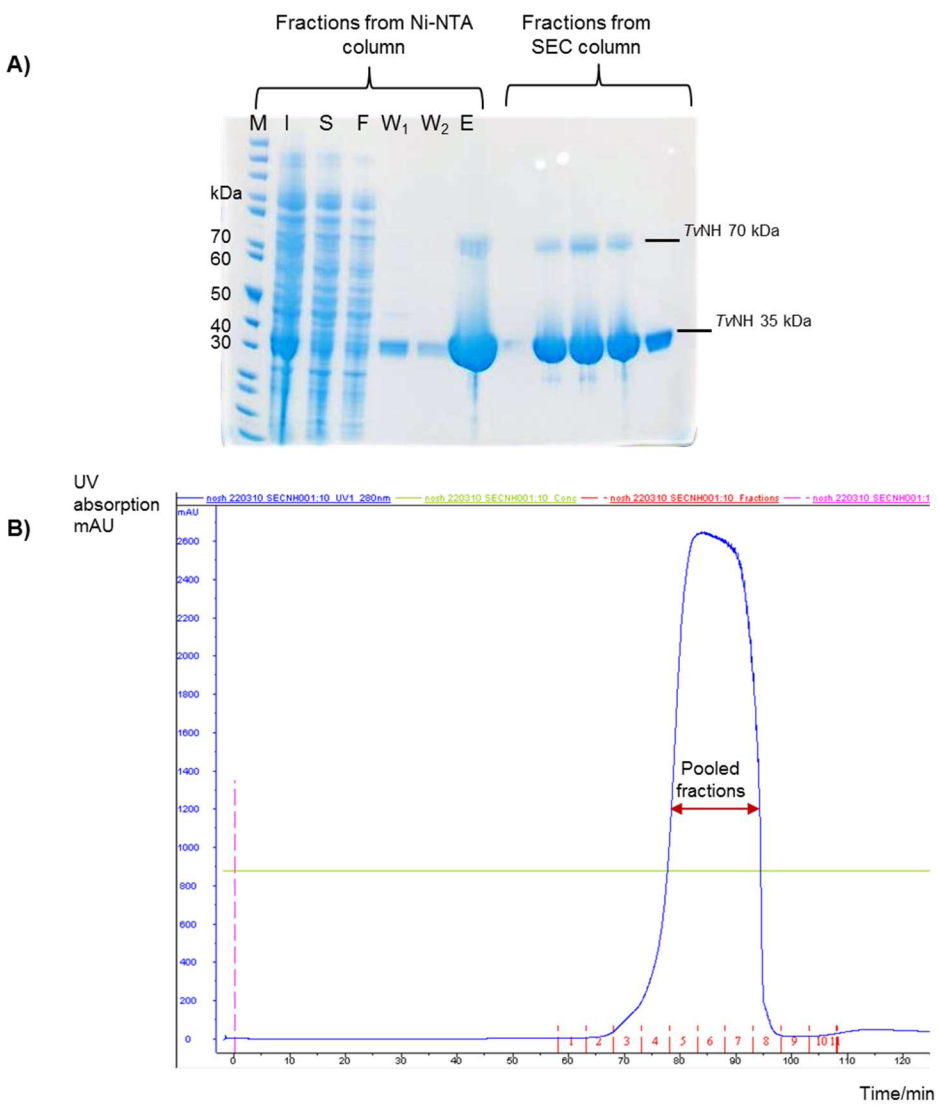


Figure 8: **A)** The SDS-PAGE gel electrophoresis of fractions obtained from Ni^{2+} affinity purification M- un-stained PageRuler protein ladder (Thermo scientific), I- insoluble fraction, S- soluble fraction F- flow-through, W_1- wash 1, W_2- Wash 2 E- elution and fractions obtained from SEC of *Tv*NH. **B)** Size-exclusion column chromatography (SEC) of *Tv*NH showing the UV absorption in blue, flow rate in green and fractions in red.

2.4.3 Enzyme immobilisation

In order to produce a user-friendly PET 'enzyme kit' it was a requirement to immobilise fluorinase and *Tv*NH on a solid support as the aqueous protein may decompose during transportation and storage. It was envisaged that a Ni-NTA resin would be suitable to trap the purified fluorinase and *Tv*NH through their His_6-tags which had been used to purify the proteins. These enzyme-bound resins could then be incubated with a reaction mixture containing the substrates, SAM-Cl **22** and KF and washed with buffer containing no imidazole **40**, which would elute the products/unreacted substrates but the enzymes will remain secured.

Purified fluorinase (2 mg) and *Tv*NH (6 mg) was applied to a cartridge containing 1 ml of Ni-NTA agarose™ (Qiagen). A cocktail containing SAM-Cl **22** (20 μM) and KF (100 mM) in 20 mM Tris-HCl pH7 buffer with 10 mM NaCl (Figure 9, reaction mixture) was applied to the column and the flow-through was collected. Prior to loading the column a sample of the reaction mixture was analysed by HPLC where as expected, only **22** was observed.

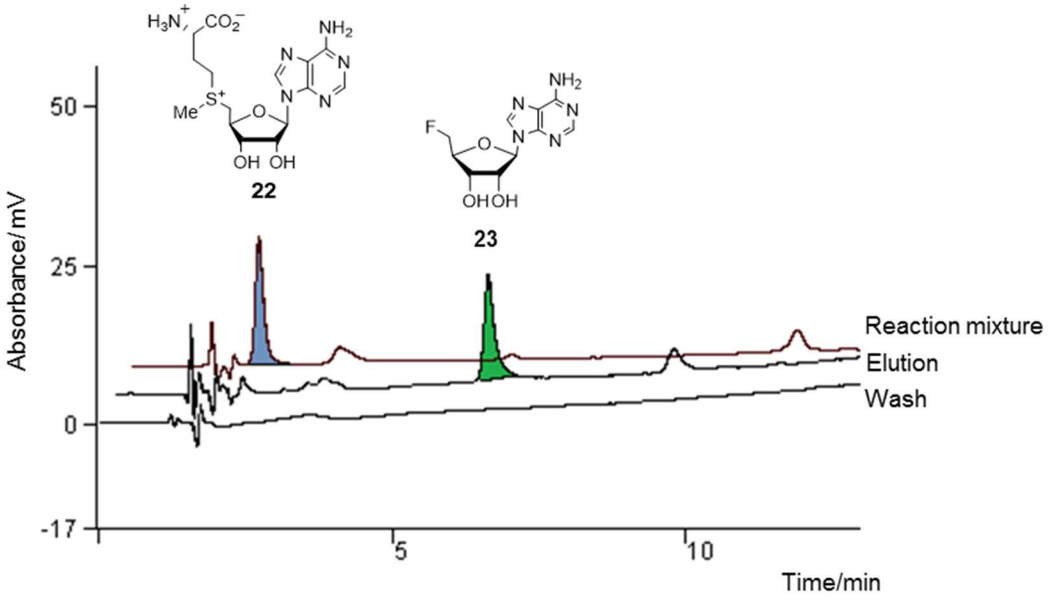
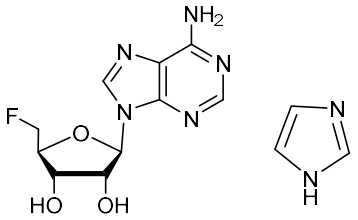


Figure 9: HPLC analysis showing the reaction mixture containing SAM **22** and KF prior to loading the cartridge containing immobilised fluorinase, first wash with buffer and elution with high imidazole containing buffer.

The cartridge was loaded with the reaction mixture and incubated for 1 h, after which was washed with 20 mM Tris-HCl (5 ml) (Figure 9, wash) following this the sample was freeze-dried and analysed by HPLC. The wash did not contain detectable quantities of 5′-

FDA **23** or adenine **29**, demonstrating that either the substrates, (**22** and **23**) were inaccessible to fluorinase/*Tv*NH or any products generated were trapped within the Ni-NTA agarose. The cartridge was washed with buffer containing high imidazole **40** concentration (400 mM) to elute both bound enzymes/substrates/products. The eluted fraction was heat treated (95 ºC) to precipitate the enzymes and the aqueous fraction analysed by HPLC (Figure 9, elution). The HPLC data showed only the presence of **23** in the eluted fraction suggesting **22** was converted to **23**. This suggests that **23** was possibly trapped in the immobilised fluorinase inaccessible to *Tv*NH or **23** released from the fluorinase was bound by the Ni^{2+} *via* it adenosine moiety. The adenosine moiety of **23** and imidazole **40** are structurally similar (Figure 10) and is possible that **23** released from fluorinase was trapped by Ni-NTA beads. It is possible that the nitrogen atoms in the adenosine ring of **23** could participate in Ni^{2+} chelation, resulting in **23** being bound the Ni-NTA beads. Maher *et al.* had observed this previously and exploited it in trapping DNA and RNA molecules containing an oligo(adenosine) tract using Ni^{2+}-NTA.[67] These data suggest that **23** is produced, but trapped and is not processed by immobilised *Tv*NH.

As such an alternative method of 'immobilising' the two enzymes was considered. A freeze-dried preparation of enzymes would potentially provide convenience in both transportation and may increase the shelf life.



24 imidazole **40**

Figure 10: Structural similarity between 5′-FDA **23** and imidazole **40**.

2.4.4 Lyophilisation of the enzymes

In protein pharmaceuticals it is important to design a formulation which is stable during transport and storage. During production, an aqueous solution of protein is easier to handle, economical and is the most convenient for the end user. However, protein solutions are inclined to chemical (eg: deamidation or oxidation) and physical (eg: aggregation and precipitation) degradation.[68] The degradation of protein solutions can be reduced by storing at low temperature (~ 4 ºC) with minimal agitation, however these conditions cannot be guaranteed during transport and handling.

Lyophilisation is a technique used widely in the pharmaceuticals and food industry in order to increase the shelf life, stability and transportability of products. It is a process where the water in the sample is frozen and then removed by sublimation in a three step process consisting of freezing, primary drying and secondary drying. The freezing stage takes several hours as the water is solidified and separates solutes from ice. The primary drying involves the sublimation of ice which commences as the pressure in the chamber is reduced to well below the vapour pressure (partial vacuum) of ice and sufficient heat is supplied to assist the sublimation process. Secondary drying then commences as most of the water had sublimed and at this stage any residual water bound to protein is removed by further increasing the temperature and further lowering the pressure. [69,70]

Most commercially available enzymes are available in their freeze-dried form for storage, handling and transport. Although lyophilisation has its advantages, the use of cold temperature in the freezing process can denature proteins due to cold shock,[69] along with pH changes and dehydration stress. The cold denaturation is dependent on the temperature, pH and the nature of additives. Some proteins are freeze-dried in the presence of an excipient system to stabilise the protein against the stress induced by freezing. Mostly proteins are susceptible to denaturing during the secondary drying stage due to the removal of the water layer hydrating to the protein, a process which de-stabilises the native conformation of the protein. Stabilising agents such as sucrose, mannitol and trehalose (lyoprotectants) are used to avoid protein denaturation during freeze-drying.[71] Carpenter and Crowe *et al.* [72,73] have shown by fourier-transform IR (FITR) that specific carbohydrates have the capability to protect proteins from denaturing by forming H-bonds substituting the role played by water. Some enzymes are shown to lose over 50% of their activity in a concentration dependent manner during the lyophilisation process in the absence of lyoprotectants.[74] This suggests that freeze-dried enzymes will benefit from added lyoprotectants.

2.4.5 Heat mediated SAM-Cl 22 decomposition

All of the fluorinase assays monitored by HPLC were observed to have a distinct peak with a retention time of around 11.4 min. Assays conducted with control samples of each compound added to the reaction showed it was originating from **22**. It was observed that this was a heat decomposition product of **22** (Figure 11) and was identified as 5′-deoxy-5′-(methylthio)adenosine (MTA) **41**.

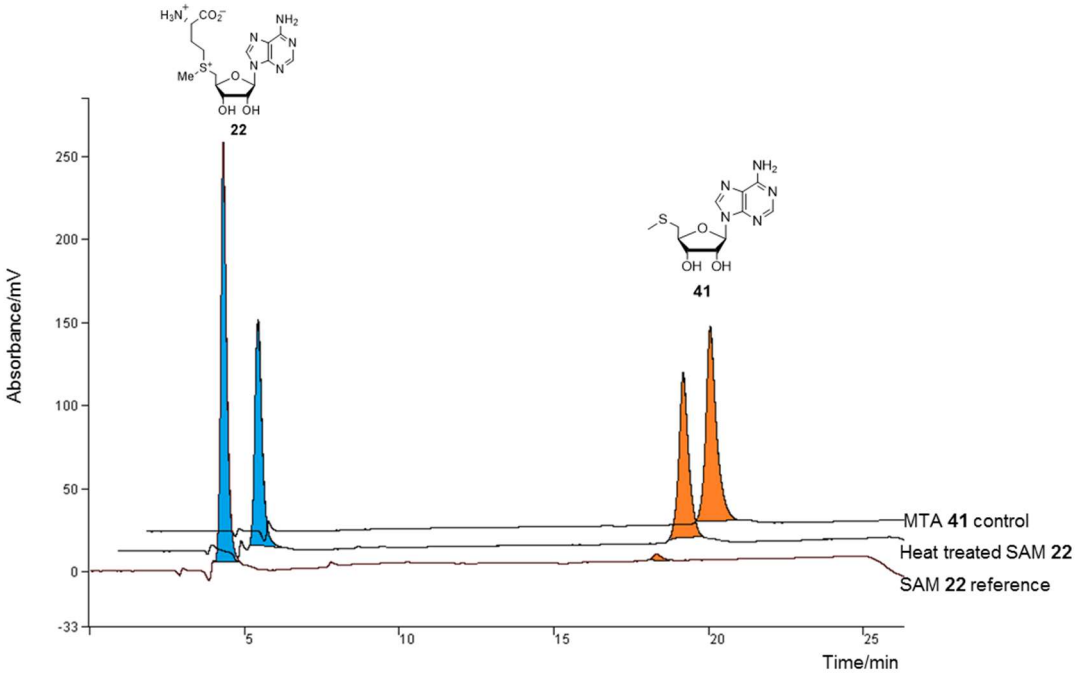
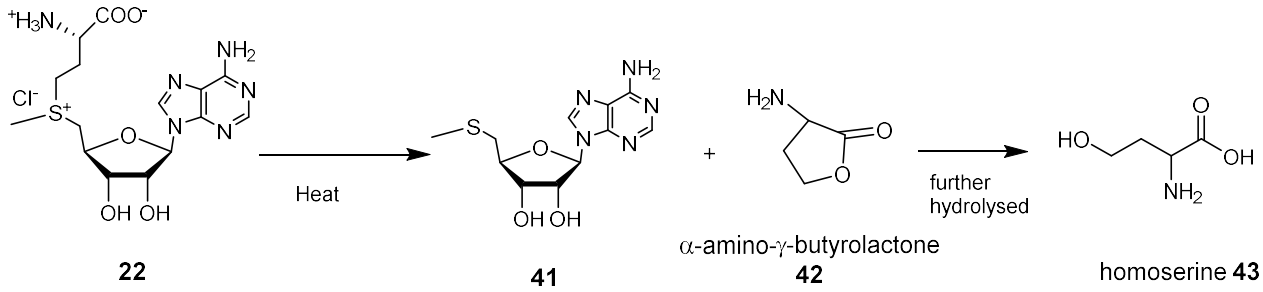


Figure 11: HPLC chromatogram demonstrating that SAM-Cl **22** decomposes to MTA **41** when heated.

The stability and hydrolysis products of **22** have been described in literature.[75,76] According to Stolowitz and Minch, heating **22** at 100 °C in the pH range of 4-7 produces **41** and α-amino-γ-butyrolactone **42**, followed by further decomposition to homoserine **43** (Scheme 5). SAM-Cl **22** decomposition was explored by analysing samples of heat treated **22** (95 °C for 10 min) along with reference compounds of **22** and **41** by HPLC (Figure 11). The reference **22** sample which was not heat treated produced negligible amount of **41** relative to the heat treated **22** which produced considerable amount of **41**. During the fluorinase assays the reactions are stopped by denaturing the enzyme *via* heating, during which residual **22** decomposes to **41** and α-amino-γ-butyrolactone **42**. [76]



Scheme 5: Heat induced decomposition products of SAM-Cl **22**.[76]

2.4.6 Use of a sucrose-glycine excipient system for enzyme lyophilisation

A sucrose-glycine excipient with the addition of polysorbate 80, enabled complete recovery of model proteins as described in the literature.[71,77] Therefore, fluorinase and *Tv*NH were freeze-dried in their respective buffers at 0.2 mg/ml concentrations containing a lyoprotectants cocktail of sucrose (18 mg/ml) and glycine (2 mg/ml) and polysorbate 80 (0.1 mg/ml) as described in the Experimental 2.6.4. When freeze-dried with these lyoprotectants the enzymes were observed to exist as a highly hygroscopic white powder (Figure 12).

Figure 12: Freeze-dried fluorinase in an eppendorf tube.

The activity of the freeze-dried fluorinase and *Tv*NH was analysed before (in the aqueous form) and after freeze-drying. Fluorinase (0.2 mg/ml), either freeze-dried in a sucrose-excipient system or in its aqueous form was incubated with SAM-Cl **22** (20 μM) and KF (100 mM) for 30 min at 37 °C. Each reaction was heat treated (95 ºC) and centrifuged to precipitate the enzyme. Each reaction was monitored by HPLC which revealed the production of 5′-FDA **23** and unreacted SAM-Cl **22** with its heat-mediated, decomposition product MTA **41** (Figure 13A). Some adenosine **47** was also visible by HPLC which was residual from fluorinase precipitation. Similarly the activity of *Tv*NH was analysed by incubating with adenosine **47** (20μM), which is a natural substrate[55] for *Tv*NH and monitoring its conversion to adenine **29** by HPLC (Figure 13B). The activity of the freeze-dried enzymes relative to aqueous enzymes was analysed by HPLC. The production of **23** and adenine **29** was observed by HPLC in multiple runs and the activity of the enzymes was calculated as a percentage of the activity of the aqueous enzymes. The data indicates a clear reduction in enzyme activity for both freeze-dried enzymes relative to aqueous enzymes. A 19% reduction of activity for freeze-dried fluorinase and 11% reduction of activity for freeze-dried *Tv*NH relative to aqueous enzymes were observed. Despite the loss of activity observed when both enzymes were freeze-dried, freeze drying was found to be a better method for enzyme preparation for the enzyme kit considering the advantages of

freeze-drying which may prolong the stability during storage [78,79] of the enzyme and may also aid in transportation.

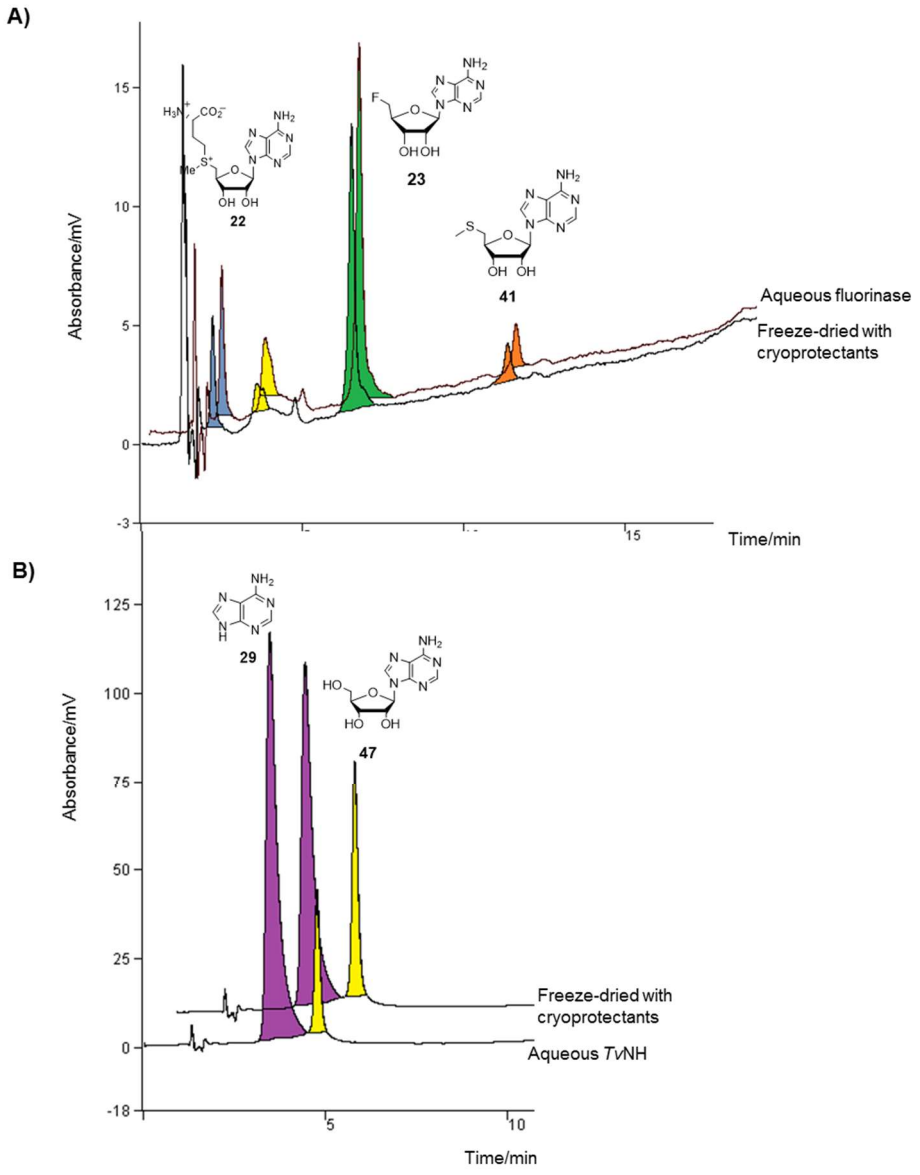


Figure 13: A) HPLC chromatogram of fluorinase enzyme assays with either freeze-dried or aqueous enzyme. (SAM-Cl **22**; t_R = 2.3 min (blue), adenosine; t_R= 3.7 min (yellow), 5′-FDA **23**; t_R= 13.7 min (green), MTA **41**; t_R = 11.4 min (orange) **B)** HPLC chromatogram of *Tv*NH enzyme assay with either freeze-dried or aqueous enzyme (adenine **29** t_R = 3.5min (purple) adenosine **47** t_R= 4.7 min (yellow)).

Similar activity studies were carried out to assess if freeze-drying with or without sucrose-glycine excipient system would have a significant effect on the activity of either enzymes. Multiple HPLC assays were carried out on both fluorinase and *Tv*NH that were either freeze-dried with or without a sucrose-glycine excipient system. The production of 5′-FDA **23** (Figure 14A) or adenine **29** (Figure 14B) was monitored as before and the activity of the enzymes were calculated as a percentage of the freeze-dried enzymes with lyoprotectants. A 5% loss of activity was observed in fluorinase freeze-dried without lyoprotectants and ~2% loss of activity was observed for *Tv*NH when freeze-dried without any lyoprotectants.

However, when using the freeze-dried enzyme preparations with lyoprotectants for generating 5-FDR **37** the formulation was found to be very viscous. This was not an issue when performing cold reactions with non-labelled F^- however when using labelled $[^{18}F]F^-$ where the reaction volumes were reduced to 1/2 that of the volumes used in cold reactions this stickiness posed a problem in the handling and transferring of samples. Therefore it was decided not to use any lyoprotectants for the freeze-drying preparation for future radiolabelled studies.

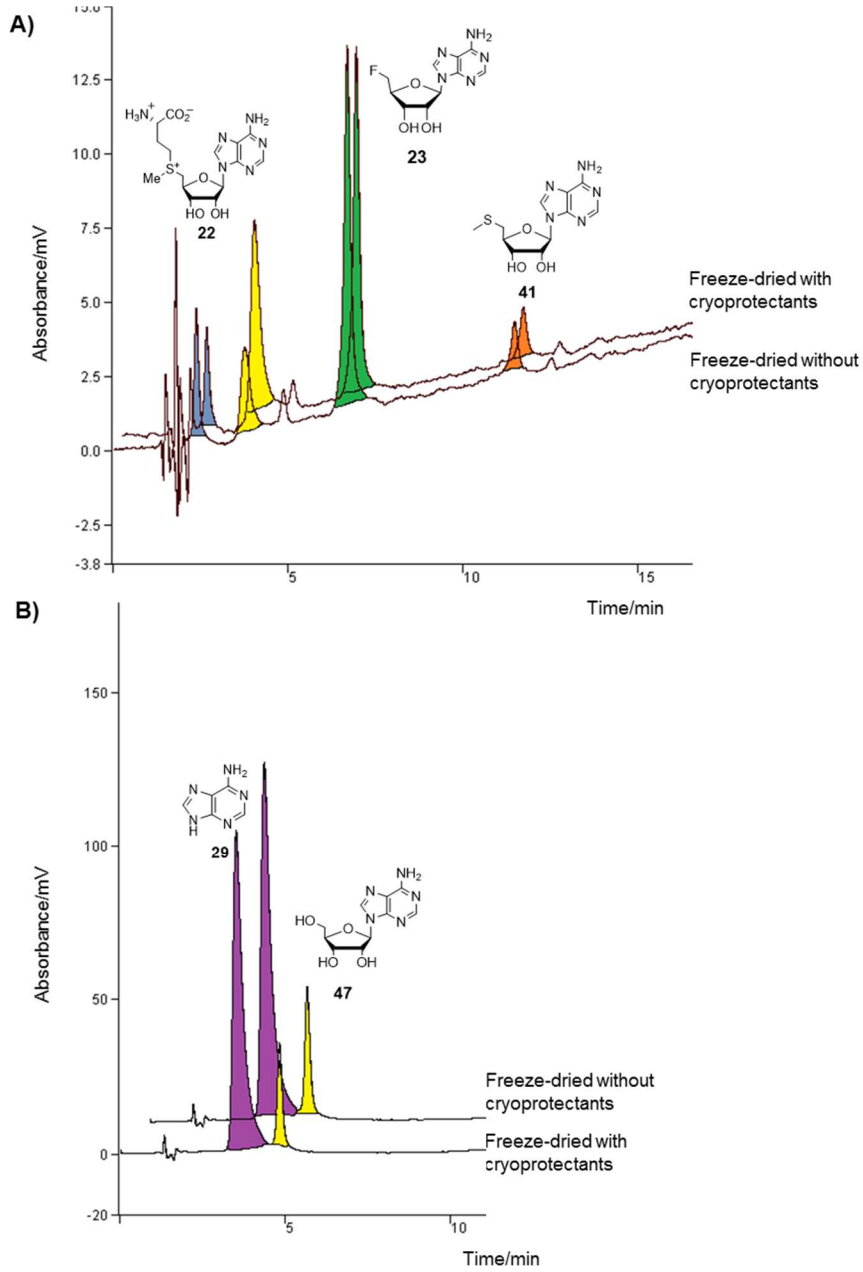


Figure 14: A) HPLC chromatogram of fluorinase assay using either enzyme freeze-dried with lyoprotectants or freeze-dried without lyoprotectants. SAM-Cl **22**; t_R = 2.3 min (blue), adenosine **47**; t_R= 3.7 min (yellow), 5′-FDA **23**; t_R= 13.7 min (green), MTA **41**; t_R = 11.4 min (orange) **B)** HPLC chromatogram of *Tv*NH enzyme assay either freeze-dried with lyoprotectants or freeze-dried without lyoprotectants adenine **29** t_R= 3.5min (purple) adenosine **47** t_R= 4.7 min (yellow).

2.4.7 Evaluating long term enzyme stability during storage

The major requirement in the designing of an enzyme kit for the efficient production of [^{18}F]5-FDR [^{18}F]-**37**, is that the enzymes should be stable over a few weeks during storage. Thus the activities of the freeze-dried enzymes were assayed over a 4 week period as described in the Experimental 2.6.6.

The activity of both enzymes was monitored either by HPLC and ^{19}F NMR for the fluorinase and *Tv*NH, respectively. Fluorinase was purified and freeze-dried with buffer (1 mg) and stored at -20 °C over eight weeks. Every week an aliquot of the fluorinase was removed and incubated with SAM **22** (0.2 mM), KF (100 mM) and sodium benzoate **44** (2 mM) (as an internal standard). This reaction was incubated at 37 °C and every 5 min an aliquot was removed, heat-treated, centrifuged and the supernatant was analysed by HPLC. In order to assess the activity of the enzyme, it was necessary to quantify the concentration of **23**. For this an internal standard of **44** was used in each reaction. Sodium benzoate **44** was chosen as the internal standard as its retention time (t_R= 9.8 min) is close to the retention time of **23** (t_R= 13.2 min), and is a non-participating reference. The concentration of **23** produced at each time point was calculated relative to the known concentration (2 mM) of **44**. The level of **23** was measured at each time point as a ratio of the peak height of **44**. Figure 15A depicts the increase in **23** production over time. Some un-reacted **22** along with the decomposed **22** derivative, MTA **41** could also be seen in the HPLC traces.

*Tv*NH was purified and freeze-dried in aliquots (20 mg) and stored at -20 °C over eight weeks. Every week an aliquot of *Tv*NH was removed and assayed. The progress of the assays were monitored by ^{19}F NMR over 3 h (Figure 15B). *Tv*NH was incubated with 5′-FDA **23** (0.5 mg) at 37 °C and samples were removed every hour for 3 h. The samples were heat treated to remove the enzyme, supplemented with D_2O and analysed by ^{19}F NMR. In order to quantify the 5-FDR **37** production a known concentration of an internal standard, 2-fluoroethanol **45** (1 mM) was used. Integration of the **37** peaks relative to the standard, provided the concentration of the **37** α/β anomers depicted in the ^{19}F NMR spectra as the reaction progressed with time, the production of both anomers of **37** can be observed as the concentration of **23** decreases (Figure 15B).

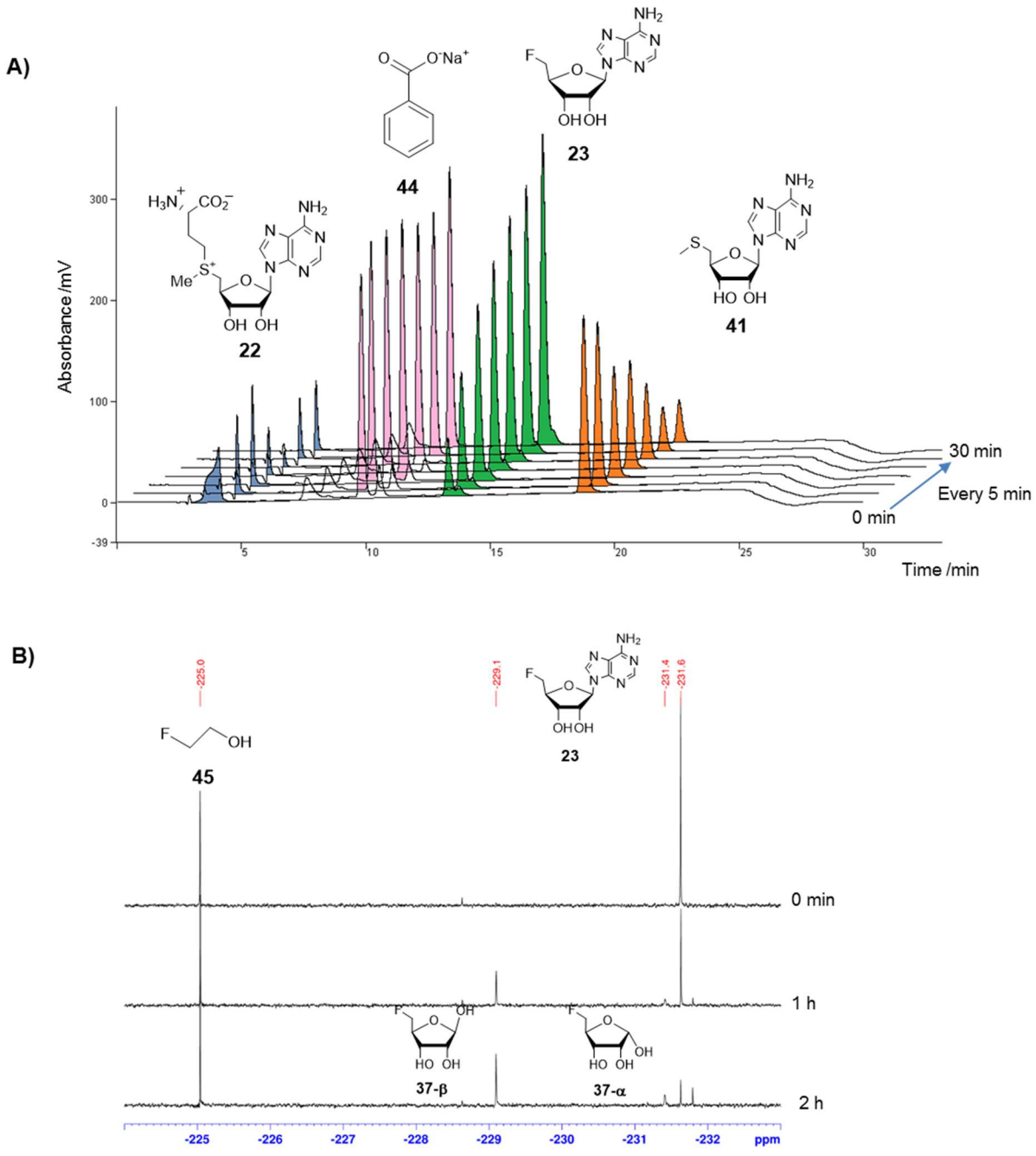


Figure 15 : A) A time course of HPLC assays over 30 min showing the production of 5′-FDA **23** by freeze-dried fluorinase which was stored at -20 °C for a week.(SAM-Cl **22**; t_R = 4.14 min (blue), $NaC_7H_5O_2$; t_R= 9.8 min (pink); 5′-FDA **23**; t_R= 13.21 min (green); MTA **41**; t_R = 18.68 min (orange)). **B)** A time course of ^{19}F {^1H} NMR assays over 2 h showing the production of two 5-FDR **37** anomers from freeze-dried *Tv*NH which was stored at -20 °C for a week. (2-fluorethanol = -225.0 ppm, 5′-FDA **23** = -231.6 ppm, 5-FDR **37** α= -231.4 ppm, 5-FDR **37** β= -229.1 ppm)

The concentration of 5′-FDA **23** and the two 5-FDR **37** anomers α/β were calculated as a ratio with respect to internal standards for each enzyme reaction. This data was used to produce a graph showing the rate of each enzyme reaction assayed on successive weeks. Absorbance of 5′-FDA **23** relative to the absorbance of the known concentration of sodium benzoate **44** was recorded against time (Figure 16). And for *Tv*NH the integration of 2-fluoroethanol **45** was used as a reference and the NMR signal for **37** anomers were integrated relative to the internal standard to obtain the concentrations of both anomers (Figure 17 A and B).

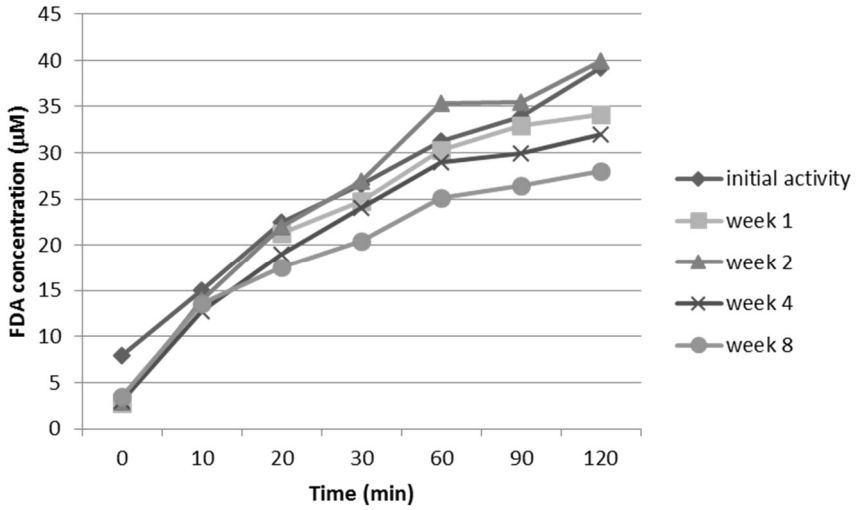


Figure 16: The initial rate of reaction of freeze-dried fluorinase determined by measuring the concentration of 5′-FDA **23** over 8 weeks.

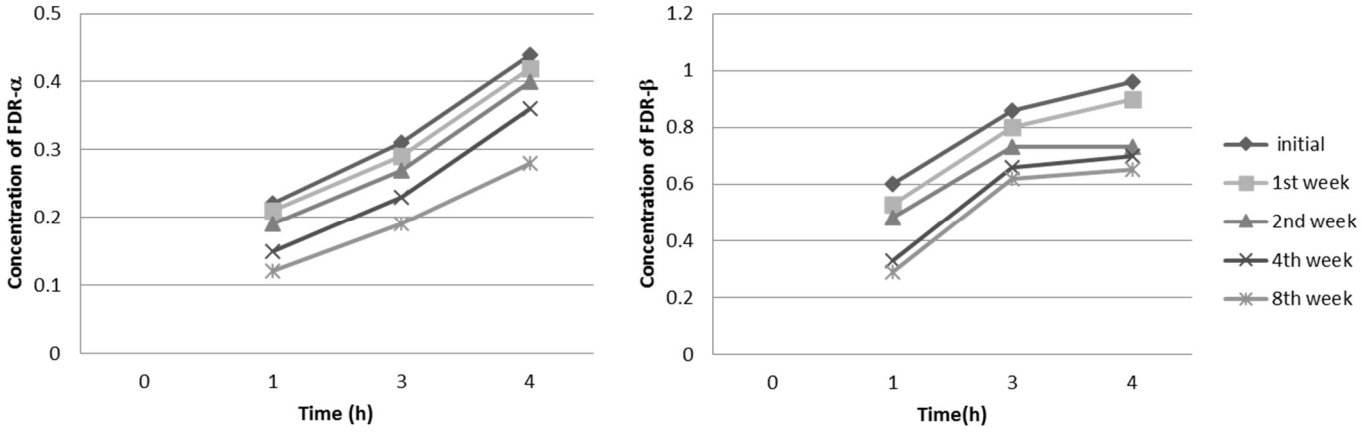


Figure 17: The initial rate of reaction of *Tv*NH over 4 h from assay samples analysed over 8 weeks. **A)** Rate of the reaction determined by measuring the concentration of 5-FDR-α **37**-α **B)** Rate of the reaction determined by measuring the concentration of 5-FDR-β **37**-β.

The activity for fluorinase and *Tv*NH is depicted as a percentage of the original activity over time in Figure 18. The stability of the freeze-dried fluorinase and *Tv*NH when stored at -20 °C over 8 weeks was mostly retained, albeit with a modest loss of ~20% reported for both the enzymes (Figure 18).

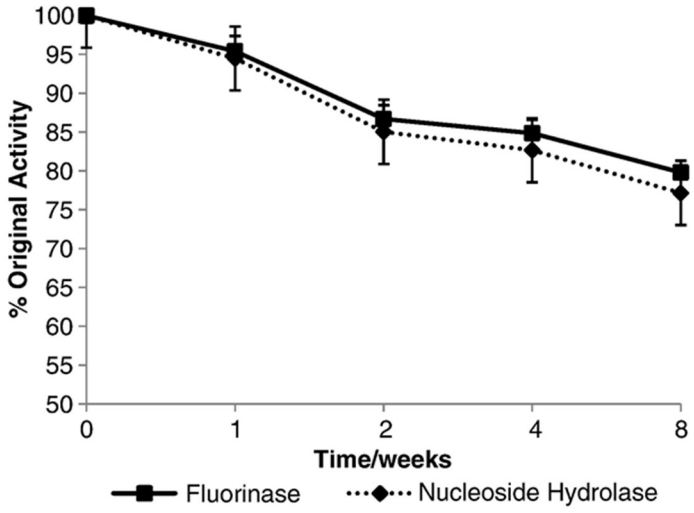


Figure 18: Activity profiles of freeze-dried fluorinase and *Tv*NH when stored at -20 °C over a 8 week period.

The freeze-dried enzymes were used in producing [^{18}F]5-FDR [^{18}F]-**37** for PET imaging.

2.4.8 Non-labelled enzyme assays 'cold reactions'

Prior to performing hot labelling assays, cold assays were carried out with freeze-dried fluorinase (of varying concentrations) incubated with SAM-Cl **22** (0.1 mM) and KF (100 mM). The reaction was incubated for 20 min at 37 ºC, heat denatured (95 ºC), centrifuged and the supernatant was added to freeze-dried *Tv*NH (of varying concentrations). The reaction was incubated for 40 min, heat denatured and centrifuged as before, prior to analysing on HPLC for the production of FDA **23** and adenine **29**. Various concentration combinations of fluorinase and *Tv*NH were tried in order to obtain the optimal ratio of fluorinase to *Tv*NH to provide the best conversions in a total reaction time of 1 h. These cold assays demonstrated that 3 times more *Tv*NH, relative to fluorinase was necessary for best conversions in the absence of L-AAO. [80]

2.4.9 [^{18}F]-5-FDR [^{18}F]-37 production using freeze-dried fluorinase and *Tv*NH

Cold assays demonstrated that 3 times more *Tv*NH was required than fluorinase for optimal conversions in the absence of L-AAO. However addition of L-AAO is required for hot labelling experiments in order to drive the reaction forward. Initially it was necessary to evaluate the minimum concentration of freeze-dried enzymes necessary for [^{18}F]-**37** production and develop a method by which unreacted [^{18}F]F-**23** and [^{18}F]F$^-$ could be removed to obtain pure [^{18}F]-**37**. Performing numerous trial hot experiments was not possible due to limited access to the radioactivity lab. Hot labelling experiments are performed under significantly different reaction conditions to the cold reactions and require very high concentrations of *Tv*NH relative to fluorinase. [80]

All experiments with [^{18}F]F$^-$ were carried out by Dr Sergio Dall'Angelo assisted by Dr Lutz Schweiger at the John Mallard PET Center (Aberdeen University) and described in detail in Experimental 2.6.7. Preliminary hot experiments with freeze-dried fluorinase (3 mg) and *Tv*NH (10 mg) demonstrated an RCY of ~66% of [^{18}F]-**37** within 2 h (Figure 19). It was observed that increasing the enzyme concentrations would shorten the reaction time and increase the RCY therefore a range of freeze-dried enzyme concentrations were explored (varying from 5 mg to 60 mg).

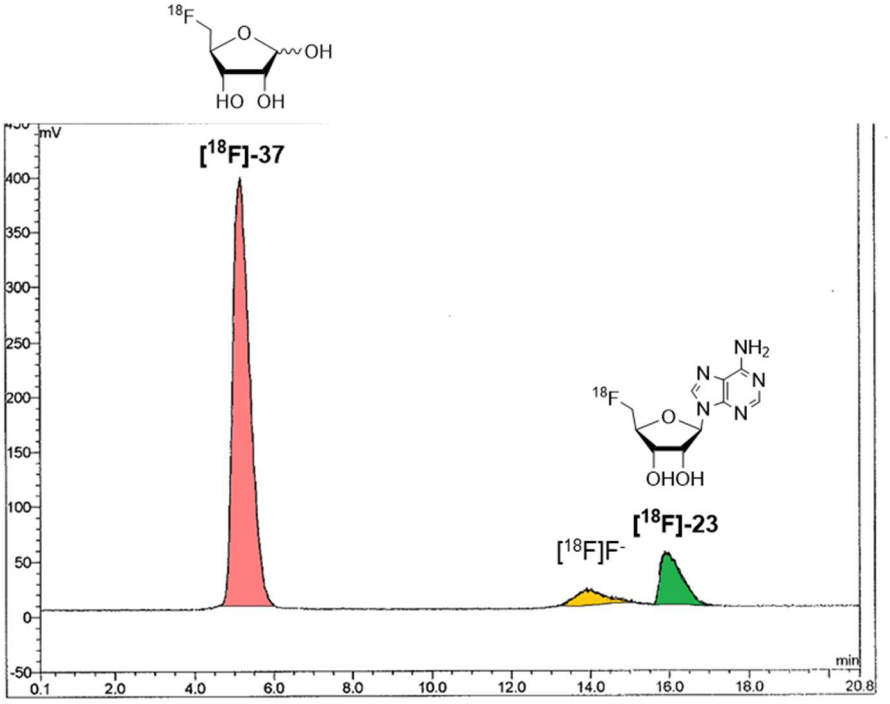


Figure 19: Preliminary radiolabelling studies using freeze-dried fluorinase (3 mg) and freeze-dried *Tv*NH (10 mg) to generate ~66% conversion to [^{18}F]5-FDR [^{18}F]-**37** within 2 h. Experiments were conducted by Dr Sergio Dall'Angelo at University of Aberdeen.

The optimal RCY was obtained using freeze-dried fluorinase (6.5 mg) and *Tv*NH (39 mg) (the reaction was set up following Experimental 2.6.8). Freeze-dried fluorinase (6.5 mg) was re-hydrated and incubated with SAM-Cl **22** (0.16 mM), L-AAO (1 mg) and [^{18}F]F^{-}at 37 °C (pH 7.0) for 20-30 min. This generated a ~98% RCY of [^{18}F]F-**23** (Figure 20A). Subsequently the enzyme mixture was heated and centrifuged to remove the precipitated fluorinase and the supernatant was added to the freeze-dried *Tv*NH (39 mg). The reaction was then incubated for 1 h at 37 ºC to produce [^{18}F]-**37** at a RCY of 98% (Figure 20B).[56]

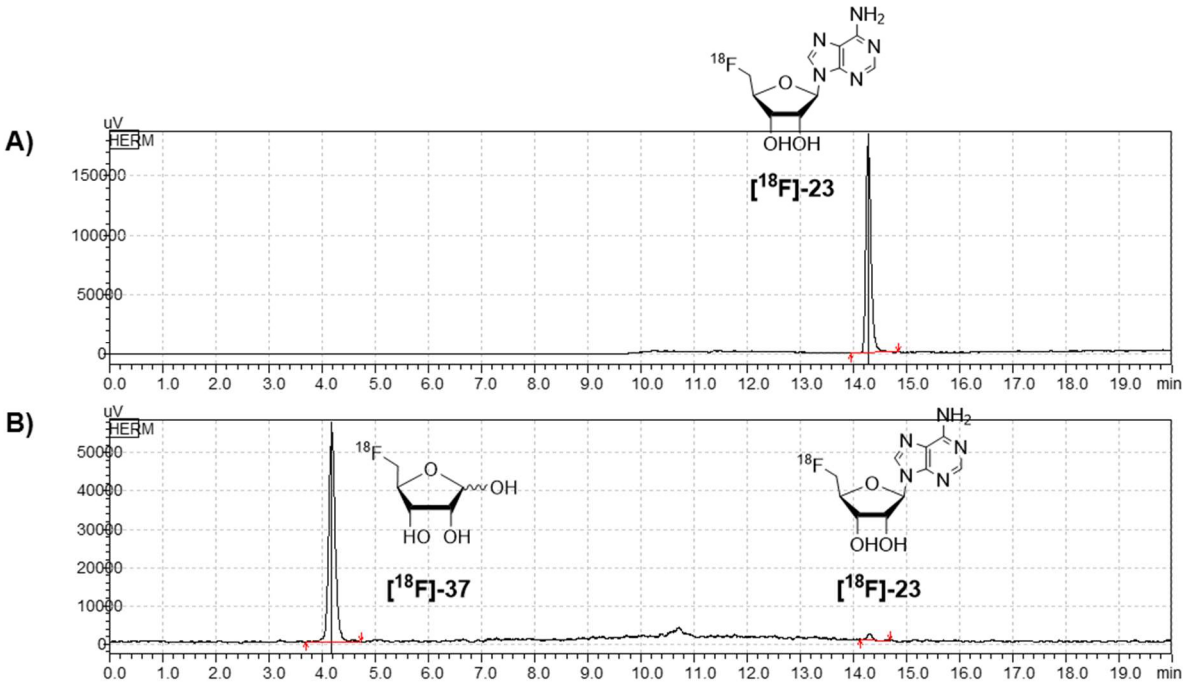


Figure 20: A) Radio-HPLC showing the production of [^{18}F]-FDA [^{18}F]F-**23** after 30 min incubation with [^{18}F]F^{-}, SAM-Cl **22** and freeze-dried fluorinase **B)** Radio-HPLC showing the production of [^{18}F]-**37** followed by 1 h incubation of the supernatant obtained from fluorinase reaction containing [^{18}F]F-**23** with freeze-dried *Tv*NH. Experiments were conducted by Dr Sergio Dall'Angelo at VU University Medical Centre in Amsterdam.[56]

2.4.10 PET studies on tumour mice using enzymatically produced [^{18}F]5-FDR [^{18}F]-37 as a radiotracer.

An *in vitro* imaging experiment was conducted to explore the enzymatically produced [^{18}F]5-FDR [^{18}F]-**37**, as a radiotracer for tumours. This was conducted at the VU University of Medical Center in Amsterdam by Dr Sergio Dall'Angelo, with the assistance of Dr Danielle Vugts and Prof A. Windhorst. The freeze-dried enzyme preparations were prepared in St Andrews (described in detail in Experimental 2.6.3) and the [^{18}F]-**37** production using the freeze-dried enzymes was slightly modified as described in Experimental 2.6.8.

For the imaging study four mice were injected with the A431 human epithelial carcinoma cell line[81] in their right and left thighs. All mice developed tumours and the freeze-dried enzyme kit was used to generate [^{18}F]5-FDR [^{18}F]-**37** which was injected to mice with tumour and PET imaged. The following day the same mice were injected with [^{18}F]FDG **32** in order to compare the [^{18}F]-**37** uptake with this common radiotracer. Cross-sectional images were taken of the whole body of the mouse following the [^{18}F]-**37** injection at a set of time frames (600-1200 sec after injection) (Figure 21A) (described in detail in Experimental 2.6.9). [56]

As shown in the Figure 21A, the [^{18}F]-**37** uptake was clearly observed in the PET images. The relative uptake of [^{18}F]-**37** against [^{18}F]-**32** was analysed (Figure 21B). In the initial 400 sec (5 min) post injection, [^{18}F]-**37** uptake was higher compared to that of [^{18}F]-**32** (Figure 21C). The [^{18}F]-**37** uptake remained constant until around 1200 sec (20 min) however from this point the efflux rate of [^{18}F]-**37** from the cell increases and eventually the efflux rate is greater than the influx rate, reducing the image intensity. Meanwhile the [^{18}F]-**32** uptake steadily increases as does the image intensity. This is consistent with [^{18}F]-**32** becoming metabolically trapped in the cells due to phosphorylation. In Figure 21C the comparative uptake of [^{18}F]-**37** against [^{18}F]-**32** is plotted.

There was no observable accumulation of [^{18}F]F^- in the bone during these imaging studies suggesting that [^{18}F]-**37** is metabolically stable and resistant to fluoride elimination unlike [^{18}F]F-**23**.[34] In addition [^{18}F]-**37** was not observed in the brain, suggesting inefficient uptake of [^{18}F]-**37** by trans-membrane GLUT transporters. High [^{18}F]-**37** uptake was observed in the bladder of the mice and [^{18}F]-**37** was observed to be excreted faster than [^{18}F]-**32**. [56]

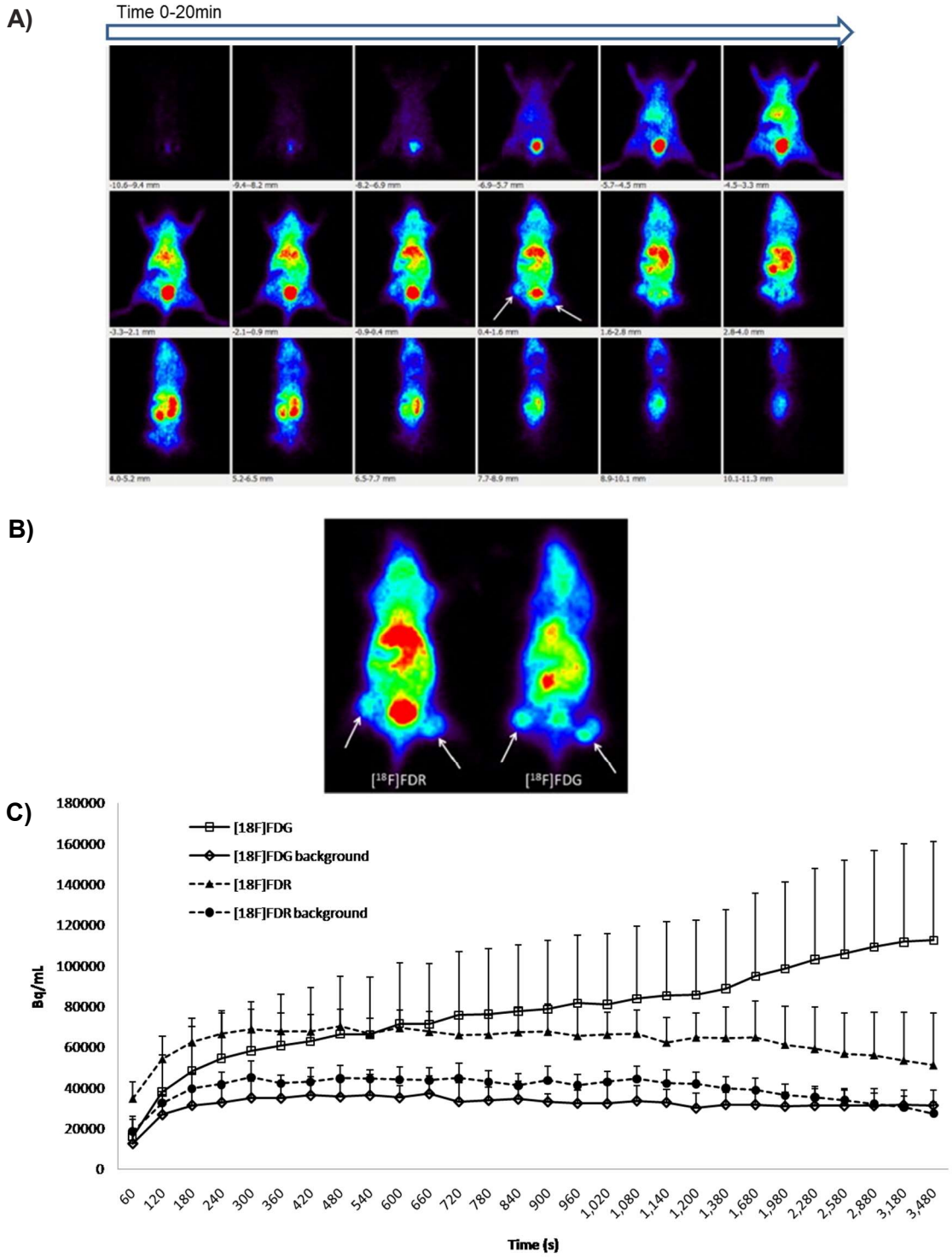


Figure 21: A) A tumour induced mouse was injected with [^{18}F] 5-FDR [^{18}F]-**37** prepared using freeze-dried enzymes and PET images were obtained in a time frame of 20 min. Two tumours on the legs are shown in arrows. **B)** The uptake of [^{18}F]-**37** into the tumours of the mouse at a given time frame was compared to the uptake of [^{18}F]-**32** in the same mouse. **C)** Uptake of [^{18}F]-**37** and [^{18}F]-**32** was plotted against their backgrounds to obtain the corrected reading for each uptake. Experiments were conducted by Dr Sergio Dall'Angelo at VU University Medical Centre in Amsterdam.[56]

2.5 Conclusions

A user-friendly, shelf-stable method using the fluorinase and *Tv*NH was demonstrated for the efficient production of [^{18}F]-**37**. [^{18}F]5-FDR [^{18}F]-**37** could be prepared within 1 h by this method, and when injected into tumour induced mice (A431 cells) PET imaging demonstrated the initial uptake of [^{18}F]-**37** was higher than that of [^{18}F]FDG [^{18}F]-**32**. The [^{18}F]-**37** uptake remained stable for up to 20 min before the signal diminished as the efflux rate of [^{18}F]-**37** was increased during which the [^{18}F]-**32** was constantly up taken. Therefore [^{18}F]-**32** was found to be a better PET radiotracer overall when compared to [^{18}F]-**37**, as the rate of [^{18}F]-**37** efflux from the cells competed with its influx, reducing image resolution. Nonetheless [^{18}F]-**37** provided reasonable imaging resolution as a PET tracer for the initial 10-20 min post injection and during this period [^{18}F]-**37** was as effective as [^{18}F]-**32**. However [^{18}F]-**32** proved to be the more effective radiotracer in imaging tumours over a longer period of time. The stability analysis of the freeze-dried enzymes showed that they were stable over 8 week period at -20 °C with a 20% loss in activity.

[^{18}F]5-FDR [^{18}F]-**37** may have some prospects as a PET radiotracer for imaging tumours which are not responsive to [^{18}F]-**32** such as prostate and neuroendocrine tumours.[3] This study has shown that a user friendly preparation of the fluorinase can be utilised in remote radiochemical labs to generate a radiotracer, which then can be used to conduct clinical imaging studies.

2.6 Experimental

2.6.1 Purification of Fluorinase enzyme in *E.coli*

The fluorinase was purified using a modified method from a previously established protocol.[63] *E. coli* BL21 (DE3) Gold cells were transformed with pET-28a(+)-flA recombinant vector inoculated into a small culture of LB (Luria Broth) media (10 ml) and incubated overnight at 37 °C. This culture was used to inoculate 2 L baffled flasks containing 1 L of LB media supplemented with kanamycin (100 μg/ml) and incubated at 37 °C until the OD_{600} had reached 0.6. These cultures were then induced by addition of isopropylthiogalactoside (IPTG; 0.5 mM) and incubated for 16 h at 25 °C. Cells were harvested by centrifugation (3200 rpm; 20 min) and the pellet was re-suspended in lysis buffer (20 mM Tris-HCl pH 8, 150 mM NaCl; 50 ml) supplemented with protease inhibitor cocktail (1 x Complete Mini-EDTA free, Roche) and Deoxyribonuclease I from bovine pancreas (Sigma; 0.1 mg/ml). The cells were lysed by sonication using SoniPrep 150 Plus (MSE™, UK) pulsing for 15-20

sec with 2 min pauses at 4 ºC. The lysate was centrifuged (15000 rpm; 20 min) and the supernatant was filtered through a syringe filter (0.45μm) before loading onto a column packed with Ni-NTA resin™ (Qiagen; 5ml). The column was washed twice with wash buffer (20 mM Tris-HCl pH 8, 400 mM NaCl, 20 mM imidazole **40** pH 8) and then eluted with elution buffer (20 mM Tris-HCl pH 8, 400 mM NaCl, 400 mM imidazole **40** pH 8). The protein containing fractions were analysed by SDS-PAGE then dialysed into buffer 20 mM Tris-HCl pH 7, 150 mM NaCl at 4 ºC for 12 h. The protein was concentrated using a centrifugal filter unit (10000 Da, GE Healthcare). The concentrated protein was loaded on to a gel filtration column Superdex S-200(HR 16/60) column, Pharmacia Biotech) connected to an FPLC (ÄKTA purifier GE healthcare) pre-equilibrated with buffer (10 mM Tris-HCl pH 7, 30 mM NaCl) and the protein was eluted in the same buffer. Elution fractions were analysed by SDS-PAGE and the purified protein was concentrated using a centrifugal filter unit (10000 Da, GE Healthcare) and the protein concentration was measured by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.) [56].

2.6.2 Purification of Nucleoside hydrolase in *E.coli*

*Tv*NH was purified according to a modified method from a previously established protocol[55,56]. *E. coli* WK6 cells were transformed with pQE-30-*Tv*NH and small LB cultures were inoculated and grown at 37 ºC overnight. These cultures were used to inoculate 2L baffled flasks containing LB (1 L) supplemented with ampicillin (50 μg/ml). The cultures were incubated at 37 °C until the OD_{600} reached 0.6 and then induced with IPTG (0.5 mM). The cultures were incubated for 16 h at 28 °C with shaking. The cells from the cultures were harvested by centrifugation (3200 rpm; 20 min) and the cell pellet was re-suspended in a lysis buffer (10 mM phosphate buffer pH 7.8, 150 mM NaCl) containing protease inhibitor cocktail (1 x Complete Mini -EDTA free; Roche) and Deoxyribonuclease I from bovine pancrease (Sigma; 0.1 mg/ml). The cells were then lysed by sonication using SoniPrep 150 Plus (MSE™, UK) pulsing for 15-20 sec with 2 min pauses at 4ºC followed by centrifugation (20000 rpm; 20 min). The supernatant was filtered with a syringe filter (0.45 μm) and loaded on to a column packed with Ni-NTA agarose™ (Qiagen; 5ml). The column was washed with wash buffer (10 mM phosphate buffer pH 7.8, 150 mM NaCl, 20 mM imidazole **40** pH 7.8) and then eluted with elution buffer (10 mM phosphate buffer pH 7.8, 150 mM NaCl, 500 mM Imidazole pH 7.8). Fractions were analysed by SDS-PAGE, pooled, dialysed against 20 mM HEPES, 1 mM $CaCl_2$, 150 mM NaCl and then concentrated using centrifugal filter unit (10000 Da, GE Healthcare) prior to loading on to a gel filtration column Superdex S-200(HR 16/60) column; Pharmacia Biotech) connected to an FPLC (ÄKTA purifier GE healthcare) pre-equilibrated with buffer (50 mM HEPES pH 8, 10 mM NaCl). The protein was eluted in the same buffer. Fractions containing the protein were analysed by SDS-

PAGE and concentrated (10000 Da, GE Healthcare) prior to measuring the concentration by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.). [55]

2.6.3 Freeze drying enzymes

The purified fluorinase (10 mM Tris-HCl pH 7, 30 mM NaCl) and *Tv*NH (20 mM HEPES, 1 mM $CaCl_2$, 150 mM NaCl) in their final buffers were freeze-dried (Christ Alpha 1-2 LD freeze drier) in eppendorf tubes at -54 ºC under 0.27 mBar vacuum for 17-20 h until the protein was a white amorphous solid was achived**.**

2.6.4 Freeze-drying with sucrose excipient system

Purified fluorinase and *Tv*NH in their respective buffers in an excipient system consisting of glycine (2.5 mg/ml), sucrose (50 mg/ml) and polysorbate 80 (0.01%) were freeze-dried (Christ Alpha 1-2 LD freeze drier) in eppendorf tubes at -54 ºC under 0.27 mbar vacuum for 17-20 h until a white amorphous solid was achieved**.**

2.6.5 Non-labelled enzyme assays 'cold reactions'

Fluorinase (1 mg, 2 mg and 3 mg) was incubated with SAM-Cl **22** (0.1 mM), KF (100 mM) and 20 mM Tris pH 7, 10 mM NaCl in a total volume of 500 μl for 20 min at 37 ºC. The reaction was heat denatured (95 °C; 10 min), centrifuged (13000 rpm; 5 min) and the supernatant was added to freeze-dried *Tv*NH (1 mg, 2 mg and 3 mg) and the resulting solution incubated at 37 ºC for 40 min. The reaction was analysed by HPLC for the production of FDA **23** and adenine **29**. The supernatant (20 μl) was injected into a Kinetix 5μm XB-C_{18} 100 Å (150 mm x 4.6 mm) column with a mobile phase of: 50 mM KH_2PO_4, 5% MeCN in water (solvent A) and 50 mM KH_2PO_4, 20% MeCN (solvent B) using a linear gradient from 0% to 100% solvent B in 20 min was used. The HPLC analysis was performed on a Varian Prostar system with a flow rate of 1 ml/min and the UV absorbance was measured at 260 nm.

2.6.6 Assaying both enzymes for stability

Fluorinase

Fluorinase (1 mg) was freeze-dried and stored at -20 °C for over 8 weeks. Each aliquot was removed every week and incubated with SAM-Cl **22** (0.2 mM) and potassium fluoride (100 mM) in a final volume of 1 ml. Sodium benzoate **44** (2 mM) was added to each reaction as an internal standard and the reactions were incubated at 37 °C. Samples (140 μl) were removed every 30 min for 2 h, heat treated (95 °C; 5 min) and centrifuged (13,000 rpm; 10 min). The supernatant (20 μl) was injected into a Kinetix 5μm XB-C_{18} 100 Å (150 mm x 4.6 mm) column with a mobile phase of : 50 mM KH_2PO_4, 5% MeCN in water (solvent A) and

50 mM KH_2PO_4, 20% MeCN (solvent B) using a linear gradient from 0% to 100% solvent B in 20 min. The HPLC analysis was performed on a Varian Prostar system with a flow rate of 1 ml/min and the UV absorbance was measured at 260 nm.

Nucleoside hydrolase

*Tv*NH (20 mg) was freeze-dried and stored at -20 °C for over 8 weeks. Each aliquot was removed and incubated with 5′-FDA **23** (0.5 mg) in a total reaction volume of 2 ml. 2-Fluoroethanol (1 mM) was added to each reaction as an internal standard, the reactions were incubated at 37 °C and samples (500 μl) were removed every 1 h for 4 h and heat treated (95 °C; 5 min) and centrifuged (13,000 rpm; 10 min). The supernatant (400 μl) was diluted with D_2O (300 μl). The reaction was assayed by ^{19}F {^1H} NMR on a Bruker AVIII-HD 500 MHz with a SmartProbe BBFO+ and all experiments were conducted at 298K.

2.6.7 Enzymatic preparation of [^{18}F]5-FDR [^{18}F]-37 (experiments were carried out by Dr Sergio Dall'Angelo in University of Aberdeen)

Freeze-dried fluorinase (6.5 mg) was incubated with SAM-Cl **22**(10 mM), L-AAO (1 mg) and a solution of [^{18}F]fluoride (200–350 MBq), to reach a final reaction volume of 600 μl. The reaction was incubated at 37 °C for 30 min and a sample of ~1-5 μl was removed and diluted with 45-49 μl of distilled water. The sample was heated (100 °C for 1 min) and centrifuged (13,000rpm, 5 min) and supernatant analysed by radio-HPLC (20 μl injection) for [^{18}F]F-**23** production.

To the reaction mixture freeze-dried *Tv*NH (39 mg) reconstituted in 50μl of water was added. The reaction was incubated at 37 °C for 1 h and an aliquot ~1-5 μl was removed and diluted with 45-49 μl of distilled water, heat treated (100 °C for 1 min) and centrifuged (13,000rpm, 5 min). The sample was analysed by radio-HPLC (20 μl injection) for [^{18}F]-**37** production. [56]

2.6.8 Modified method for enzymatic preparation of [^{18}F]5-FDR [^{18}F]-37 for tumour mice imaging (experiments were carried out by Dr Sergio Dall'Angelo at VU University Medical Centre in Amsterdam)

The previously mentioned method was altered to perform tumour mice imaging studies at VU University Medical Centre in Amsterdam. The freeze-dried fluorinase (10 mg) was resuspended in saline solution and a glass reactor vessel designed to fit the radiosynthesis system was used to perform the reaction. In the reactor 0.16 mM SAM-Cl **22** and 1 mg L-AAO was added to the fluorinase mixture followed by the addition of 600 μl of 10 GBq

[^{18}F]F$^-$ the reaction mixture was then incubated at 37 °C for 30 min. The entire reaction vessel was heated to 100 °C for 2 min to denature the protein. The reactor was then cooled to 37 °C and a buffered solution of freeze-dried *Tv*NH (60 mg) was added directly to the reaction mixture. The reaction was incubated at 37 °C for 30min and the reactor was heated as before. The reactor was cooled to 20 °C before the precipitated protein was removed by diluting with sterile water and filtering through a DLL filter (0.22μm). The sample was passed through a pre-equilibrated solid phase extraction (SPE) cartridge system consisting of PS-Ag$^+$ (Macherey-Nagel, Germany), C_{18}Plus (Waters, Netherlands) and PS-HCO_3 (Macherey-Nagel, Germany) to remove any unreacted [^{18}F]F$^-$, [^{18}F]-**23** and unreacted **22**. The purified [^{18}F]-**37** was eluted in sterile water (4 ml).[56]

HPLC monitoring used to detect the radiochemical and UV purity of [^{18}F]-**37**. The RP-HPLC analysis was carried out using a Phenomenex Luna column and radioactivity detector. Two detection methods were utilised, a refractive index detector and a Photodiode array (PDA). A synthetically prepared reference 5-FDR **37** was used to identify [^{18}F]-**37** peak. [56]

2.6.9 PET imaging of tumour induced mice using [^{18}F]5-FDR [^{18}F]-37 (experiments carried out by Dr Sergio Dall'Angelo at VU University Medical Centre in Amsterdam)

Four mice (nu/nu) were injected in their thighs with A431 cells (tumour inducing cells) and the tumours were left to grow for 1 week until they reached a size of 5–7 mm. Prior to the PET scan, mice were provided with plenty of food and then subsequently fasted for 2 h. They were anesthetised with isoflurane (2% in oxygen, 1 l/min *via* a nose mask) and cannulated in the jugular vein for administration of the radiopharmaceuticals. Following cannulation the mice were mounted onto the PET scanner and the body temperature was maintained between 35 °C and 37 °C. All animal experiments were carried out in compliance with Dutch law and approved by the VU University Animal Ethics Committee.[56]

Four mice were scanned using a double LSO/LYSO layer High Resolution Research Tomograph (HRRT; CTI/Siemens, Knoxville, USA). For attenuation and scatter correction, a transmission scan was acquired using a 740 MBq 2-dimensional (2D) fancollimated ^{137}Cs (662 keV) moving point source. Dynamic emission scans were then acquired and data were acquired in list mode and allocated into 28 frames, 20 frames of 60 s and 8 frames of 300 s. Following corrections for decay, dead time, scatter and random scans were reconstructed using an iterative-3D ordered-subsets weighted least-squares (3D-OSWLS) method. On the initial day [^{18}F]5-FDR [^{18}F]-**37** was administered at ~3 MBq per mouse and in the second day [^{18}F]FDG [^{18}F]-**32** was administered at a similar dose (as a control). PET image data

were analysed using the publically available software package, Amide 0.8.22. Volumes of interest (VOIs) were drawn by hand over the tumours, in the X, Y and Z direction. The VOIs were then projected onto all frames of the scan, resulting in time– activity curves for each scan, tumour and mouse. These time–activity curves were then normalized for the injected activity and animal weight to yield standardized uptake values (SUVs) evaluated over time. Whole body uptake was determined by a whole body VOI and this value was taken as 100% of the total activity, background uptake was defined in a VOI adjacent to the tumours.[56]

2.7 Chapter 2 References

1 M. A. Pysz, S. S. Gambhir and J. K. Willmann, *Clin. Radiol.*, 2010, **65**, 500–516.

2 Z. Li and P. S. Conti, *Adv. Drug Deliv. Rev.*, 2010, **62**, 1031–1051.

3 S. Fanti, M. Farsad and L. Mansi, *PET-CT Beyond FDG A Quick Guide to Image Interpretation*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

4 I. Y. Chen and J. C. Wu, *Circulation*, 2011, **123**, 425–443.

5 S. S. Gambhir, *Nat. Rev. Cancer*, 2002, **2**, 683–693.

6 S. Biswal and S. S. Gambhir, in *Gene and Cell Therapy*, ed. N.S. Templeton, Marcel Dekker, New york, 4th edn., 2015, 447–481.

7 P. W. Miller, N. J. Long, R. Vilar and A. D. Gee, *Angew. Chemie Int. Ed.*, 2008, **47**, 8998–9033.

8 H. H. Coenen, P. H. Elsinga, R. Iwata, M. R. Kilbourn, M. R. A. Pillai, M. G. R. Rajan, H. N. Wagner and J. J. Zaknun, *Nucl. Med. Biol.*, 2010, **37**, 727–740.

9 L. B. Been, A. J. H. Suurmeijer, D. C. P. Cobben, P. L. Jager, H. J. Hoekstra and P. H. Elsinga, *Eur. J. Nucl. Med. Mol. Imaging*, 2004, **31**, 1659–1672.

10 J. R. Grierson and A. F. Shields, *Nucl. Med. Biol.*, 2000, **27**, 143–156.

11 B. T. Christian, D. S. Lehrer, B. Shi, T. K. Narayanan, P. S. Strohmeyer, M. S. Buchsbaum and J. C. Mantil, *Neuroimage*, 2006, **31**, 139–152.

12 G. A. Digenis, S. H. Vincent, C. S. Kook, R. E. Reiman, G. A. Russ and R. S. Tilbury, *J. Pharm. Sci.*, 1981, **70**, 985–986.

13 J. R. Barrio, N. Satyamurthy, S. C. Huang, R. E. Keen, C. H. Nissenson, J. M. Hoffman, R. F. Ackermann, M. M. Bahn, J. C. Mazziotta and M. E. Phelps, *J. Cereb. Blood Flow Metab.*, 1989, **9**, 830–839.

14 D. Sorger, M. Patt, P. Kumar, L. I. Wiebe, H. Barthel, A. Seese, C. Dannenberg, A. Tannapfel, R. Kluge and O. Sabri, *Nucl. Med. Biol.*, 2003, **30**, 317–26.

15 V. W. Pike, *Trends Pharmacol. Sci.*, 2009, **30**, 431–440.

16 B. M. Gallagher, A. Ansari, H. Atkins, V. Casella, D. R. Christman, J. S. Fowler, T. Ido, R. R. MacGregor, P. Som, C. N. Wan, P. Wolf, D. E. Kuhl and M. Reivich, *J Nucl Med*, 1977, **18**, 990–996.

17 E. K. Pauwels, E. J. Sturm, E. Bombardieri, F. J. Cleton and M. P. Stokkel, *J. Cancer Res. Clin. Oncol.*, 2000, **126**, 549–559.

18 E. K. J. Pauwels, *Drugs Future*, 2001, **26**, 659.

19 R. E. Ehrenkaufer, J. F. Potocki and D. M. Jewett, *J. Nucl. Med.*, 1984, **25**, 333–337.

20 K. Hamacher, H. H. Coenen and G. Stöcklin, *J. Nucl. Med.*, 1986, **27**, 235–238.

21 L. Cai, S. Lu and V. W. Pike, *European J. Org. Chem.*, 2008, **2008**, 2853–2873.

22 P. Bjurling, Y. Watanabe, M. Tokushige, T. Oda and B. Långström, *J. Chem. Soc., Perkin Trans. 1*, 1989.

23 P. Bjurling, G. Antoni, Y. Watanabe and B. Långström, *Acta Chem. Scand*, 1990, 178–182.

24 E. Lui, R. Chirakal and G. Firnau, *J. Label. Compd. Radiopharm.*, 1998, **41**, 503–521.

25 M. Sasaki, M. Ikemoto, M. Mutoh, T. Haradahira, A. Tanaka, Y. Watanabe and K. Suzuki, *Appl. Radiat. Isot.*, 2000, **52**, 199–204.

26 G. Slegers, R. H. Lambrecht, T. Vandewalle, L. Meulewaeter and C. Vandecasteele, *J. Nucl. Med.*, 1984, **25**, 338–342.

27 E. S. da Silva, V. Gómez-Vallejo, J. Llop and F. López-Gallego, *Catal. Sci. Technol.*, 2015, **5**, 2705–2713.

28 C. Kim, D. J. Yang, E. E. Kim, A. Cherif, L. Kuang, C. Li, W. Tansey, C. W. Liu, S. C. Li, S. Wallace and D. A. Podoloff, *J. Pharm. Sci.*, 1996, **85**, 339–344.

29 L. Martarello, C. Schaffrath, H. Deng, A. D. Gee, A. Lockhart and D. O'Hagan, *J. Label. Compd. Radiopharm.*, 2003, **46**, 1181–1189.

30 H. Deng, S. L. Cobb, A. D. Gee, A. Lockhart, L. Martarello, R. P. McGlinchey, D. O'Hagan and M. Onega, *Chem. Commun. (Camb).*, 2006, 652–654.

31 M. Onega, J. Domarkas, H. Deng, L. F. Schweiger, T. a D. Smith, A. E. Welch, C. Plisson, A. D. Gee and D. O'Hagan, *Chem. Commun. (Camb).*, 2010, **46**, 139–141.

32 M. Winkler, J. Domarkas, L. F. Schweiger and D. O'Hagan, *Angew. Chemie - Int. Ed.*, 2008, **47**, 10141–10143.

33 H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, *Angew. Chemie Int. Ed.*, 2006, **45**, 759–762.

34 M. Onega, *PhD thesis, Studies and application of the enzymes of fluorometabolite biosynthesis in Streptomyces cattleya*, University of St Andrews, 2009.

35 T. Smith, M. Onega and D. O'Hagan, *Unpubl. results*.

36 J. Marik, A. Ogasawara, B. Martin-McNulty, J. Ross, J. E. Flores, H. S. Gill, J. N. Tinianow, A. N. Vanderbilt, M. Nishimura, F. Peale, C. Pastuskovas, J. M. Greve, N. van Bruggen and S. P. Williams, *J. Nucl. Med.*, 2009, **50**, 982–990.

37 D. E. Ponde, C. S. Dence, N. Oyama, J. Kim, Y. Tai, R. Laforest, B. A. Siegel and M. J. Welch, *J. Nucl. Med.*, 2007, **48**, 420–428.

38 X.-G. Li, J. Domarkas and D. O'Hagan, *Chem. Commun. (Camb).*, 2010, **46**, 7819–7821.

39 C. Petersen and L. B. Møller, *J. Biol. Chem.*, 2001, **276**, 884–894.

40 J. Ogawa, S. Takeda, S. X. Xie, H. Hatanaka, T. Ashikari, T. Amachi and S. Shimizu, *Appl. Environ. Microbiol.*, 2001, **67**, 1783–1787.

41 X. Kang, Y. Zhao, D. Jiang, X. Li, X. Wang, Y. Wu, Z. Chen and X. C. Zhang, *Biochem. Biophys. Res. Commun.*, 2014, **446**, 965–970.

42 W. Shi, V. L. Schramm and S. C. Almo, *J. Biol. Chem.*, 1999, **274**, 21114–21120.

43 D. W. Parkin, B. A. Horenstein, D. R. Abdulah, B. Estupiñán and V. L. Schramm, *J. Biol. Chem.*, 1991, **266**, 20658–20665.

44 J. E. Kurtz, F. Exinger, P. Erbs and R. Jund, *Curr. Genet.*, 1999, **36**, 130–136.

45 J. M. C. Ribeiro and J. G. Valenzuela, *Insect Biochem. Mol. Biol.*, 2003, **33**, 13–22.

46 K. K. W. Siu, K. Asmus, A. N. Zhang, C. Horvatin, S. Li, T. Liu, B. Moffatt, V. L. Woods and P. L. Howell, *J. Struct. Biol.*, 2011, **173**, 86–98.

47 M. Szuwart, E. Starzyńska, M. Pietrowska-Borek and A. Guranowski, *Phytochemistry*, 2006, **67**, 1476–1485.

48 M. Porcelli, L. Concilio, I. Peluso, A. Marabotti, A. Facchiano and G. Cacciapuoti, *FEBS J.*, 2008, **275**, 1900–1914.

49 C. Minici, G. Cacciapuoti, E. De Leo, M. Porcelli and M. Degano, *Biochemistry*, 2012, **51**, 4590–4599.

50 F. Giannese, M. Berg, P. Van der Veken, V. Castagna, P. Tornaghi, K. Augustyns and M. Degano, *Acta Crystallogr. D. Biol. Crystallogr.*, 2013, **69**, 1553–1566.

51 D. W. Parkin, *J. Biol. Chem.*, 1996, **271**, 21713–21719.

52 B. Estupiñán and V. L. Schramm, *J. Biol. Chem.*, 1994, **269**, 23068–23073.

53 B. Giabbai and M. Degano, *Structure*, 2004, **12**, 739–749.

54 L. Muzzolini, W. Versées, P. Tornaghi, E. Van Holsbeke, J. Steyaert and M.

Degano, *Biochemistry*, 2006, **45**, 773–782.

55 W. Versées, K. Decanniere, R. Pellé, J. Depoorter, E. Brosens, D. W. Parkin and J. Steyaert, *J. Mol. Biol.*, 2001, **307**, 1363–1379.

56 S. Dall'Angelo, N. Bandaranayaka, A. D. Windhorst, D. J. Vugts, D. van der Born, M. Onega, L. F. Schweiger, M. Zanda and D. O'Hagan, *Nucl. Med. Biol.*, 2013, **40**, 464–470.

57 A. Illanes, *Enzyme Biocatalysis*, Springer Netherlands, Dordrecht, 2008.

58 A. Illanes, *Electron. J. Biotechnol.*, 1999, **2**.

59 A. Bommarius and B. Riebel-Bommarius, *Biocatalysis: Fundamentals and Applications*, Wiley-Blackwel, 1st edn., 2004.

60 Fagain, Ciaran O., *Protein Stability And Stabilization Of Protein Function (biotechnology Intelligence Unit)*, Springer-Verlag, 1997.

61 C. Schaffrath, H. Deng and D. O'Hagan, *FEBS Lett.*, 2003, **547**, 111–114.

62 K. C. Duong-Ly and S. B. Gabelli, *Methods Enzymol.*, 2014, **541**, 85–94.

63 C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer and J. H. Naismith, *Nature*, 2004, **427**, 561–565.

64 A. F. Weir, *Enzymes of Molecular Biology*, Methods in Molecular Biology, 16(1) edn., 1993.

65 A. Goeminne, M. McNaughton, G. Bal, G. Surpateanu, P. Van Der Veken, S. De Prol, W. Versées, J. Steyaert, A. Haemers and K. Augustyns, *Eur. J. Med. Chem.*, 2008, **43**, 315–326.

66 H. E. Bergmans, I. M. van Die and W. P. Hoekstra, *J. Bacteriol.*, 1981, **146**, 564–570.

67 B. Nastasijevic, N. A. Becker, S. E. Wurster and L. James Maher, *Biochem. Biophys. Res. Commun.*, 2008, **366**, 420–425.

68 J. F. Carpenter, M. J. Pikal, B. S. Chang and T. W. Randolph, *Pharm. Res.*, 1997, 14, 969–975.

69 X. Tang and M. J. Pikal, *Pharm. Res.*, 2004, **21**, 191–200.

70 I. Roy and M. N. Gupta, *Biotechnol. Appl. Biochem.*, 2004, **39**, 165–177.

71 W. Liu, D. Q. Wang and S. L. Nail, *AAPS PharmSciTech*, 2005, **6**, E150–E157.

72 J. F. Carpenter and J. H. Crowe, *Cryobiology*, 1988, **25**, 459–470.

73 J. F. Carpenter and J. H. Crowe, *Biochemistry*, 1989, **28**, 3916–3922.

74 A. Bansal, S. Lale and M. Goyal, *Int. J. Pharm. Investig.*, 2011, **1**, 214–221.

75 L. W. Parks and F. Schlenk, *J. Biol. Chem.*, 1958, **230**, 295–305.

76 M. L. Stolowitz and M. J. Minch, *J. Am. Chem. Soc.*, 1981, **103**, 6015–6019.

77 D. Q. Wang, J. M. Hey and S. L. Nail, *J. Pharm. Sci.*, 2004, **93**, 1253–1263.

78 A. Capolongo, A. A. Barresi and G. Rovero, *J. Chem. Technol. Biotechnol.*, 2003, **78**, 56–63.

79 M. J. Pikal, in *Formulation and Delivery of Proteins and Peptides*, American Chemical Society, 1994, 120–133.

80 L. A. Mounter, *Radiat. Res.*, 1960, **12**, 487–494.

81 J. Theobald, A. Hanby, K. Patel and S. E. Moss, *Br. J. Cancer*, 1995, **71**, 786–788.

3. Identification of other fluorinases

3.1 *S. cattleya* fluorometabolite gene cluster

The physical clustering of genes involved in secondary metabolite biosynthetic pathways in bacteria, filamentous fungi and in some plants is a common occurrence as identified by whole genome sequencing.[1] Genes encoding consecutive steps in a biosynthetic pathway are prone to group together in this manner on the chromosome, although their organisation within these clusters has been found to be highly variable.[2] Secondary metabolite biosynthetic gene clusters most often operate as self-contained cassettes comprising of the genes required for various processes, such as; forming the frame of the secondary metabolite, subsequent modifications of this metabolite, as well as pathway specific regulators and often resistance genes for coping with any toxic traits of the final product. These secondary metabolite gene clusters often lay dormant and are only expressed in response to specific metabolic stress.[2,3]

In order to explore if the genes involved in fluorometabolite production by *S. cattleya* might be clustered, Dr Jonathan Spencer and co-workers (University of Cambridge) used genomic library screening to identify the genes clustered around the fluorinase gene (*flA*) of *S. cattleya*. The study revealed a 12kb gene cluster (the 'Spencer cluster', Figure 1) containing twelve open reading frames (ORF's) centered around the *flA* gene.[4]

To determine the function of the genes, each one was knocked-out and the resulting mutant was analysed for its ability to produce fluorometabolites. When the *flA* gene was knocked-out, the bacterium's ability to produce 4-fluorothreonine (4-FT) **14** and fluoroacetate (FAc) **10** was lost. Further investigation found that the addition of synthetically prepared 5′-FDA **23** to cultures of this mutant restored the production of these fluorometabolites. The second enzyme in this pathway, PNP (Scheme 5; Chapter 1) is encoded by the *flB* gene, which lies adjacent to the *flA* gene. Inactivation of *flB* was found to arrest the fluorometabolite production similar to *flA* inactivation.[4]

The Spencer cluster also contains genes involved in the regulation of this fluorometabolite pathway including *flI*, encoding for an adenosylhomocysteine hydrolase involved in the removal of excess *S*-adenosyl homocysteine (SAH) **46**, a potent inhibitor of the fluorinase. The gene *flK* encodes for a 139 amino acid thioesterase, which catalyses the selective breakdown of fluoroacetyl-CoA. This enzyme is proposed to confer resistance to *S. cattleya* against FAc **10** toxicity.[5] The knockout mutant of *flK* was

unable to produce any fluorometabolite in the presence of fluoride, confirming a tight regulation of fluorometabolite production. The genes *flE*, *flF* and *flG* are also located within this cluster and were found to encode for DNA binding regulatory proteins.[4]

The expression of many genes involved in this fluorometabolite pathway was found to be dependent on fluoride ion levels. For instance; an increase in fluoride concentrations was shown to upregulate the expression of *flG* and *flA* gene over 2 fold.[6] The *flH* gene was found to encode a transmembrane Na^+/H^+ antiporter. Na^+/H^+ antiporters transport protons out of the cell in exchange for Na^+, which is essential for regulation of cell volume and pH. It was observed that fluoride uptake in resting cells of *S. cattleya* peaked at pH 6.0 and dropped significantly as pH increased, indicating that a certain level of extracellular H^+ was required for transport of F^- in to cells. It follows that Na^+/H^+ antiporters may buffer the accumulation of HF by exchanging intracellular H^+ for a Na^+, and thus regulate the intracellular pH and net accumulation of NaF.[7] However, a *flH* gene knockout mutant did produce fluorometabolites suggesting that *flH* is not essential for fluorometabolite biosynthesis.[4]

ORF	Start–Stop (bp)	Length (aa)	Function / Homology
flE	127–795c	222	DNA binding regulatory protein
flD	854–1504c	216	Dehalogenase / Phosphatase
flC	1845–3038	397	MFS permease
flB	3053–3952c	299	5′-FDA phosphorylase
flA	4172–5071	299	5′-FDA synthase (fluorinase)
flF	5196–5753	185	DNA binding regulatory protein
flG	5947–6651c	234	DNA binding regulatory protein
flH	6648–8051c	467	Na^+ / H^+ antiporter
flI	8317–9786	489	Adenosylhomocysteine hydrolase
flJ	9802–10197	131	Regulatory protein
flK	10172–10591c	139	Fluoroacetyl-CoA thioesterase
flL	10696–11373c	225	DNA binding regulatory protein

Figure 1: The Spencer cluster, showing genes encoding for enzymes involved in fluorometabolite biosynthesis in *S. cattleya*. Image reproduced with permission from O'Hagan *et al.* [8]

Two research groups collaborated in completing the full genome sequence of *S. cattleya* which was released in 2011.[9,10] The complete genome sequence of *S. cattleya* consists of a linear chromosome of 6.28 Mb (containing 5822 genes) and a linear mega-plasmid of 1.81 Mb (containing 1747 genes) (Figure 2).

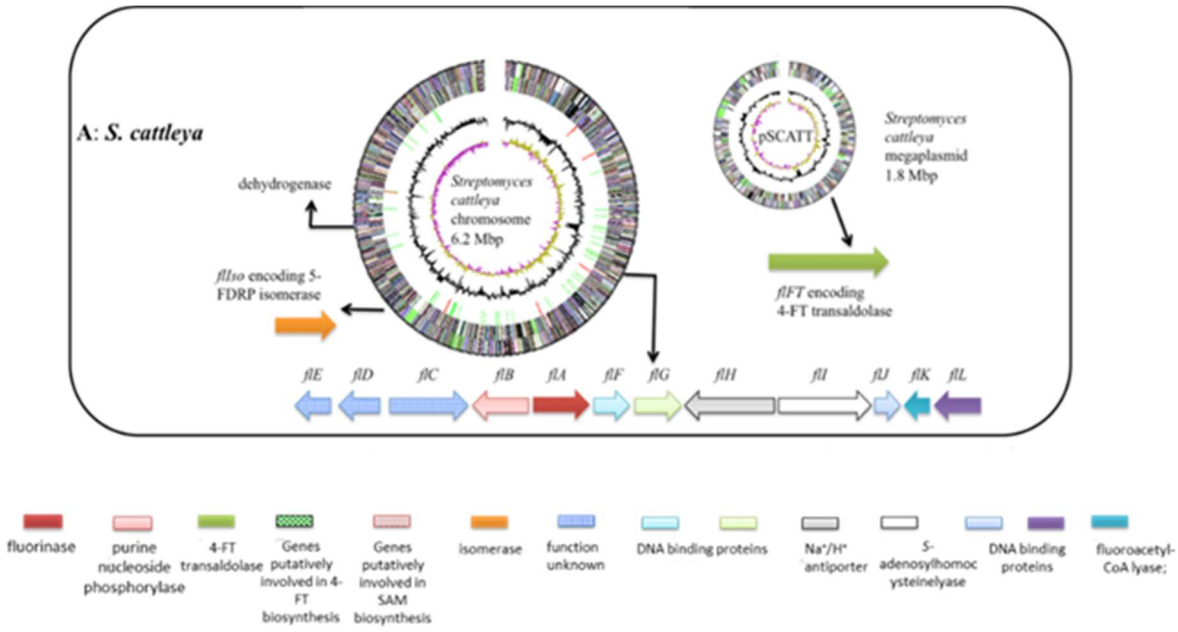


Figure 2: The genome map showing the arrangement of biosynthetic genes involved in fluorometabolite biosynthesis of *S. cattleya*. Image reproduced and modified with permission from O'Hagan *et al.*[11]

Genes encoding the remaining enzymes involved in the biosynthetic pathway of fluorometabolites in *S. cattleya* are absent from the Spencer cluster. Full genome sequencing has found that they are distributed over the chromosome.

The mega-plasmid was revealed to contain only two genes involved in the fluorometabolite biosynthesis, the 4FT-transaldolase and a putative aldolase, whereas most of the genes required for fluorometabolite production are located on the chromosome or within the Spencer cluster. The *flFT* gene knock-out mutant was able to produce FAc **10** but had lost the ability to produce 4-FT **14** consistent with a branch point on the fluorometabolite pathway (Scheme 5; Chapter 1).[4]

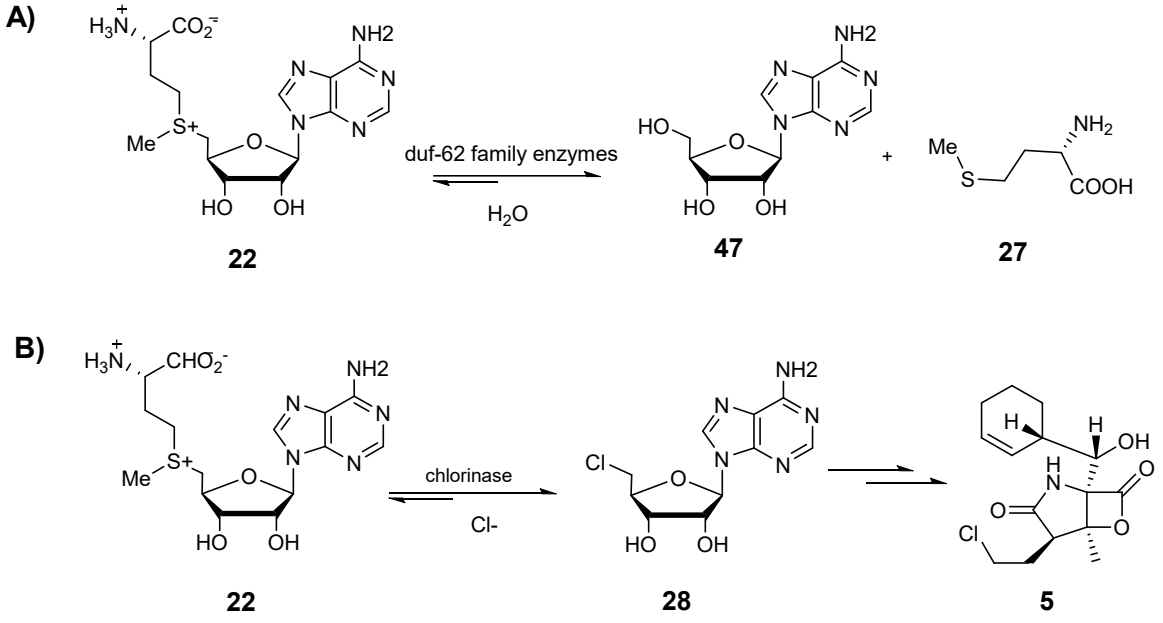
The genes encoding for isomerase (*flIso*) and dehydrogenase enzymes are located on the chromosome, but they are distant from the Spencer cluster. Inactivation of the *flIso* gene also generated mutants unable to produce any fluorometabolite.[4,7]

3.1.1 Enzymes with homology to fluorinase

The fluorinase is a unique enzyme in that it catalyses a nucleophilic substitution reaction involving fluoride ion and the C-5′ carbon of SAM **22**, as discussed in Chapter 1.[12] Only a small number of genes possess some sequence similarities to the fluorinase, these consist of the larger class of proteins expressed by the domains of unknown fuction-62 (duf-62) genes present in many prokaryotes and the chlorinase isolated from the marine bacterium

Salinispora tropica. The duf-62 family genes have ~25-32 % gene homology to the fluorinase whereas chlorinase has ~40% gene homology.[13]

Over 300 duf-62 genes are evident in publicly available genome sequences with their distribution limited to mostly extremophiles and pathogen-related microbes. Duf-62 genes encode for a protein, which mediates the hydrolytic conversion of SAM **22** to adenosine **47** and L-methionine in what is believed to be an S_N2 reaction requiring hydroxide as the nucleophile (Scheme 1A). Although functionally the duf-62 enzyme mediates a reaction very similar to that of the fluorinase and chlorinase, they are distinctive both in the organisation of their active site and the mechanism by which they activate the nucleophile for attack on **22**. Fluorinase and chlorinase activate halide ions through dehydration *via* a combination of hydrogen bonds to active site residues which is further reinforced by the binding of **22**. However the duf-62 proteins use a highly conserved amino acid triad (His-Arg-Asp) which assists in activating water towards nucleophilic attack on the adenosine **47** C-5′ carbon of SAM **22**.[13,14] Mutation of His-127 in the catalytic triad was shown to retain some enzyme activity, however any mutation to either Asp-68 or Arg-75 of the catalytic triad abolished the enzyme activity completely.



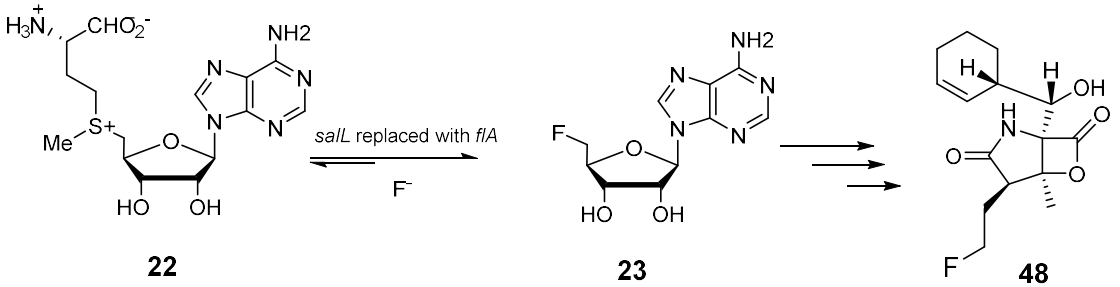
22 **47** **27**

22 **28** **5**

Scheme 1: A) The duf-62 family of enzymes use water as a nucleophile to convert SAM **22** to adenosine **47**. **B)** Chlorinase (from *Salinispora tropica*) catalyses the conversion of **22** to ClDA **28**, several subsequent enzymatic steps lead to the production of the final chlorometabolite, salinosporamide A **5**.

The chlorinase enzyme was isolated during a genome mining study carried out in 2008 by the Bradley Moore group at the Scripps Institution of Oceanography, San Diego. An entire gene cluster was identified from marine actinomycete *Salinospora tropica* in which the

salL gene encodes for the SAM **22**-dependent chlorinase. This enzyme mediates the reaction between chloride ion and **22** to form 5′-chlorodeoxy adenosine (ClDA) **28**, as shown in Scheme 1. The chlorinase is capable of utilising both Br^- and I^- for halogenation, however it is unable to utilise fluoride as the halide source for these reactions. The chlorinase is the first enzyme in a chlorometabolite biosynthesis pathway, in which the final product is salinosporamide A **5**, which is an effective proteasome inhibitor and a potential anticancer drug.[15] The chlorinase is a homotrimer similar to the fluorinase although the latter organises as a hexamer constructed as a dimer of trimers. Although the reaction catalysed by the fluorinase and chlorinase follow similar mechanisms, the key residues co-ordinating the halide ions are different. During fluorinase-catalysed halogenation, Ser-158 aids in fluoride ion binding, for the chlorinase Gly-131 serves the same function.[16] Site directed mutagenesis has been carried out on the chlorinase replacing Gly-131 with serine in order to analyse if this mutant would be able to accommodate a fluoride ion, however this was not the case and the mutant was found to inhibit chlorination activity. The chlorinase also lacks a 21 amino acid loop unique to the fluorinase enzyme, suggesting the importance of this motif in the selectivity of fluoride over other halides. Enzymes located downstream of the chlorinase involved in salinosporamide biosynthesis are capable of accepting fluorinated substrates evident by the production of fluorosalinosporamide **48** on addition of **23** (Scheme 2).[17] A genetically engineered *S. tropica* mutant with *salL* replaced by *flA* (encoding for fluorinase from *S. cattleya*) was cultured in the presence of inorganic fluoride. This mutant was capable of producing **48** as a novel fluorinated natural product. Despite this observation *S. tropica* was seen to be extremely sensitive to fluoride concentrations during these experiments, limiting the prospect of large scale production of **48**.[17]



Scheme 2: Novel fluorinated natural product, fluorosalinosporamide A **48** produced by genetically altered *S. tropica* mutant.[17]

The fluorinase isolated from *S. cattleya* was the only known fluorinase enzyme for over a decade, since its discovery in 2003.[18] However in 2014, three more fluorinases were identified, through genome mining, one of which was found in *Streptomyces* sp. MA37

recently isolated from Ghana.[11] The crystal structure, characteristics and kinetics of the new fluorinase from *Streptomyces* sp. MA37 in comparison to other fluorinases are discussed below.

3.2 Characterisation of a fluorinase isolated from *Streptomyces* sp. MA37.

3.2.1 Isolation and growth of *Streptomyces* sp. MA37

Streptomyces sp. MA37 is a soil bacterium, isolated from the rhizosphere of a *Moraceae* Bark Cloth tree (*Antiaris toxicaria*, Africa) in 2011 by Dr Hai Deng of the University of Aberdeen and Dr Kwaku Kyeremeh of the University of Ghana. The tree from which the *Streptomyces* sp. MA37 was isolated was growing in the grounds of the University of Ghana Botanical Gardens.

Figure 3: *Streptomyces* sp. *MA37* cultured on ISP2 media.

The *Streptomyces* sp. MA37 bacterium was cultured (Figure 3) by Dr Hai Deng and Zhiwei Qin at the University of Aberdeen. The cultures were supplemented with 2 mM potassium fluoride and grown at 28 °C for 7 days, following which the supernatant was analysed for fluorometabolites by $^{19}F\{^{1}H\}$ NMR. The extract was observed to contain FAc **10** similar to *S. cattleya* and therefore it was of interest to analyse the genome of *Streptomyces* sp. MA37 for genes related to synthesis of FAc **10**.

Full genome sequencing of *Streptomyces* sp. MA37 was carried out by Dr Yi Yu at Wuhan University, using paired-end sequencing (Illumina). In addition, 16S ribosomal RNA (rRNA) sequencing of *Streptomyces* sp. MA37 was performed to identify the evolutionary relationship of this organism relative to other *Streptomyces* species. The 16S rRNA gene sequences of the small subunit of prokaryotic ribosomes are highly conserved within prokaryotes. This relationship is exploited in creating phylogenetic trees and identifying novel bacteria.[19] A phylogenetic tree was reconstructed based on the 16S rRNA gene sequences of *Streptomyces* sp. MA37 where it was observed to share a distant relationship

to a clade consisting of *S. cattleya*, *S. coelicolor, S. avermitilis* and *S. scabiei*. Consequently *Streptomyces* sp. MA37 was proposed to have diverged at an earlier branch than *S. cattleya*, from a closely related common ancestor (Figure 4).[11]

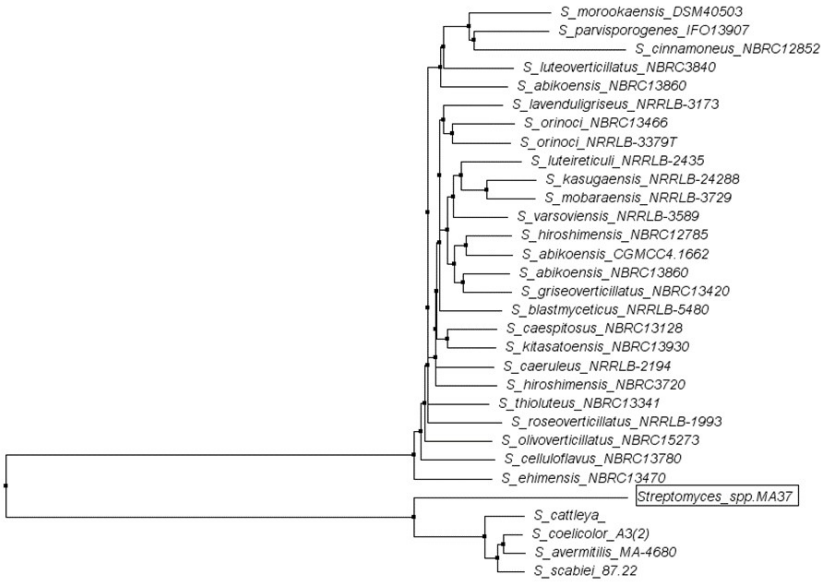


Figure 4: Phylogenetic relationship between *Streptomyces* sp. *MA37* and other *Streptomyces* strains based on 16S rRNA gene sequencing. Image reproduced with permission from Deng *et al*.[11]

3.2.2 Genome analysis of *Streptomyces* sp. *MA37*

Full genome sequencing of *Streptomyces* sp. *MA37* identified a *flA1* gene, which possesses an 87% homology to the *flA* gene of *S. cattleya*. *flA1* also contained a signature 21 amino acid loop, which is unique to the fluorinase (*flA*) and absent from the chlorinase and duf-62 enzymes.

Streptomyces sp. MA37 was also found to contain the 4-FT **14** transaldolase gene (*flFT*1), responsible for catalysing the final enzymatic step in the biosynthetic pathway of **14** in *S. cattleya.*[8] This suggested that *Streptomyces* sp. MA37 was able to synthesise **14** in addition to its previously established FAc **10** production. When comparing the gene organisation of *Streptomyces* sp. MA37 relative to that of *S. cattleya*, the *flFT*1 gene was located adjacent to the *flA*1 gene. This is a significant observation, as the *S. cattleya* gene is located in a megaplasmid isolated from the chromosome. This indicated advanced gene clustering in *Streptomyces* sp. MA37. The ability of *Streptomyces* sp. MA37 to produce **14** was subsequently confirmed by Dr Long Ma at the University of St Andrews, by analysing the

supernatant of cultures *via* $^{19}F\{^{1}H\}$ NMR. Production of **14** was observed which was further confirmed by GC-MS fragmentation (Figure 5).[11]

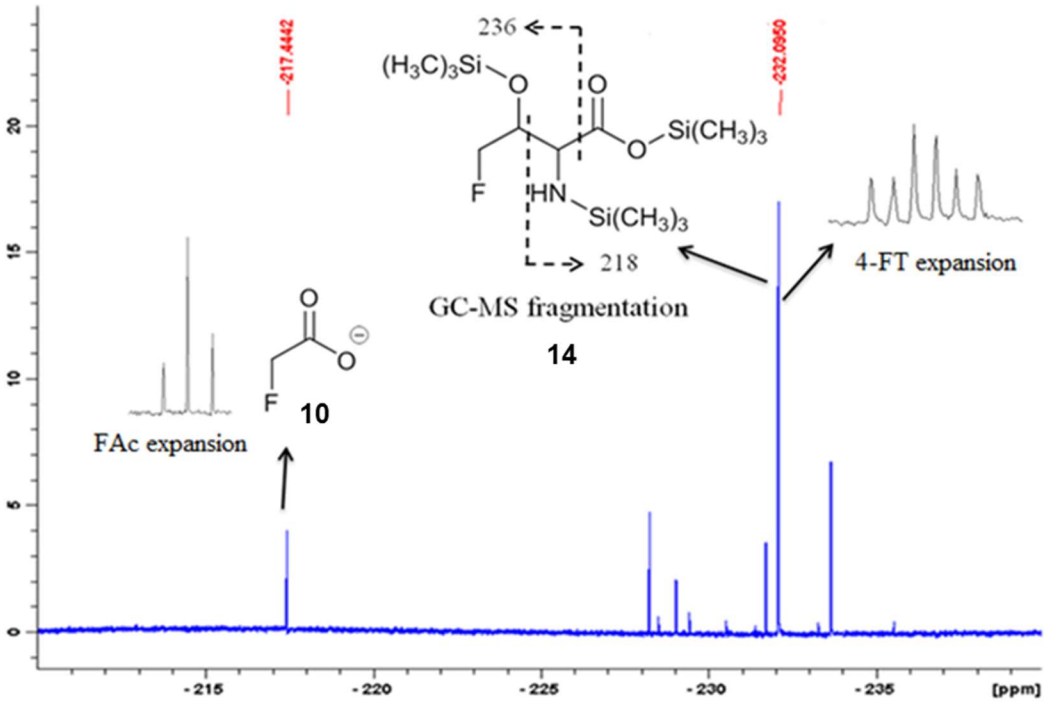
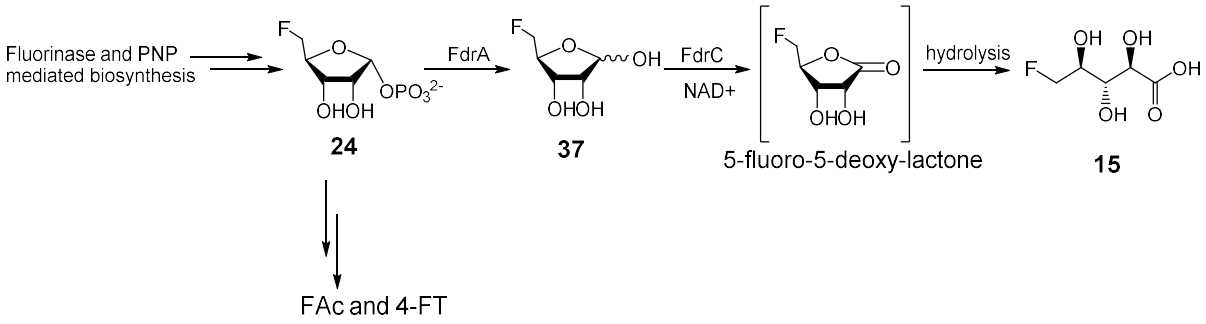


Figure 5: $^{19}F\{^{1}H\}$ NMR spectrum obtained from cultures of *Streptomyces* sp. MA37 showing the production of **14** and **10** along with new fluorometabolites.[11]

^{19}F NMR analysis of a CFE of *Streptomyces* sp. MA37 revealed the production of several new fluorometabolites in addition to **10** and **14**. One of these fluorometabolites was identified as (2*R,* 3*S,* 4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid (5-FHPA) **15** (Figure 4, Chapter 1). 5-FHPA **15** production was also observed during these experiments when 5-FDR **37** was added to the CFE, suggesting a **37** mediated biosynthesis. Detailed analysis of the genome of *Streptomyces* sp. MA37 revealed an *fdr* gene cluster which possesses a homologous gene cluster to that which has been extensively studied in *S. tropica* (the salinosporamide A producer). The *fdr* gene cluster consists of *fdrA*, encoding a metal-dependent phosphoesterase, FdrA with high sequence identity (56%) to SalN of the salinosporamide biosynthetic pathway. Immediately downstream of *fdrA*, two ORF's, *fdrB* and *fdrC* was found to encode two proteins FdrB and FdrC which share high sequence identity to SalH (68%) and SalM (69%), respectively. Therefore, a related biosynthetic pathway analogous to that in *S. tropica* was proposed for *Streptomyces* sp. MA37, which was subsequently confirmed by ^{19}F NMR.[11] Following this a candidate gene *fdrC* expressing an NAD^{+} dependent dehydrogenase was investigated as the enzyme responsible for the

conversion of 5-FDR **37** to 5-FHPA **15**. Indeed over-expressed FdrC was found to catalyse the NAD^+ dependent oxidation of 5-FDR **37** to an intermediate 5-fluoro-5-deoxy-lactone which then underwent hydrolysis to form **15** (Scheme 3) and these reactions could be followed by ^{19}F-NMR.[20]



Scheme 3: A) The biosynthesis pathway of 5-FHPA **15** in *Streptomyces* sp. MA37.[20]

3.2.3 Gene mining for the *flA* gene

In 2012, the genome sequence of a hospital pathogen, *Nocardia brasiliensis* HUJEG-1 was added to the National Center for Biotechnology Information GenBank database.[10,11] Genome analysis of *Nocardia brasiliensis* revealed a *flA*2 gene at an ORF (YP_006809254) which had 81% homology to *flA* of *S. cattleya.* The predicted amino acid sequence of the *flA2* gene contained the 21 amino acid loop along with some conserved amino acids found within the active site of the *S. cattleya flA* gene (Figure 6). A gene encoding a 4-FT **14** transaldolase (*flFT*2) was also located close to the *flA*2 gene, similar to the gene organisation of *Streptomyces* sp. MA37 (Figure 7C). *N. brasiliensis* is an actinomycete belonging to family *Norcardia*, a soil bacterium which degrades soil matter. It is an opportunistic bacterium which can cause sporadic diseases in humans and animals. In humans, it can cause nocardiosis, an infectious disease affecting the lungs or other organs, mostly affecting immunocompromised patients.[21]

Figure 6: Sequence identity alignment of FlA from *S. cattleya* (WP_014144878.1), FlA 1 from *Streptomyces* sp. MA37 (CDH39444.1), FlA2 from *Nocardia brasiliensis* (WP_014985135.1), FlA3 from *Actinoplanes* sp. N902-109 (WP_015619887.1), FlA4 from *Streptomyces xinghaiensis* (WP_019711456.1) and SalL (chlorinase from *S. tropica*). Over 80% sequence identity is represented by dark blue and less than 40% identity is represented by white, lighter shades represent moderate sequence identity. The 21 amino acid loop unique to fluorinases is highlighted by the red box.

Furthermore, in 2012 an *Actinoplanes* sp. N902-109 was found to contain a *flA*3 gene with 80% homology to the *flA* gene of *S. cattleya* which also possessed the conserved 21 amino acid loop (Figure 6). *Actinoplanes* sp. N902-109 is an actinomycete, first isolated from a soil sample collected in Shizuoka Prefecture, Japan in 1995 and investigated for its ability to produce large quantities of rapamycin in comparison with *Streptomyces hygroscopicus*.[22] The genome sequence of this organism also revealed a 4-FT transaldolase (*flFT3*), organised in close proximity to the *flA*3 gene, suggesting the possibility of the biosynthesis of **14**. (Figure 7D).[7]

More recently in 2014, gene mining revealed a *flA4* gene in a marine bacterium, *Streptomyces xinghaiensis*.[23] This *flA4* gene also contained the 21 amino acid loop specific to fluorinases (Figure 6). This bacterium was isolated and cultured for its ability to produce a novel broad-spectrum antibiotic, xinghaiamine A, and it was also discovered to be a fluoroacetate producer. *S. xinghaiensis* was first isolated in 2009 from a marine sediment sample collected from Dalian, China.[24] This *S. xinghaiensis* gene cluster (Figure 7E) also contained a truncated *flFT4* gene located adjacent to *flA4* which suggests that this is a

pseudogene,[25] particularly as the strain could not produce **14** when grown in the presence of fluoride. The FAc **10** production by *S. xinghaiensis* was also found to be sea salt dependent.[23]

In *S. cattleya* the PNPase (*flB*) is located immediately adjacent to the *flA* gene. Upon comparing the genome sequences of *Streptomyces* sp. MA37, *Nocardia brasiliensis*, *Actinoplanes* sp., and *S. xinghaiensis* it was apparent that all four organisms possess the PNPase gene (*flB* 1, 2, 3, 4 respectively) located in close proximity to the fluorinase gene (Figure 7).[11] The arrangement of the genes involved in fluorometabolite biosynthesis in the more recently discovered fluoroacetate producing organisms appear to have a higher degree of clustering when compared to that of *S. cattleya.*[11]

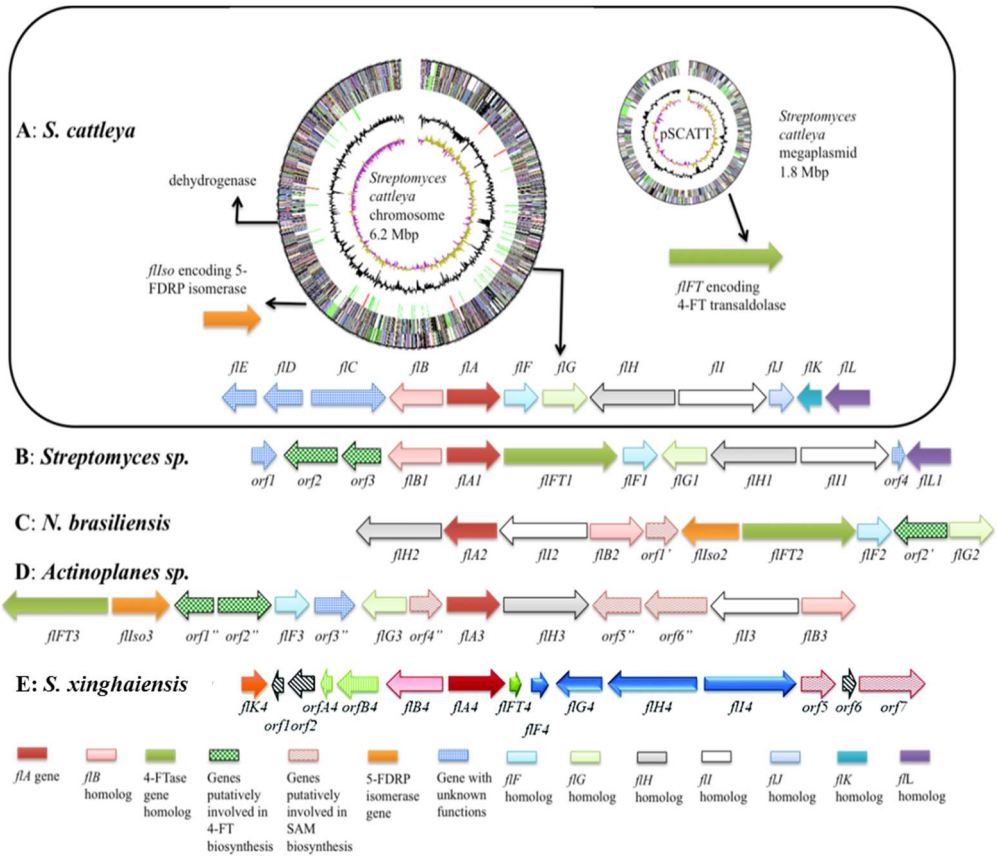


Figure 7: A) Genome map of *S. cattleya* showing the location of fluorometabolite biosynthetic genes. 4-FTtransalodolase and isomerase are remote from the Spencer cluster containing the *flA* gene. **B)** *Streptomyces* sp. MA37 **C)** *N. brasiliensis*; **D)** *Actinoplanes* sp. N902-109 **E)** *S. xinghaiensis*. Image reproduced and modified with permission from Deng *et al*.[23]

Through these studies, three new candidate fluorinase genes were identified (*flA*2, *flA*3 and *flA*4) *via* gene mining in three different species of bacteria, each belonging to different

families in the actinomycetales order. Genes isolated with similarities to the fluorinase could possibly be pseudogenes which possess similar gene sequences but are non-functional. Therefore it was necessary to express and purify the gene products of *flA*1, *flA*2 and *flA*3 and determine if they are functional fluorinases. If all three fluorinase homologs are functional it would suggest that the fluorinase gene either evolved independently (parallel evolution) in three different families of actinomycetales or was acquired through an ancient common ancestor (collateral evolution).[26]

3.3 Project aims

The aim was to express, purify and characterise the fluorinase (*flA1*) of *Streptomyces* sp. MA37 and establish it as a 'true' fluorinase.[12]

As such, an initial objective was to clone the *flA1* gene into a suitable vector and over-express it in *E.coli*. The over-expressed protein would then be purified *via* affinity chromatography and assayed for fluorinase activity to confirm it as a functional fluorinase. It was also important to analyse the kinetic parameters (K_m, V_{max}) of the *Streptomyces* sp. MA37 fluorinase, and attempt crystallisation of the protein in order to compare the crystallography structures with that of the *S. cattleya* fluorinase.

3.4 Results and discussion

3.4.1 Isolation and Cloning of *flA1*

The gross DNA was extracted (described in detail in Experimental 3.6.1) from a *Streptomyces* sp. MA37 culture which was donated by Dr Hai Deng at the University of St Andrews. The purified DNA and PCR primers were used to set up PCR reactions to amplify the *flA*1 gene (sequence of the gene is mentioned in Appendix 3) as detailed in the Experimental 3.6.2. The pEHISTEV plasmid was chosen as the vector for this study, which is derived from a pET vector containing an amino terminal hexahistidine tag (His_6-tag) and a tobacco etch virus (TEV) protease recognition site [27] between the His_6-tag and the inserted gene. This design allows for the expression of a protein of interest with a TEV protease cleavable His_6-tag.[28] The resultant PCR product (900 bp) and the pEHISTEV plasmid were digested with restriction enzymes NcoI and HindIII and analysed by DNA gel electrophoresis (Figure 8A). The digested plasmid and PCR products were gel purified.

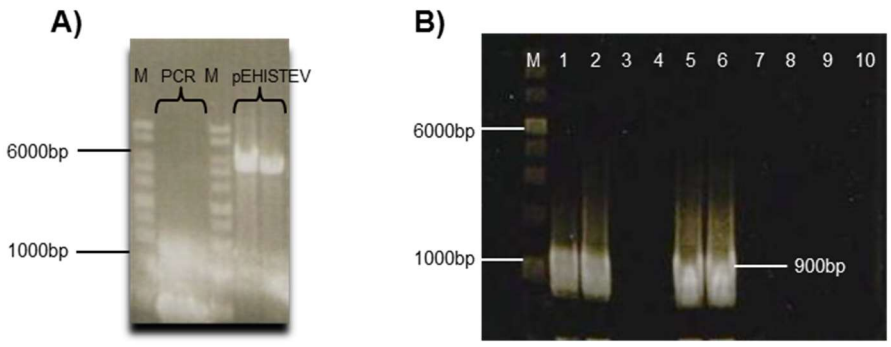


Figure 8: A) DNA gel electrophoresis of PCR product digested with NcoI and HindIII (900 bp); pEHISTEV digested with NcoI and HindIII (5369bp); M- marker **B)** Colony PCR of 10 *E.coli* DH5α transformants out of which lanes 1,2,5 and 6 are positives with the PCR product ligated to the vector.

The digested PCR product was ligated into digested pEHISTEV[28] and the ligation mixture was transformed[29] into DH5α *E.coli* cells[13], plated on agar containing kanamycin. Twelve transformants were isolated and analysed for the presence of the PCR product (900 bp) using colony PCR and DNA gel electrophoresis analysis (Figure 8A). From the ten transformants, four were found to contain the PCR product at 900 bp (Figure 8B). Subsequently all four were minipreped to obtain plasmid DNA and sequenced. Three of these transformants (1, 2 and 6, Figure 8B) did not possess mutations and thus the PCR product was ligated in the correct orientation. Transformant 5 however contained a small number of mutations within the gene itself. Transformant 2 was ultimately chosen for further study, and the recombinant plasmid was transformed[29] into *E.coli* BL21-Gold (DE3).

3.4.2 Over-expression and purification of MA37 fluorinase

Studies were carried out to determine if the fluorinase enzyme was over-expressed as a soluble protein. Initially the protein was expressed in mg/ml quantities, however over 80% was obtained as insoluble protein. Two key factors that affect protein solubility are temperature and inducer concentration. An over-expression study was carried out using cultures of BL21-Gold (DE3) transformed with pEHISTEV-*flA1.* These cultures were induced with either 1mM or 0.2 mM IPTG concentration and grown at either 16 ºC or 30 ºC for 24 h shaking at 180 rpm. Mature cultures were analysed *via* SDS-PAGE for soluble and insoluble protein (Figure 9). High FlA1 (MA37 fluorinase) protein expression was observed with 1 mM IPTG induced cultures grown at 30 ºC when compared to those grown at 16 ºC.

However most of the expressed protein was still observed *via* SDS-PAGE in the insoluble fraction. A moderate amount of soluble protein was also observed, however a second

unwanted *E.coli* protein (28 kDa) was also detected at significant levels. Therefore the IPTG concentration was further lowered and an over-expression with 0.03 mM IPTG at 16 °C was performed. As before an SDS-PAGE was run on the soluble and insoluble fractions (Figure 10). These conditions also produced some insoluble MA37 fluorinase, however sufficient soluble protein was observed and the expression of the unwanted *E.coli* protein was abolished.

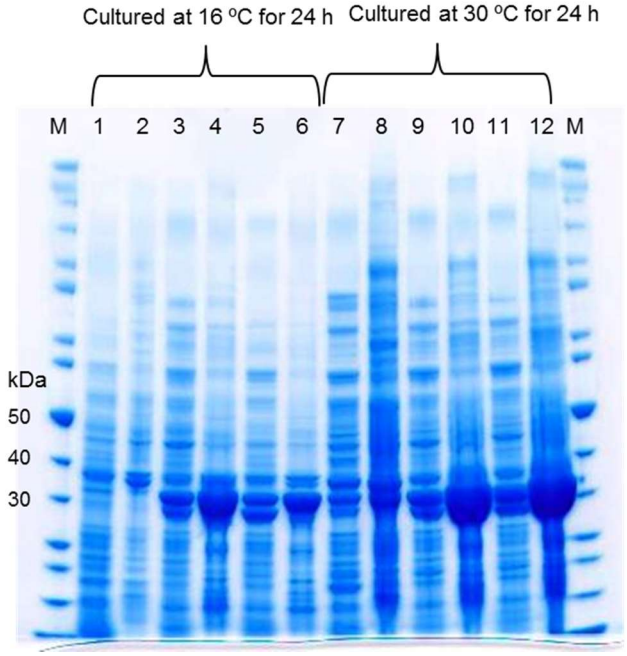


Figure 9: SDS-PAGE of samples from the over-expression study of *Streptomyces* sp. MA37 fluorinase in BL21-Gold (DE3) *E.coli* cells with varying temperatures and IPTG concentrations. M: un-stained PageRuler protein ladder.

Cultured at 16 ºC for 24 h 1: no IPTG; soluble fraction, 2: no IPTG; insoluble fraction, 3: 1 mM IPTG; soluble fraction, 4: 1 mM IPTG; insoluble fraction, 5: 0.2 mM IPTG; soluble fraction, 6: 0.2 mM IPTG; insoluble fraction.

Cultured at 30 ºC for 24 h 7: no IPTG; soluble fraction, 8: no IPTG; insoluble fraction, 9: 1 mM IPTG; soluble fraction, 10: 1 mM IPTG; insoluble fraction, 11: 0.2 mM IPTG; soluble fraction, 12: 0.2 mM IPTG; insoluble fraction.

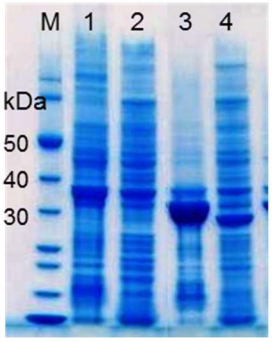


Figure 10: SDS-PAGE of the samples of over-expressed BL21-Gold (DE3) *E.coli* cells containing pEHISTEV-*flA1* induced with 0.03 mM IPTG and grown at 16 ºC.
M: un-stained PageRuler protein ladder (Thermo scientific), 1: no IPTG; insoluble fraction, 2: no IPTG; soluble fraction, 3: 0.03 mM IPTG; insoluble fraction, 4: 0.03 mM IPTG; soluble fraction.

Further screening was carried out to optimise the soluble protein concentration of the MA37 fluorinase by altering the buffers used during cell lysis. Many methods for solubilising protein are available; expressing in specialised *E.coli* strains,[30] creating fusion protein constructs (such as using maltose binding protein[31,32] and Small Ubiquitin-like Modifiers[33]) and including molecular chaperones[34]. However each additional method requires significant optimisation without guaranteed success. Studies have shown that some insoluble protein can be solubilised during the purification process by utilising specific buffer conditions[35] and by addition of co-solvents such as detergents, salts, buffers and reducing agents during the cell lysis.[36]

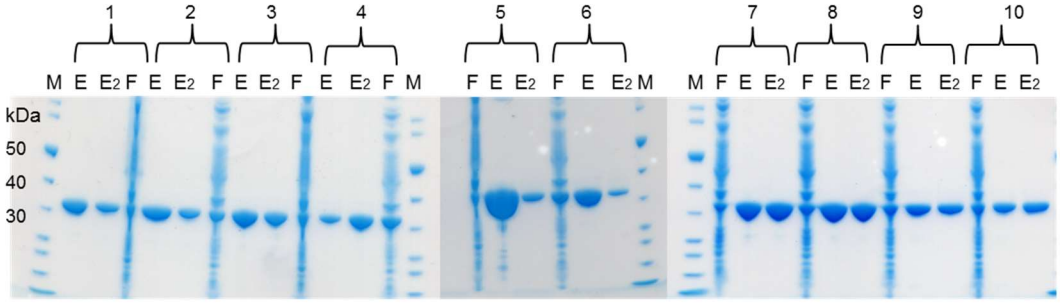
Additional screening was therefore performed using two buffers containing varying salt concentrations and additives. A BioSprint 96® was used to perform the screening. This is an automated system capable of performing rapid small scale affinity purifications independently. *E.coli* cultures expressing MA37 fluorinase were first lysed and processed through the BioSprint 96® system using ten different buffers, each supplemented with different salt and additive combinations, composed of either 20 mM Tris-HCl or 20 mM Bis-Tris at pH 8.0 or 6.8 along with additives such as glycine, 2-mercaptoethanol (β-ME), NP40 (Nonidet P40) and Triton X-100.[35]

The pH and ionic strength of the solution both have a strong influence on protein solubility, as such specific buffers and salts are commonly added to sample preparation solutions. Salts such as $MgSO_4$, NaCl and KCl stabilise the native intramolecular protein interactions, and thus outcompete the protein-protein intermolecular interactions that lead to aggregation.[37–39] Small amino acids such as glycine and L-arginine are aggregation

inhibitors stabilising the unfolded proteins by interacting favourably with majority of amino acid side chains on the protein surface.[39] Reducing agents such as β-ME assist in maintaining the protein in a reduced state by preventing oxidation of cysteine residues, which may ultimately cause protein aggregation through intermolecular disulphide bridges.[35]

Non-ionic detergents such as NP40 and Triton X-100 have uncharged and hydrophilic head groups. They are considered mild surfactants as they break protein-lipid and lipid-lipid associations, but not protein-protein interactions, and most do not denature proteins. Therefore, through use of these detergents, proteins are solubilized and isolated in their native and active forms, retaining important intermolecular interactions.[40]

Ten buffers supplemented with various co-solvents (discussed above), were analysed as potential lysis buffers, in order to optimise the solubility of the MA37 fluorinase. Soluble fluorinase was obtained when cells were lysed and purified in 20 mM Bis-Tris pH 6.8, 150 mM NaCl and 0.1 M Glycine (Figure 11; buffer #5). Therefore these conditions were used in subsequent large scale affinity purifications.



F-flow-through ; E- elution 1 ; E2- elution 2

1. 20mM Bis-Tris pH6.8 150mM NaCl
2. 20mMTris-HCl pH8.0 150mMNaCl
3. 20mM BisTris pH 6.8 + 150mM NaCl + βME 2.5mM
4. 20mM Tris pH 8.0+ 150mM NaCl + βME 2.5mM
5. **20mM Bis-Tris pH6.8 150mM NaCl + 0.1M Glycine**
6. 20mM Tris-HCl pH8.0 150mM NaCl + 0.1M Glycine
7. 20mMBis-TrispH6.8 150mM NaCl + 0.1%(v/v) NP40
8. 20mM Tris-HCl pH8.0 150mM NaCl + 0.1%(v/v) Triton X100
9. 20mM Bis-Tris pH6.8 150mM NaCl + 0.1%(v/v) Triton X100
10. 20mM Tris-HCl pH8.0 150mM NaCl + 0.1%(v/v) NP40

Figure 11: SDS-PAGE of fractions obtained from Biosprint 96® (small scale affinity purifications), investigating the optimal buffer-additive combination which solubilise the MA37 fluorinase. The ten buffer conditions are listed, from which #5 yielded the most soluble protein (20 mM Bis-Tris pH 6.8, 150 mM NaCl, 0.1 M Glycine buffer).

Streptomyces sp. MA-37 fluorinase was purified using the protocol reported in the Experimental 3.6.3. The purified MA-37 fluorinase was relatively unstable compared to the

more robust *S. cattleya* fluorinase, and was found to be sensitive to short term interaction with high imidazole **40** and high salt concentrations, as well as exposure to mechanical agitation, causing the protein to precipitate over time. Therefore, it was necessary for the protein to be desalted immediately following affinity purification, as removing the imidazole **40** through dialysis was too slow a process and resulted in aggregated protein. Through trial and error, the optimised protocol (Experimental 3.6.3) was utilised to obtain pure *Streptomyces* sp. MA37 fluorinase for assays and crystallization studies.

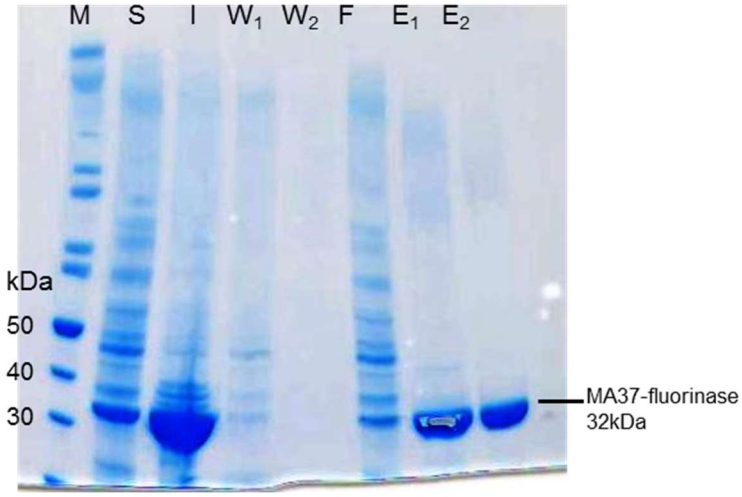


Figure 12: SDS-PAGE of fractions obtained from Ni-NTA column for MA37 fluorinase. M- marker, S- soluble fraction, I- insoluble fraction, W_1- wash 1, W_2- wash 2, F- flow through, E_1- elution 1, E_2- elution 2

E.coli cells transformed [29] with pEHISTEV-*flA1* and were grown in 1 L cultures, harvested and lysed in Bis-Tris (20 mM) pH 6.8 buffer supplemented with glycine to obtain the soluble lysate. The lysate was applied to a Ni His-trap column which was first equilibrated with wash buffer containing Bis-Tris (20 mM) pH 6.8, supplemented with glycine (0.1M) and imidazole **40** (40 mM). The column was then flushed using wash buffer to remove any proteins bound by non-specific interactions. The protein was eluted using Bis-Tris (20 mM) pH 6.8 buffer supplemented with glycine (0.1M) and imidazole **40** (400 mM) (Figure 12). This process was followed by an immediate desalting column using wash buffer to reduce the final imidazole **40** concentration.

The *S. cattleya* fluorinase purifies with adenosine **47** bound to its active site. A successful strategy is to remove the **47** by treatment with adenosine deaminase.[41] Similarly MA37 fluorinase was also found to contain bound **47** following purification, as shown by HPLC analysis carried out following heat denaturing of purified MA37 fluorinase. Therefore the desalted MA37 fluorinase was incubated with adenosine deaminase which converts **47** to

inosine, which does not bind to fluorinase and can be readily removed *via* dialysis, thus generating apo- MA37 fluorinase.[15] HPLC analysis was carried out during (12 h) and post (22 h) adenosine deaminase treatment to assess the progress of adenosine **47** removal (Figure 13). Post treatment, a Ni His-trap column was used to separate the adenosine deaminase (non His_6-tagged) from treated MA37 fluorinase (His_6-tagged). The treated fluorinase was eluted as before and desalted immediately.

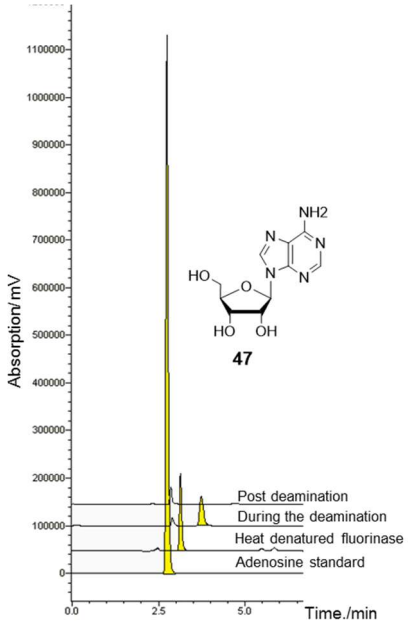


Figure 13: HPLC analysis of heat denatured MA37 fluorinase showing the presence of adenosine **47** t_R= 2.8 min. A standard sample of **47** was run to confirm the peak.

His_6-tags are significantly smaller than most affinity tags and are thought to cause minimal disturbance to protein structure and function. However some instances are reported where the His_6-tag may affect the protein structure, activity and binding affinity and[42] can also be a disadvantage during protein crystallisation.[43]

The pEHSITEV-flA1 contains a TEV recognition site located between the His_6-tag and the main polypeptide sequence, which is identified by the TEV protease. Therefore the deaminase-treated MA37 fluorinase was incubated with a TEV protease and purified by another Ni His-trap column to perform a 'subtractive' purification, where cleaved MA37 fluorinase is collected with the flow through. Following this the deaminase-treated, His_6-tag-cleaved, MA37 fluorinase was concentrated and injected into a gel filtration column as a refining step.

An equilibrated size exclusion column (SEC) was loaded with the concentrated protein and eluted with HEPES buffer (Figure 14A). The MA37 fluorinase was then dialysed into phosphate buffer (20 mM) using Vivaspin concentrators (GE Healthcare). Fractions obtained from the SEC column were pooled and analysed by SDS-PAGE gel to confirm the purity of the protein. The SDS-PAGE revealed a single band at ~32 kDa (Figure 14B), the gel band was collected and analysed by MALDI-TOF to confirm the identity of the protein (Figure 15). The trypsin digested protein gel band was analysed for distinct cleavage patterns dictated by the amino acid sequence of the protein.[44] The sequence coverage for *Streptomyces* sp. MA-37 fluorinase was over 66% confirming the identity of the protein (Figure 13).

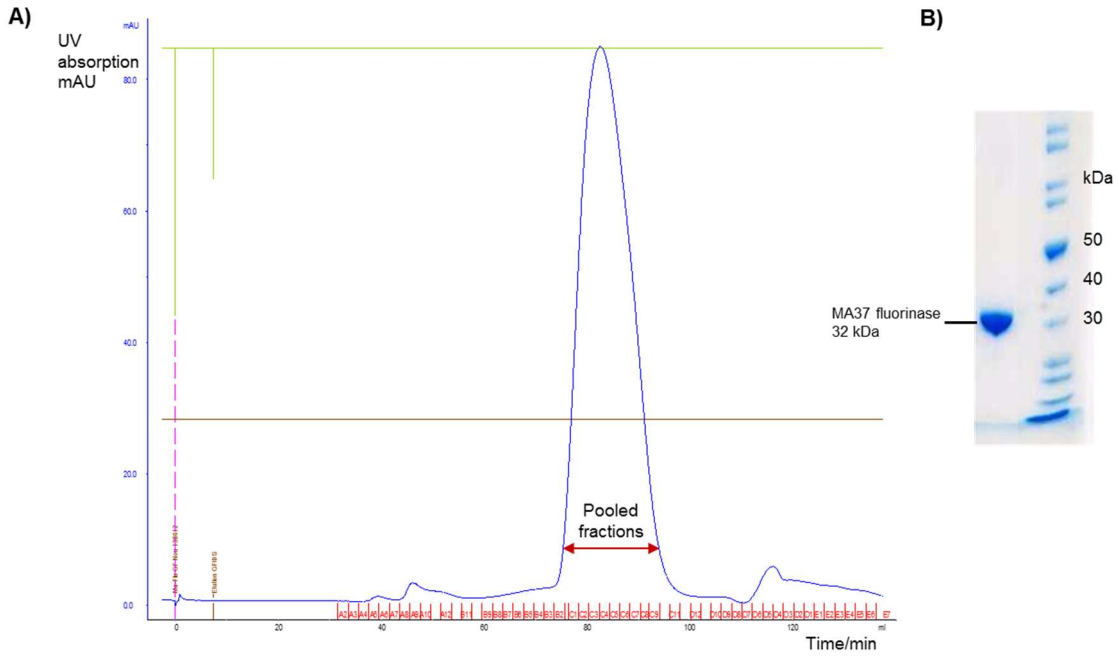


Figure 14: A) Size exclusion column chromatography of MA37 fluorinase showing the UV absorption in blue, concentration of the buffer in brown, flow rate in green and fractions in red. **B)** SDS-PAGE of the pooled fractions of MA37 fluorinase following size exclusion column chromatography.

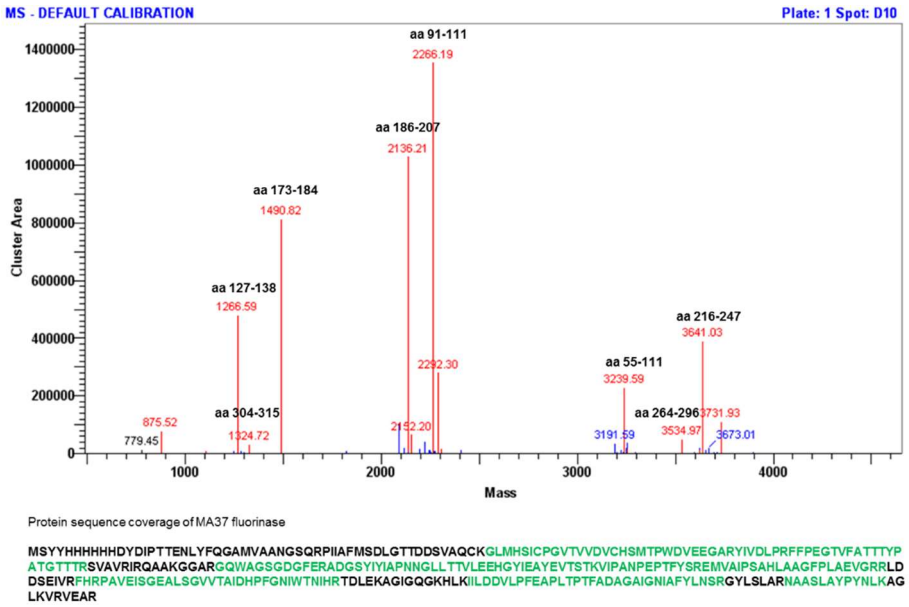


Figure 15: MALDI-TOF mass spectrum of trypisin digested SDS-PAGE gel band of MA37 fluorinase. Red peaks depict the peptide ions which matched the amino acid sequence of His_6-tagged MA37 fluorinase which were over 66% for MA37 fluorinase.

3.4.3 Crystallization of MA37 fluorinase

Purified MA37 fluorinase was used in crystallisation trials as reported in the Experimental section 3.6.4. The success of protein crystallisation depends upon many variables such as; pH, ionic strength, temperature, protein concentration, additives or ligands and salts. Therefore it is difficult to predict the best crystallisation conditions prior to optimisation studies for a given protein.

Initial crystallisation screens for the MA37 fluorinase were set up using pre-made commercial screens JSCG+ and PEGs II suite and incubated at 20 ºC and 4 ºC for 3-5 days. Screens incubated at 4 ºC produced mostly spherulites and contained precipitated protein. The screens incubated at 20 ºC all contained PEG in which thin needle like crystals were observed after 5 days. Therefore, optimisation screens were prepared for all four conditions with varying concentrations of PEG and additives (Experimental 3.6.4; Table 2). The best crystals (most suitable for diffraction) were obtained in 7% *w*/v PEG 6000, 0.5 M NaCl with 6.6 mg/ml of apo-MA37 fluorinase, incubated at 20 ºC. Octahedron shaped crystals appeared after ~3 days which grew in size to approximately 0.03- 0.05 mm. They were harvested after 1.5- 2 weeks (Figure 16). The crystals were cryoprotected and shot in an in-house X-ray diffractometer at 100K. The data was used to determine the structure up to 2.5 Å resolution with an R value of 0.21 (a measure of the quality of the atomic model

obtained from the crystallographic data). The crystal structure were resolved by molecular replacement from previous fluorinase structures by Greg Mann (Professor J. Naismith group) at the University of St Andrews (X-ray diffraction data is presented in Appendix 1). The crystal structure of the MA37 fluorinase with bound adenosine **47** is shown in Figure 17. The X-ray diffraction data is presented in Appendix 1.

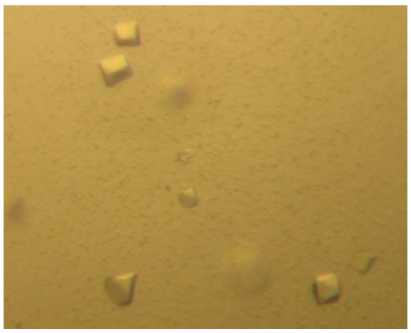


Figure 16: Octahedron shaped MA37 fluorinase crystals.

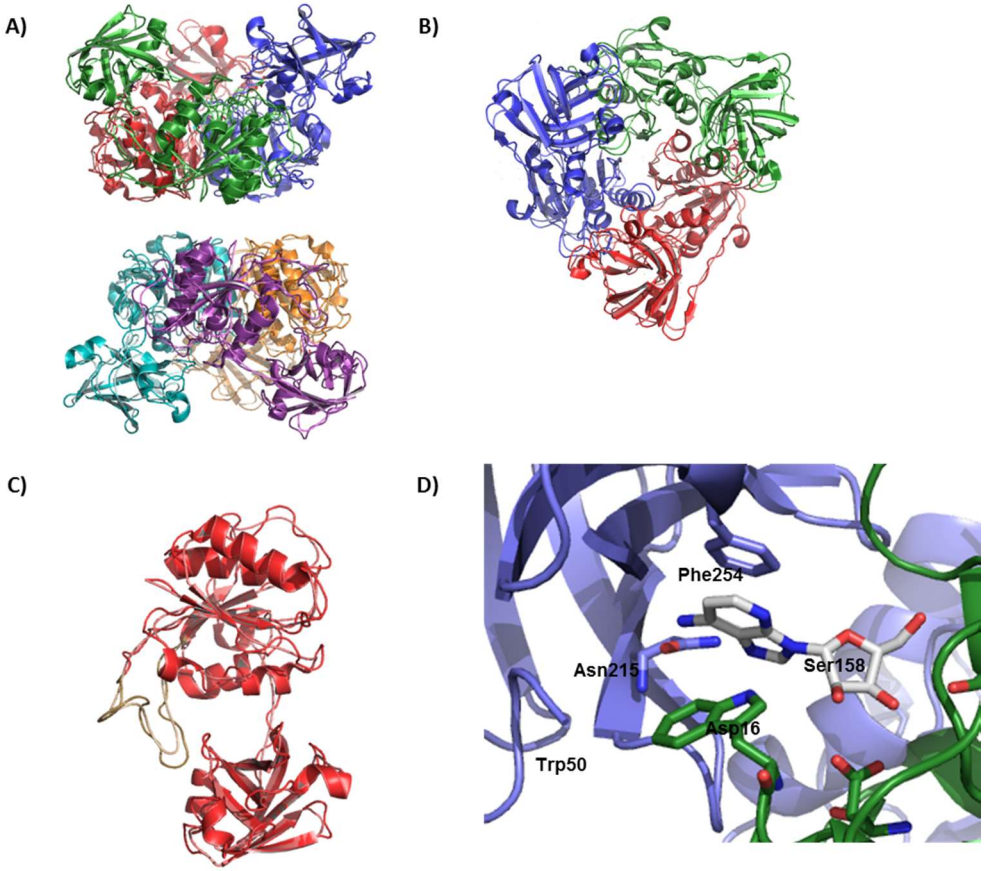


Figure 17: A) Hexameric arrangement of *Streptomyces.* sp. MA37 fluorinase superimposed on *S. cattleya* fluorinase **B)** Superimposition of *Streptomyces* sp. MA37 fluorinase on *S. cattleya* fluorinase. **C)** Superimposition of monomeric unit of *Streptomyces* sp. MA37 fluorinase to *S. cattleya* fluorinase. (*Streptomyces* sp. MA37 fluorinase is represented in dark colours (blue, green, red, orange, purple and cyan) and *S. cattleya*

fluorinase in corresponding light colours) **D)** The active site of the *Streptomyces* sp. MA37 fluorinase with bound adenosine **47**. Residues Asp16, Trp50 , Asn 215 and Phe254 play a key role in substrate recognition, while Ser158 is reported to bind fluoride, all of which are highly conserved.[12] (The crystal structure was resolved by molecular replacement by Greg Mann (Professor J. Naismith group) at the University of St Andrews).

Despite deaminase treatment of the MA37 fluorinase, the solved crystal structure had adenosine **47** within active site of the enzyme. This clearly suggests that the removal of bound **47** using adenosine deaminase was an inefficient process.[15] The adenosine-MA37 fluorinase co-crystal structure, was then compared to that of the adenosine **47** bound *S. cattleya* fluorinase (Figure 17D). The two structures are almost identical at the monomeric level. FLA1 was also found to be arranged into a hexameric structure (Figure 17A) formed by two dimers (Figure 17B), where each dimer unit is a trimer of three monomeric units. The residues that form the active site of the *Streptomyces* sp. MA37 fluorinase are shared between two monomers and their arrangement is the same as that in *S. cattleya* fluorinase (Figure 17D). Only a small number of amino acids are dissimilar, and as these are situated remote from the active site they and do not cause any major structural changes to FLA. The conserved 21 amino acid loop is also present in FLA1, albeit in a more disordered manner than that found in FLA (Figure 17C).

3.4.4 Enzymatic assays- evaluation of kinetic parameters

The MA37 fluorinase was assayed, varying the concentrations of SAM **22**, and the kinetic parameters (K_m, V_{max}) were evaluated in order to perform a comparative study to the *S. cattleya* fluorinase.[45,46]

A standard curve for 5′-FDA **23** was prepared by measuring the relative absorption of known concentrations of synthetic **23** (Figure 18). The data obtained was plotted as UV absorbance against the concentration of **23**. This was used to quantify the concentration of **23** produced in each HPLC analysed enzyme assay.

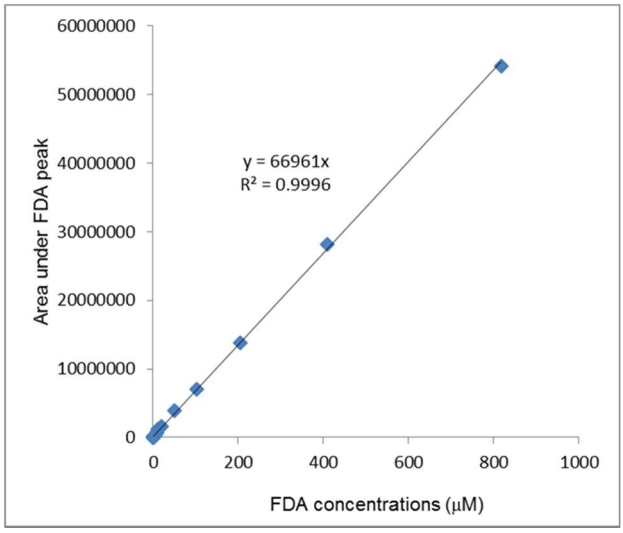


Figure 18: A standard curve obtained for 5′-FDA **23** concentrations vs UV.

Enzyme assays were set up using 0.5 mg/ml of purified *Streptomyces* sp. MA-37 fluorinase, KF (200 mM) and varying concentrations of SAM **22** (2 µM to 1000 µM) in phosphate buffer at pH 7.8, as reported in Experimental 3.6.5. Periodically, small aliquots of each assay were removed and analysed by HPLC. Figure 19A shows one such HPLC reaction profile where the MA-37 fluorinase was incubated with 40 µM SAM **22** and an HPLC profile analysed over 30 min. Multiple reactions were performed for each reaction to correct for any pipetting or handling errors and the data was also corrected for any dilutions performed.

The data obtained from multiple runs were averaged, and the initial rates of reaction determined by plotting 5′-FDA **23** concentration against time for varying concentrations of SAM **22** in a linear regression plot (using GraphPad Prism software, Figure 19B). The gradient of each line in Figure 19B represents the rate of each reaction at a given concentration of **22**. A general increase in **23** production was observed over time as the concentration of **22** was increased up to 500 µM, after which the maximum rate of reaction (V_{max}) was achieved, most likely due to substrate saturation.

The rate of reaction of each assay was plotted against corresponding **22** concentrations and fitted to non-linear regression model for enzyme kinetics using GraphPad prism. The best fit was obtained for Michaelis-Menten kinetic profile (Figure 19C) which revealed the same kinetic trend as for *S. cattleya* fluorinase (Figure 20).

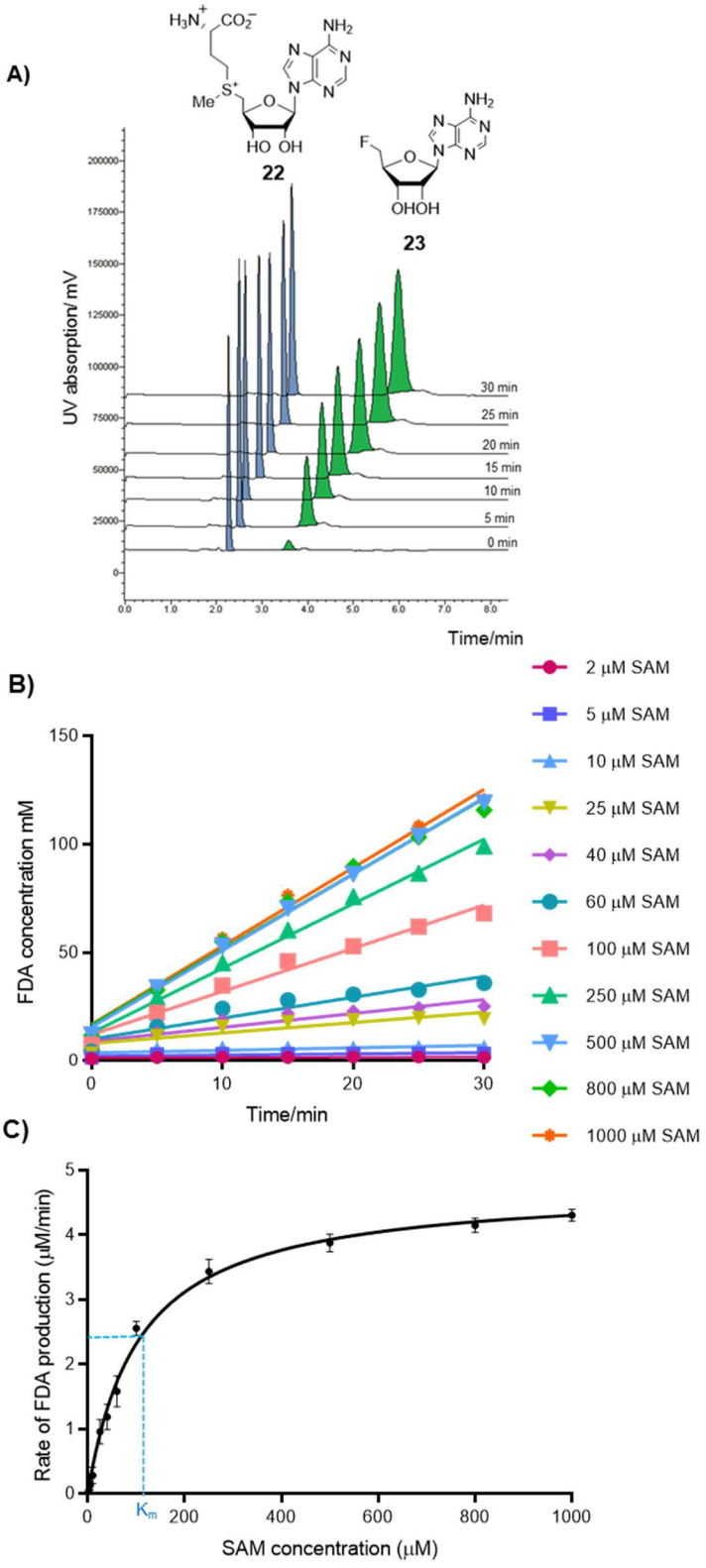


Figure 19: **A)** HPLC time course of a single set of assays where the MA37 fluorinase was incubated with 40 µM SAM **22** and assayed over 30min. **B)** Linear regression model fitting 5′-FDA **23** concentration against time for varying concentrations of **22** **C)** Non-linear regression model fitting initial rate of reaction of each reaction against corresponding **22** concentrations.

The K_m of MA-37 fluorinase for SAM was calculated from data shown in Figure 19C, (the substrate concentration at half maximum velocity). The maximum reaction rate for MA-37 fluorinase was reported as 4.75 ±0.01 μM/min and the approximate K_m for SAM **22** was calculated to be 105.5 μM ±7.68 μM. Since the **22** (Sigma) used in all enzyme reactions was 80% pure, the approximate K_m was corrected to 84.4 μM ±6.12μM. The h value (the hill slope that accounts for the steepness of the curve) was equal to 1 (1.004 ± 0.08739) which is consistent with non co-operative binding.

The same kinetic analysis was performed for the *S. cattleya* fluorinase and the data was fitted also using GraphPad Prism software to calculate K_m (Figure 20).

The maximum reaction rate for *S. cattleya* fluorinase was observed to be 1.147 ± 0.1 μM/min and the approximate K_m for **22** was calculated to be 29.2 μM ± 2.41 μM, in line with previously calculated values for this fluorinase.[45,46]

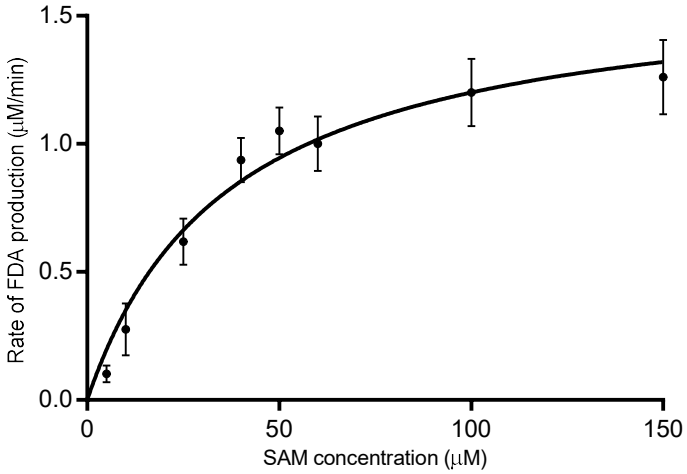


Figure 20: Non-linear regression model fitting of the initial reaction rate against corresponding SAM **22** concentrations from which K_m and k_{cat} of *S. cattleya* fluorinase was calculated.

S. cattleya fluorinase was reported to possess a k_{cat} (turnover number) of approximately 0.07 min^{-1}, confirming the fluorinase to be a slow enzyme.[45] The k_{cat} defines the number of substrate molecules that each catalytic site of the enzyme converts per unit of time, and is usually expressed in s^{-1}. The turnover number for *Streptomyces* sp. MA-37 fluorinase was calculated to be around 0.292 min^{-1}, three times faster than the *S. cattleya* fluorinase (0.083 min^{-1}). However the *Streptomyces* sp. MA-37 fluorinase was also observed to have a higher K_m (84.4μM) compared to K_m of *S. cattleya* (29.2 μM). Therefore, the *S. cattleya* fluorinase has a higher affinity for SAM **22** compared to the *Streptomyces* sp. MA-37 fluorinase. The specificity constant of an enzyme is defined by the ratio of the kinetic constants k_{cat}/K_m. It is

useful as a guide to compare the relative efficiencies of enzymes. *Streptomyces cattleya* fluorinase was calculated to have a specificity constant of 2.84 mM^{-1} min^{-1} less than the *Streptomyces* sp. MA-37 fluorinase at 3.6 mM^{-1} min^{-1}, suggesting that the MA37 fluorinase is a little more efficient.

3.4.5 Comparison of all fluorinases

The two *S. cattleya* fluorinase gene homologs identified from *Nocardia brasiliensis* (*flA2*) and *Actinoplanes* sp. N902-109 (*flA3*) were also successfully over-expressed in *E.coli* using codon optimised synthetic genes. Both flA2 and flA3 gene products were purified and found to be functional fluorinases capable of catalysing the conversion of **22** and fluoride to **23**. The kinetic analysis of these additional fluorinase enzymes was conducted in similar manner to that used for the *Streptomyces* sp. MA-37 fluorinase. The protein purification and kinetic analysis of *Nocardia brasiliensis* and *Actinoplanes* sp. N902-109 fluorinases were carried out by Dr Long Ma at the University of St Andrews. [11]

The kinetic data of the three newly identified fluorinases (flA1, flA2, flA3) and the *S. cattleya* fluorinase (flA) were compared. The *S. cattleya* fluorinase is the least efficient. The *N. brasiliensis* and *Actinoplanes* sp. N902-109 fluorinases were the most efficient, each having specificity constants of around 4.4 mM^{-1} min^{-1}. The affinity (K_m) of *N. brasiliensis* fluorinase for SAM **22** was similar to that observed with *S. cattleya* fluorinase at around 27 μM (Table 1).

Table 1: Summary of the kinetic data for all four *flA* homologs.[11]

Organism fluorinase was isolated from	Fluorinase gene homolog	V_{max} ($\mu M^{-1}min$)	SAM K_m (μM)	k_{cat} or Turnover number (min^{-1})	Catalytic efficiency k_{cat}/K_m ($mM^{-1}min^{-1}$)
Streptomyces cattleya	*flA*	1.65 ± 0.14	29.6 ± 6.57	0.08	2.70
Streptomyces sp. MA37	*flA1*	4.76 ± 0.01	84.4 ± 6.12	0.29	3.60
Nocardia brasiliensis	*flA2*	1.70 ± 0.07	27.8 ± 4.23	0.12	4.40
Actinoplanes sp. N902-109	*flA3*	2.83 ± 0.14	45.8 ± 7.91	0.20	4.44

3.5 Conclusions

After almost a decade since the characterisation of the *S. cattleya* fluorinase (FlA)[12,17], a new bacterium *Streptomyces* sp. MA37 was identified from the soil below a Bark Cloth tree in Ghana and the genome successfully sequenced. The genome of *Streptomyces* sp. MA37 revealed a homologue of the *flA* gene with 87% homology to the original fluorinase gene. This gene (*flA1*) was cloned in *E.coli*, over-expressed and the enzyme purified for assay analysis and crystallization studies. The crystal structure of *Streptomyces* sp. MA37 fluorinase was found to be almost identical to that of the *S. cattleya* fluorinase. The k_{cat}/K_m of the *Streptomyces* sp. MA37 fluorinase was moderately more efficient than that of the *S. cattleya* fluorinase. Furthermore, three additional fluorinase homologs have since been discovered and their individual kinetics were assessed by discontinuous assay using the HPLC.

3.6 Experimental

3.6.1 Genomic DNA extraction from *Streptomyces* sp. MA-37

Streptomyces sp. MA-37 was cultured following the protocol provided by the International Streptomyces Project (ISP) 2[47] (provided by Dr Hai Deng, University of Aberdeen). The mycelium was obtained from the cultures and was flash frozen in liquid nitrogen and finely ground using pestle and mortar. The ground mycelium was re-suspended in TE25S buffer (Tris-base 25 mM, EDTA disodium salt 25 mM and sucrose 10% and adjusted to pH 7.0; 5 ml) and lysozyme was added to a final concentration of 2 mg/ml. The suspension was incubated for 40 min at 37 ºC. After which Proteinase K from *Aspergillus melleus* (Sigma, 2U) and 10% sodium dodecylsulfate was added and incubated at 55 ºC for 1 h. To this preparation, 5 M NaCl (1ml) and phenol/chloroform (1:1 v/v; 5 ml) was added and mixed by inversion for 30 min. The resulting mixture was centrifuged (5000 rpm; 5 min) and the supernatant transferred to a clean eppendorf tube. The DNA was precipitated and dried using standard isopropanol precipitation[48] before re-dissolving in water (0.3 ml) and the resulting genomic DNA of *Streptomyces* sp. MA-37 was stored at -20 ºC until further use. The purity of the gross DNA was analysed by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.).

3.6.2 Cloning of *flA1* gene into pEHISTEV plasmid

The *Streptomyces* sp. MA-37 *flA*1 gene was amplified by PCR using previously purified genomic DNA as the template. Primers pMAF-F (5′- AATCCCATGGTGGCTGCAAACGGC -3′) and pMAF-R (5′-CACAAGCTTTCAGCGCGCTTCCAC-3′) were designed to match the

template DNA incorporating restriction sites NcoI and HindIII (restriction sites are underlined). The primers were synthesised by Eurogentec, Belgium. The PCR program included 35 cycles of 1 min at 95 °C (denaturation), 1 min 30 sec at 60 °C (annealing) and 1 min at 72 °C (elongation) and a final extension for 2 min at 68 °C. The amplified PCR product and empty pEHISTEV vector[28] was digested with NcoI and HindIII (Promega) in a double restriction digest in the presence of Buffer C at 37 °C for 2h. The digested PCR product and plasmid were analysed by DNA gel electrophoresis following which the gel bands were excised and gel purified using a Wizard® SV gel and PCR clean-up system (Promega). A ligation was set up using a Takara® DNA ligation kit, between the digested PCR product and digested pEHSITEV plasmid. The ligation was transformed in to *E.coli* DH5α (Thermo Fisher) cells and plated out on LB agar medium supplemented with kanamycin. Ten colonies were isolated from the transformation and analysed for the presence of *flA1* by colony PCR. The colony PCR was conducted using pMAF-F and pMAF-R and the PCR programme included 20 cycles of 15 s at 98 ºC, 2 s at 55 ºC and 20 s at 72 ºC. The colonies containing the *flA1* gene were cultured (10 ml) in LB supplemented with kanamycin. The pEHISTEV-*flA1* plasmid was extracted from the culture using QIAprep® Spin Miniprep, Qiagen kit (Qiagen). The plasmids were sent for sequencing (DNA Sequencing & Services, University of Dundee) using standard primers T7 and T7 Term. The sequencing results were analysed by Chromas (Technelysium, Australia) for mutations and orientation of the insert and the most appropriate transformant was chosen for protein expression. The pEHISTEV-*flA1* plasmid of the chosen transformant was then transformed into *E.coli* BL21 (DE3) cells (Agilent ®).

3.6.3 Over-expression and purification of *Streptomyces* sp. MA37 fluorinase

Transformed BL21 (DE3) Gold *E.coli* containing the pEHISTEV-*flA1* plasmid were cultured in 10 ml of LB media supplemented with kanamycin, at 37 ºC overnight. This culture was used to inoculate 1 L LB media supplemented with 100 µg/ml kanamycin and incubated at 37 ºC until the OD_{600} reached 0.6. Following which the cultures were then induced by isopropylthiogalactoside (IPTG, 0.03 mM) and incubated for 24 h at 16 ºC. The cultured *E.coli* was harvested by centrifugation (3200 rpm; 20 min) and small scale affinity purifications were performed using BioSprint 96® (Qiagen) to test 10 buffers (with varying pH and additives). The list of buffers used are reported in Figure 11. The fractions obtained from BioSprint 96® were analysed by SDS-PAGE and the optimal condition (20 mM Bis-Tris pH 6.8, 150 mM NaCl and 0.1 M Glycine) providing the most amount of soluble protein was chosen to perform large scale purification.

The harvested cell pellet was re-suspended in lysis buffer (20 mM Bis-Tris pH 6.8, 150 mM NaCl, 0.1 M Glycine ; 200 ml) supplemented with protease inhibitor cocktail (1x Complete Mini -EDTA free®, Roche) and Deoxyribonuclease I from bovine pancreas (Sigma, 0.1 mg/ml). The cells were lysed using a cell disruptor (Constant systems™; UK) at 30 kPSI and 4 ºC. Subsequently, the lysate was centrifuged (20000 rpm, 20 min, 4 ºC) and the supernatant filtered (0.45 µm syringe filter) following which the lysate was loaded onto a 10 ml pre-packed Ni-Sepharose® High performance His-Trap HP (GE Healthcare) column connected to a FPLC. Protein-bound Ni^{2+} beads were washed with 20 mM Bis-Tris pH 6.8, 150 mM NaCl, 0.1 M Glycine, 40 mM imidazole **40** pH 6.8 and 2.5 mM β-ME and eluted with 20 mM Bis-Tris pH 6.8, 150 mM NaCl, 0.1 M Glycine, 400 mM imidazole **40** pH 6.8, 2.5 mM β-ME.

The eluted protein was immediately loaded to a HiPrep 26/10 Desalting column (GE Healthcare) equilibrated with buffer 20 mM Bis-Tris pH 6.8, 150 mM NaCl and 0.1M Glycine. The purified MA37 fluorinase was incubated with adenosine deaminase (bovine recombinant, Sigma, 0.8 U) for 16 h- 24 h at 4 ºC, rocking at 30 rpm. The deamination process was analysed by HPLC until completion. An enzyme sample was removed after 16 h, heat denatured (95 ºC, 5 min) and centrifuged (13000 rpm, 10 min) and the soluble fraction was analysed by HPLC for the presence of adenosine **47**. HPLC column and method used is as described in 2.7.5. Once the deaminase treatment was complete, the resulting mixture of proteins, adenosine deaminase and MA37 fluorinase was loaded on to a pre-packed Ni-Sepharose® High performance His-Trap HP (GE Healthcare) column, washed and eluted as before, followed by a desalting column as mentioned previously.

Subsequently the de-aminase treated MA37 fluorinase was incubated with (tobacco etch virus (TEV) protease (recombinant protein expressed in *E.coli* and kindly provided by Professor J. Naismith research group) at 1:20 ratio (TEV : protein) for 16 h, rocking at 30 rpm at 4 ºC. The progress of the cleavage was monitored by SDS-PAGE gel as the molecular weight of the uncut protein (His_6-tagged MA37 fluorinase) is larger than cut protein. In order to separate the TEV protease, cleaved His_6-tag and His_6-tag removed MA37 fluorinase, the enzyme mixture was loaded to a pre-packed Ni-Sepharose® High performance His-Trap HP (GE Healthcare) column equilibrated with wash buffer. The flow-through from the column was collected as the His_6-tag removed MA37 fluorinase would not interact with the Ni-NTA beads, whereas the His_6-tagged TEV and cleaved His_6-tags will be bound to the Ni-NTA beads. The MA37 fluorinase obtained in the flow-through was concentrated to ~ 7 ml using a centrifugal filter unit (30000 Da, GE Healthcare) following which the protein was loaded on to a gel filtration column Superdex S-200 (HR 16/60; GE

Healthcare), connected to a FPLC (ÄKTA express, Amersham Biosciences) equilibrated with 10 mM HEPES pH 7.4,150 mM NaCl, 1 mM TCEP (tris (2-carboxyethyl) phosphine). The fractions containing MA37 fluorinase were pooled and concentrated using a centrifugal filter unit (30000 Da, GE
Healthcare). For crystallisation trials, the buffer was exchanged using centrifugal filter unit (30000 Da, GE Healthcare) to 20 mM phosphate pH 7.8 and 150 mM NaCl. The final protein concentration was measured *via* Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.) and aliquots were flash frozen and stored in -80 ºC until further use.

3.6.4 Crystallization of *Streptomyces* sp. MA37 fluorinase

Initial screening was carried out using commercial screens in a 96-well format (JSCG+; Molecular dimensions and PEGs II suite; Qiagen) set up with 6 mg/ml and 12 mg/ml of purified *Streptomyces* sp. MA-37 fluorinase using a sitting drop vapour diffusion technique at 20 ºC and 4 ºC. All crystal trials were set up with an Arts Robbins Gryphon at protein to precipitant ratios of 2:1 and 1:1. Thin needle like crystals were observed in 4 conditions (Table 2), optimisation screens were designed for each condition varying the PEG concentrations and the additive concentrations, keeping the protein concentration at 6mg/ml. Crystals were observed in several conditions of the optimisations however the best diffracting single crystals (octahedron shaped) were obtained with 7% *w*/v PEG 6000 and 0.5 M NaCl at 20 ºC. These crystals were grown to their maximum size over 1.5- 2 weeks and a single crystal was picked and flash-frozen in cryoprotectant containing precipitant supplemented with 30% glycerol. Data were collected in-house at 100 K on a Rigaku 007 HFM rotating anode X-ray generator with a Saturn 944 CCD detector. A dataset of 236 frames of 0.5° oscillation with a crystal-to-detector distance of 60 mm and 240 sec exposure per frame was collected. The data were indexed and integrated using *MOSFLM* and scaled using *SCALA.*

The structure was solved using PHASER molecular replacement, using the *S. cattleya* fluorinase. The structure was refined iteratively through manual building in COOT (Crystallographic Object-Oriented Toolkit), and automated refinement was carried out using PHENIX.

Table 2: Crystal forming conditions of the commercial screens, PEGs II suite (Qiagen) and JSCG+ (Molecular dimensions) which were the basis to design optimisation screens.

Commercial Screen ; Well in which crystals were observed	Contents of the well	Concentration range for optimisation
PEGs II ; D6	30% *w*/v PEG 4000, NaOAc 0.1 M pH 4.6, 0.1M $MgCl_2$	25- 36% *w*/v PEG 4000 and 0.04 - 0.18 M $MgCl_2$
PEGs II ; F7	25% *w*/v PEG 4000, 0.1 M NaOAc, 0.2 M Li_2SO_4, 0.1 M HEPES pH 7.5	20- 31% *w*/v PEG 4000 and 0.02 - 0.16 M NaOAc
PEGs II ; G5	12% *w*/v PEG 6000 and 2 M NaCl	6 -17% *w*/v PEG 6000 and 0.5-4 M NaCl
JSCG+ ; B2	20% *w*/v PEG 3350, 0.2 M NaSCN	15- 26% *w*/v PEG 3350 and 0.1 -0.35 M NaSCN

3.6.5 K_m determination of *Streptomyces* sp. MA37 fluorinase

Purified MA-37 fluorinase was used to set up a series of HPLC assays to determine the K_m for SAM **22**. 0.5 mg/ml of MA-37 fluorinase was incubated with 200 mM KF, 20 mM phosphate buffer (pH 7.8) and varying concentrations of SAM **22** (2 µM, 5 µM, 10 µM, 25 µM, 40 µM, 60 µM, 100 µM, 250 µM, 500 µM, 800 µM and 1000 µM), in a total reaction volume of 1 ml and incubated at 37 ºC. Aliquots from each reaction was removed periodically (100 µl every 5 min, 10 min, 15 min, 20 min, 25 min and 30 min), heat denatured (95 ºC, 5 min) and centrifuged (13000 rpm, 10 min) to precipitate the enzyme. The supernatant (50 µl) was diluted with water (100 µl) out of which 100 µl was injected to the Phenomenex® Kinetix 5µm C_{18} 100 Å (150 mm x 4.6 mm) column using mobile phase : water with 0.05% TFA (solvent A) to MeCN with 0.05% TFA (solvent B). The HPLC analysis was performed using a Shimadzu Prominence system with a two-step gradient 5% to 23% solvent B over 14 min, then from 23% to 95% solvent B over 16 min; then 95% solvent B for 3 min followed by a re-equilibration with 5% solvent B for 5 min. The flow rate was 1ml/min and the UV absorbance was measured at 260nm.

The standard curve for 5′-FDA **23** was prepared by running known concentrations of **23** using the same HPLC method.

3.7 Chapter 3 References

1 M. H. Medema, R. Kottmann, P. Yilmaz and E. Al, *Nat. Chem. Biol.*, 2015, **11**, 625–631.

2 A. Osbourn, *Trends Genet.*, 2010, **26**, 449–457.

3 M. H. Medema, R. Breitling and E. Takano, 2011, 485–502.

4 C. Zhao, P. Li, Z. Deng, H. Y. Ou, R. P. McGlinchey and D. O'Hagan, *Bioorg. Chem.*, 2012, **44**, 1–7.

5 A. M. Weeks, S. M. Coyle, M. Jinek, J. A. Doudna and M. C. Y. Chang, *Biochemistry*, 2010, **49**, 9269–9279.

6 M. C. Walker, M. Wen, A. M. Weeks and M. C. Y. Chang, *ACS Chem. Biol.*, 2012, **7**, 1576–1585.

7 F. Huang, S. F. Haydock, D. Spiteller, T. Mironenko, T. L. Li, D. O'Hagan, P. F. Leadlay and J. B. Spencer, *Chem. Biol.*, 2006, **13**, 475–484.

8 H. Deng, S. M. Cross, R. P. McGlinchey, J. T. G. Hamilton and D. O'Hagan, *Chem. Biol.*, 2008, **15**, 1268–1276.

9 V. Barbe, M. Bouzon, S. Mangenot, B. Badet, J. Poulain, B. Segurens, D. Vallenet, P. Marliere and J. Weissenbach, *J. Bacteriol.*, 2011, **193**, 5055–5056.

10 H. Y. Ou, P. Li, C. Zhao, D. O'Hagan and Z. Deng, *Nucleotide Seq. Submitt. to EMBL/GenBank/DDBJ databases*, 2011.

11 H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith and D. O'Hagan, *ChemBioChem*, 2014, **15**, 364–368.

12 C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer and J. H. Naismith, *Nature*, 2004, **427**, 561–565.

13 H. Deng and D. O'Hagan, *Curr. Opin. Chem. Biol.*, 2008, **12**, 582–592.

14 H. Deng, C. H. Botting, J. T. G. Hamilton, R. J. M. Russell and D. O'Hagan, *Angew. Chemie - Int. Ed.*, 2008, **47**, 5357–5361.

15 R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen and W. Fenical, *Angew. Chemie Int. Ed.*, 2003, **42**, 355–357.

16 A. S. Eustáquio, F. Pojer, J. P. Noel and B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69–74.

17 A. S. Eustáquio, D. O'Hagan and B. S. Moore, *J. Nat. Prod.*, 2010, **73**, 378–382.

18 C. D. Murphy, C. Schaffrath and D. O'Hagan, *Chemosphere*, 2003, **52**, 455–461.

19 P. C. Y. Woo, S. K. P. Lau, J. L. L. Teng, H. Tse and K.-Y. Yuen, *Clin. Microbiol. Infect.*, 2008, **14**, 908–934.

20 L. Ma, A. Bartholome, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng and D. O'Hagan, *Chem. Sci.*, 2015, **6**, 1414–1419.

21 R. K.J. and R. C.G., *Sherris Medical Microbiology*, McGraw Hill, 4th edn., 2004.

22 H. Nishida, T. Sakakibara, F. Aoki, T. Saito, K. Ichikawa, T. Inagaki, Y. Kojima, Y. Yamauchi, L. H. Huang and M. A. Guadliana, *J. Antibiot. (Tokyo).*, 1995, **48**, 657–666.

23 S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan and H. Deng, *Org. Biomol. Chem.*, 2014, **12**, 4828–4831.

24 X.-Q. Zhao, W.-J. Li, W.-C. Jiao, Y. Li, W.-J. Yuan, Y.-Q. Zhang, H.-P. Klenk, J.-W. Suh and F.-W. Bai, *Int. J. Syst. Evol. Microbiol.*, 2009, **59**, 2870–2874.

25 E. F. Vanin, *Annu. Rev. Genet.*, 1985, **19**, 253–272.

26 D. L. Stern, *Nat. Rev. Genet.*, 2013, **14**, 751–764.

27 R. B. Kapust and D. S. Waugh, *Protein Expr. Purif.*, 2000, **19**, 312–318.

28 H. Liu and J. H. Naismith, *Protein Expr. Purif.*, 2009, **63**, 102–111.

29 H. E. Bergmans, I. M. van Die and W. P. Hoekstra, *J. Bacteriol.*, 1981, **146**, 564–570.

30 B. Miroux and J. E. Walker, *J. Mol. Biol.*, 1996, **260**, 289–298.

31 M. Hammarström, N. Hellgren, S. van Den Berg, H. Berglund and T. Härd, *Protein Sci.*, 2002, **11**, 313–321.

32 R. B. Kapust and D. S. Waugh, *Protein Sci.*, 1999, **8**, 1668–1674.

33 D. Kuo, M. Nie and A. J. Courey, *Methods Mol. Biol.*, 2014, **1177**, 71–80.

34 C. Schlieker, B. Bukau and A. Mogk, *J. Biotechnol.*, 2002, **96**, 13–21.

35 S. E. Bondos and A. Bicknell, *Anal. Biochem.*, 2003, **316**, 223–231.

36 D. J. Leibly, T. N. Nguyen, L. T. Kao, S. N. Hewitt, L. K. Barrett and W. C. Van Voorhis, *PLoS One*, 2012, **7**, e52482.

37 A. Neagu, M. Neagu and A. Dér, *Biophys. J.*, 2001, **81**, 1285–1294.

38 T. Arakawa, D. Ejima, K. Tsumoto, N. Obeyama, Y. Tanaka, Y. Kita and S. N. Timasheff, *Biophys. Chem.*, 2007, **127**, 1–8.

39 H. Lu, H. Zhang, Q. Wang, H. Yuan, W. He, Z. Zhao and Y. Li, *Curr. Microbiol.*, 2001, **42**, 211–216.

40 J. M. Neugebauer, *Methods Enzymol.*, 1990, **182**, 239–253.

41 S. L. Cobb, H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan and D. a Robinson, *Org. Biomol. Chem.*, 2006, **4**, 1458–1460.

42 K. A. Majorek, M. L. Kuhn, M. Chruszcz, W. F. Anderson and W. Minor, *Protein Sci.*, 2014, **23**, 1359–1368.

43 M. Carson, D. H. Johnson, H. McDonald, C. Brouillette and L. J. DeLucas, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2007, **63**, 295–301.

44 U. Leurs, U. H. Mistarz and K. D. Rand, *Eur. J. Pharm. Biopharm.*, 2015, **93**, 95–109.

45 X. Zhu, D. Robinson, A. R. McEwan, D. O'Hagan and J. H. Naismith, *J. Am. Chem. Soc.*, 2007, **129**, 14597–14604.

46 C. Schaffrath, H. Deng and D. O'Hagan, *FEBS Lett.*, 2003, **547**, 111–114.

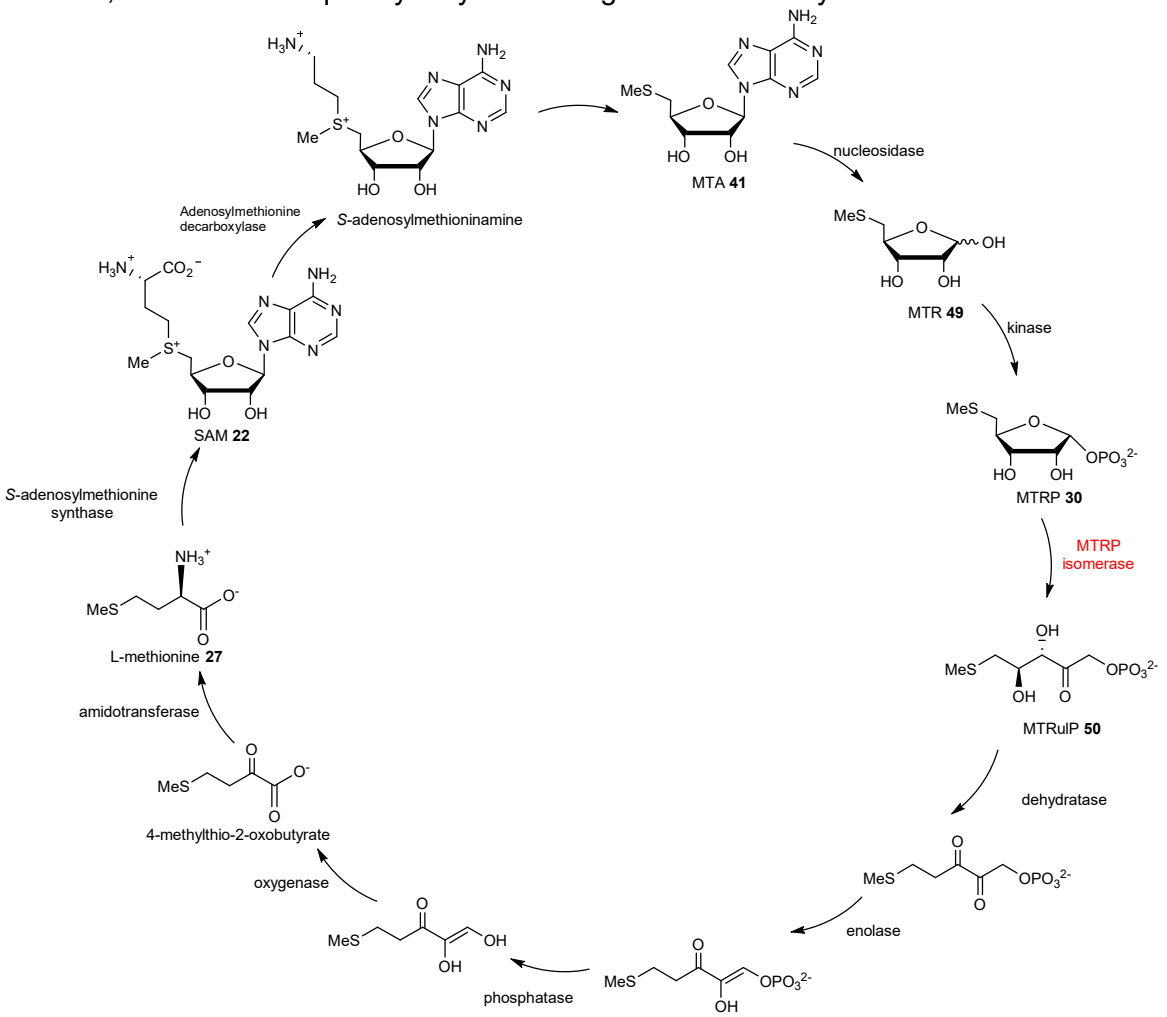
47 E. B. Shirling and D. Gottlieb, *Int. J. Syst. Bacteriol.*, 1966, **16**, 313–340.

48 J. F. Sambrook and D. W. Russell, Cold Spring Harbor Laboratory Press, New York, 4th edn., 2001, 26–28.

4. Crystallisation of 5-FDRPi Isomerase

4.1 Methionine salvage pathway

The methionine salvage pathway also known as the 5′-methylthioadenosine (MTA) **4** cycle is an essential cellular process in all organisms as it recycles sulphur-containing metabolites to regenerate L-methionine **27**. The amino acid L-methionine **27** is essential for numerous important cellular functions including: protein synthesis, the regulation of gene expression (*via* DNA and rRNA methylation), as well as the biosynthesis of secondary metabolites, hormones, phospholipids and polyamines. The cell utilises L-methionine **27** as both a proteinogenic amino acid and a component of the cofactor SAM **22**, which is formed by an enzyme catalysed condensation reaction between L-methionine **27** and ATP.[1] As the amount of **27** in a cell is limited it is essential for it to be recycled for cell survival. Most L-methionine **27** consumed in the generation of **22** is ultimately converted to the by-product, MTA **41**, which is subsequently recycled through the MTA **41** cycle.[2]

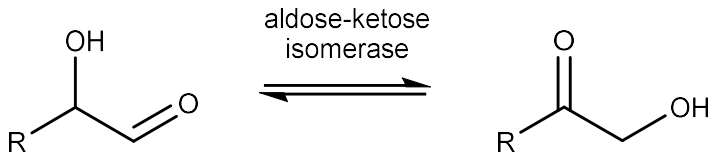


Scheme 1: The methionine salvage pathway of *B. subtilis* highlighting the MTRPi in red which has 35% sequence identity to *S. cattleya* 5-FDRPi.[7,36]

The enzymes involved in the methionine salvage pathway of model bacterial organisms, such as *E. coli* [1,3,4] and *Klebsiella pneumoniae*[5,6], have been widely studied and the pathway is well characterised. However, studies on other species of bacteria have revealed the use of alternative enzymes and different intermediates throughout the pathway, indicating a diversity of the methionine biosynthetic pathway dependent on the organism in which it resides.[7] In bacteria and plants, MTA **41** is first depurinated and phosphorylated to form 5-methylthioribose 1-phosphate (MTRP **30**) in two steps *via* 5-methylthioribose (MTR) **49**, utilising both a nucleosidase and a kinase (Scheme 1). In animals, **30** is formed in a single step reaction catalysed by a phosphorylase. Subsequently, **30** is transformed to 5-methylthioribulose 1-phosphate (MTRuP) **50** by an 5-deoxy-5-methylthioribose 1-phosphate isomerase (MTRPi) which catalyses the ring opening of **30**. The MTRulP **50** produced by MTRPi undergoes several enzyme catalysed reactions including dehydration and oxidative decarboxylation to generate 4-methylthio-2-oxobutyrate **51**, an immediate precursor for SAM **22** (Scheme 1).[1] The final step in the L-methionine **27** salvage pathway involves a transaminase which produces free methionine.

4.2 Aldose ketose isomerase

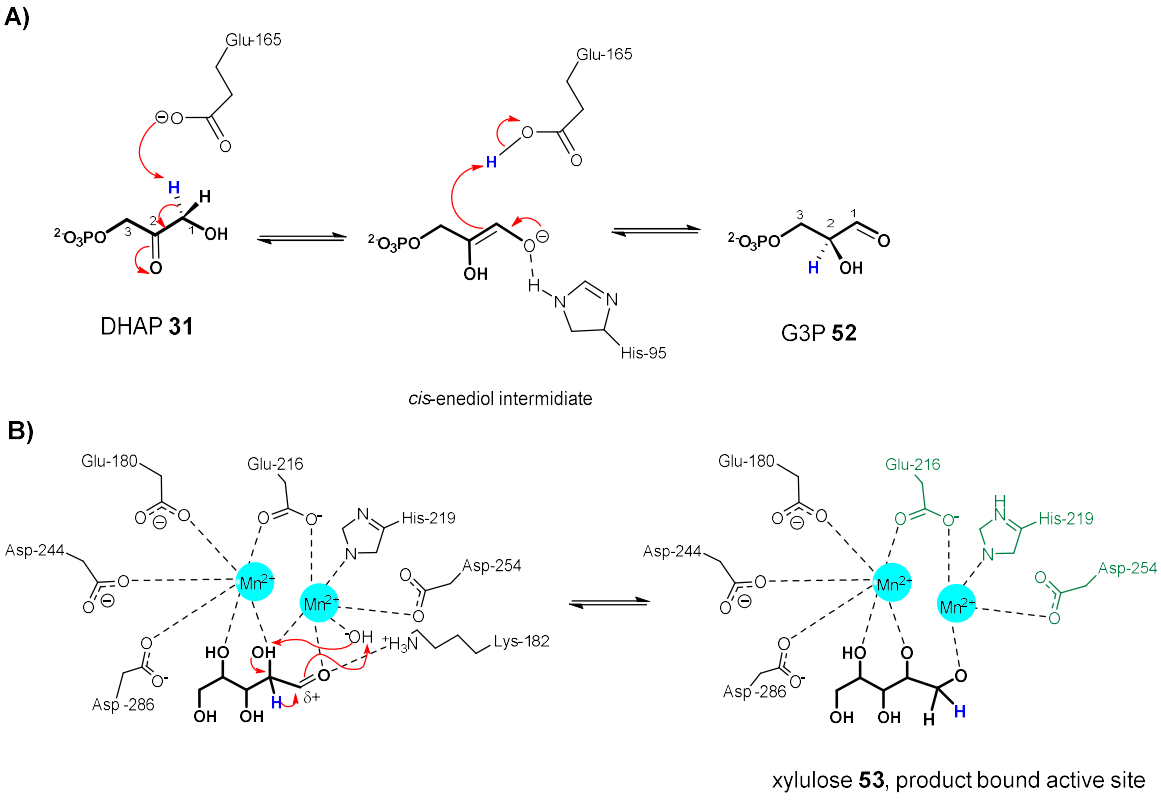
Aldose-ketose isomerases are crucial enzymes involved in the methionine salvage pathways present in all types of organisms, from unicellular bacteria to plants and animals. Such an isomerase catalyses the reversible isomerisation between α-hydroxyaldoses and α-hydroxyketoses (Scheme 2). In principle this class of enzymes can proceed *via* two mechanisms, the most common of which involves a *cis*-enediol intermediate as observed in the triose phosphate isomerase, which catalyse the interconversion of dihydroxyacetone phosphate (DHAP) **31** and glyceraldehyde 3-phosphate (G3P) **52** in the glycolysis pathway.[8]



Scheme 2: A simplified reaction scheme of the aldose-ketose isomerases.

Isotopic labelling studies investigating triose phosphate isomerase have demonstrated that the pro-*R* proton on C1 was removed by a glutamate residue (Glu165) acting as a base. As glutamate residues easily undergo proton exchange with the solvent, when the reaction is carried out in D_2O, deuterium incorporation is observed at the C1 position of DHAP **31** and

C2 of G3P **52** (Scheme 3A).[8] The enediol intermediate is stabilised by a neutral His-95 and a positively charged Lys-12 side chain residue. These *cis*-endiol mediating enzymes do not require any metal ions to catalyse the reaction. Phosphoglucose isomerase is also an aldose-ketose isomerase which proceeds *via* *cis*-enediol intermediate and catalyses the interconversion of glucose-6-phosphate and fructose-6-phosphate.[9,10]



Scheme 3 A): *Cis*-enediol mediated reaction mechanism for triose phosphate isomerase.[8] **B)** Active site arrangement of the intermediate and final product for *S. olivochromogenes* xylose isomerase showing the proposed mechanism.[9] In both cases the transferred proton is highlighted in dark blue and the catalytic triad involved in catalysis is highlighted in green.

Aldose-ketose isomerases are also found to proceed *via* metal dependent 1,2-hydride shift mechanism as observed for *S. olivochromogenes* xylose isomerase. These enzymes catalyse the interconversion of several aldose and ketose sugars. Primarily they catalyse the interconversion of xylose to xylulose **53** and glucose to fructose. Structural analysis of the xylose isomerase of *S. olivochromogenes* revealed it to be a tetrameric enzyme containing a bridged bimetallic system that acts as an electrophilic centre to promote hydride transfer within the active site. The two divalent metal ions secured by xylose coordination are thought to be involved in stabilising the transition state. During the hydride shift mediated interconversion, the hydrogen moves directly between C1 and C2 as a hydride, and as such cannot exchange with the solvent (Scheme 3B). Therefore when the

reaction is carried out in the presence of D_2O, no deuterium exchange is observed, unlike the *cis*-enediol mechanism.[11,12]

The enzyme, 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (5-FDRPi) identified in *S. cattleya* was also found to be an aldose-ketose isomerase.[13,14] The Basic Local Alignment Tool (BLAST) analysis of the protein sequence of 5-FDRPi revealed some sequence identity to the *Bacillus subtilis* MTRPi (35%) and the Ypr118w of *Saccharomyces cerevisiae* (PDB code: 1W2W; 26%) (Figure 1).[13] Therefore these MTRPi enzymes were used as reference structures for the study of 5-FDRPi, and this Chapter describes the over-expression and crystallisation of 5-FDRPi in an effort to provide an insight in to its mechanism.

4.3 Ribose 1-phosphate isomerases

4.3.1 Identification of 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (5-FDRPi) from *S. cattleya*

NMR studies conducted on cell free extracts of *S. cattleya* have confirmed the conversion of 5-FDRP **24** to 5-FDRulP **25** to be the third step in fluorometabolite production.[14] Subsequently, the *B. subtilis* MTRPi gene homolog in *S. coelicolor* (gene SCO3014) was isolated, cloned and over-expressed in *E. coli*, as the full genome sequence of *S. cattleya* was unavailable at the time, in 2008. Using the protein sequence of the *S. coelicolor* isomerase, the equivalent gene in *S. avermitilis* (gene SAV6658) was identified, as this was the most closely related organism to *S. cattleya*. Primers were designed close to highly conserved regions of SAV6658 and these primers were used to amplify the homologous isomerase in *S. cattleya* (gene SCATT 20080). Primer walking was performed on the amplified PCR product which provided the gene sequence for the *S. cattleya* isomerase. The SCATT 20080 gene was then cloned, and over-expressed in *E. coli*, and was subsequently shown (^{19}F NMR assay) to catalyse the conversion of **24** to **25**.[15]

In 2011 the full genome sequence of *S. cattleya* was published[16] following which another isomerase was identified in *S. cattleya* (SCATT 32590). SCATT 32590 was observed to be clustered among other genes expressing proteins related to the methionine salvage pathway suggesting its involvement in primary metabolism. The isomerase expressed by gene SCATT 20080 was not associated with any primary metabolic pathways and assumed to be exclusively involved in the fluorometabolite pathway. Consistent with this a SCATT 20080 gene knock out mutant of *S. cattleya* was viable, but incapable of producing the two fluorometabolites; FAc **10** and 4FT **14**. [15]

4.3.2 Enzymes related to 5-FDRPi by sequence

When 5-FDRPi from *S. cattleya* was first identified in 2008 the BLAST analysis showed that the *S. avermitilis* SAV6658 gene and *S. coelicolor* SCO3014 gene possessed the highest sequence identities to 5-FDRPi.[15] However, more current BLASTp (protein BLAST) analysis of *S. cattleya* 5-FDRPi revealed hits for hundreds of bacterial MTRPi's involved in the methionine salvage pathway, along with proteins related to the initiation factor 2 (IF2) subunit family, which are important regulators of translation initiation in eukaryotes.[17] As genome sequencing has become popular and cost effective, several thousand bacterial genomes have been sequenced and made available, and as such have led to increased hits in BLASTp analysis.[18] Most IF2 protein hits matching *S. cattleya* 5-FDRPi originate from yeast (predominantly *Saccharomyces cerevisiae*). These IF2 family proteins are thought to form a ternary complex consisting of Met-t-RNA, GTP and an initiation factor protein eIF2. Once formed, this ternary complex associates with the smaller subunit of the ribosome to form a preinitiation complex which subsequently binds to mRNA and initiates protein translation. [17,19] Prior to the crystallisation and characterisation of *B. subtilis* MTRPi[20] most MTRPi of prokaryotic origin were mis-annotated as IF2 family proteins.

Currently, 27 genes annotated as MTRPi, have been identified from various *Streptomyces* species, each of which have sequence identity ranging from 75% to 80% to 5-FDRPi and possess a 0 E value (a statistical parameter describing the probability of the number of hits one can expect to see by chance when searching a database, the closer the value to 0 the lesser the chance that it is a random hit). These *Streptomyces* species include *S. olivochromogenes* with a 75% sequence identity, the organism from which xylose isomerase was identified (as previously discussed in section 4.2). *S. aurantiacus* and *S. iranensis*, which both have over 80% sequence identity to the *S. cattleya* 5-FDRPi (Figure 1) *S. aurantiacus* was first isolated in 1969 in China[21], whereas *S. iranensis* was first isolated in 2010 from soil sample in Isfahan city, Iran. Both these locations are geographically disparate to New Jersey, USA where *S. cattleya* was first identified in 1979.[22] *S. xinghaiensis*, a marine S*treptomyces* (discussed in Chapter 2) identified in 2011 from China, also contains an MTRPi gene with 76% sequence identity to 5-FDRPi of *S. cattleya* (Figure 1). In addition, over hundreds of genes in other *Streptomyces* species annotated as either MTRPi or IF2 with acceptable E values were also observed.

Actinobacteria OK006 gene annotated as MTRPi were also shown to have approximately 75% sequence identity to 5-FDRPi (Figure 1). *Actinobacteria* OK006 belongs to the *Streptomyces* genus and was first isolated in Oregon USA in 2015.[23] Despite the existence of many genes with high sequence identity to 5-FDRPi, most of these genes have not been

explored in detail. The closest representative of 5-FDRPi, which has been studied structurally for its mechanism and binding is the MTRPi of *B. subtilis*, which has 35% sequence identity to 5-FDRPi (Figure 1).

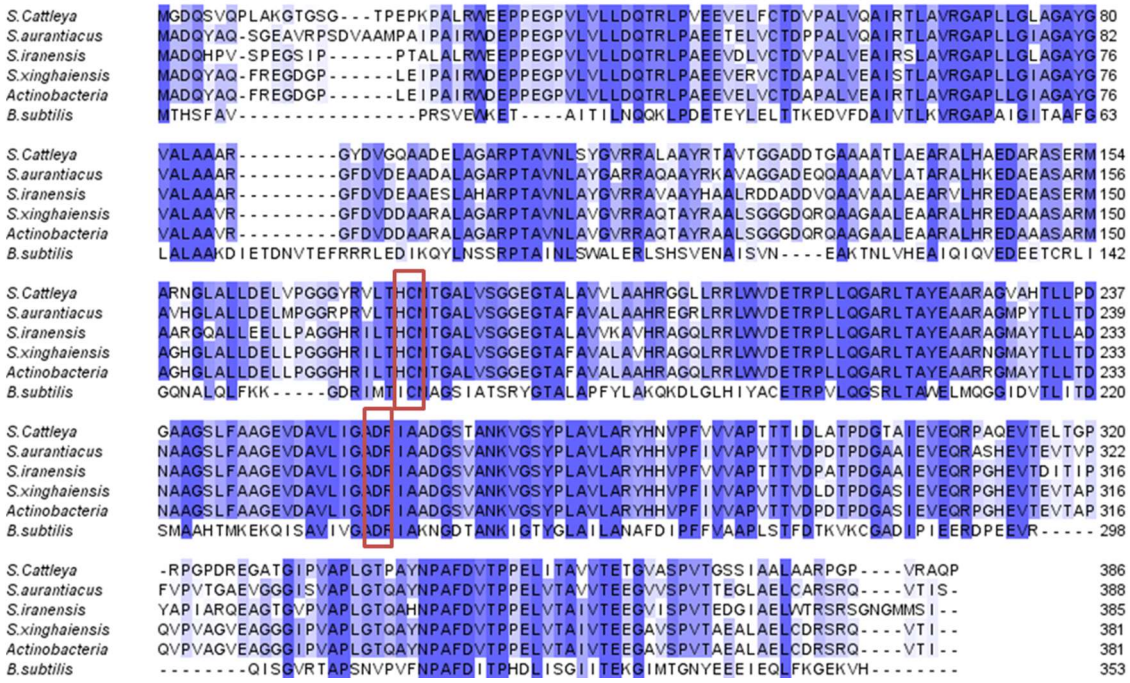


Figure 1: Sequence identity alignment of MTRPi from *S. aurantiacus* (WP_016640738.1), *S. iranensis* (WP_044574421.1), *S. xinghaiensis* (WP_019711652.1), *Actinobacteria* OK600 (WP_054229590.1), *B. subtilis* (WP_041333442.1) and 5-FDRPi of *S. cattleya (*WP_014142773.1). Over 80% sequence identity is represented by dark blue and less than 40% identity is represented by white, lighter shades represent moderate sequence identity. The red box indicates the two catalytic residues which are highly conserved in all of the above mentioned MTRPi's.

4.3.3 The MTRPi from *Bacillus subtilis*

MTRPi plays a key role in the methionine salvage pathway of *B. subtilis* by catalysing the interconversion of **30** and **50** (Scheme 1). MTRPi has protein sequence identity to other IF2 family proteins isolated from *Saccharomyces cerevisiae* (38%). In addition to this, other proteins possessing some sequence identity to *B. subtilis* MTRPi were observed in *Thermotoga maritima (50%), Archaeoglobus fulgidus (43%)* and *Pyrococcus horikoshii* (32%).

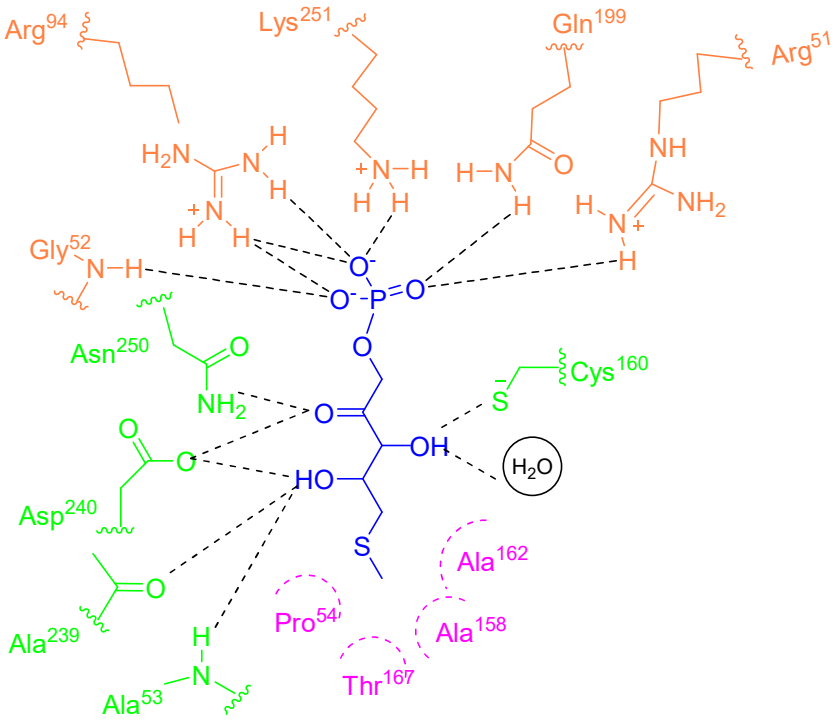
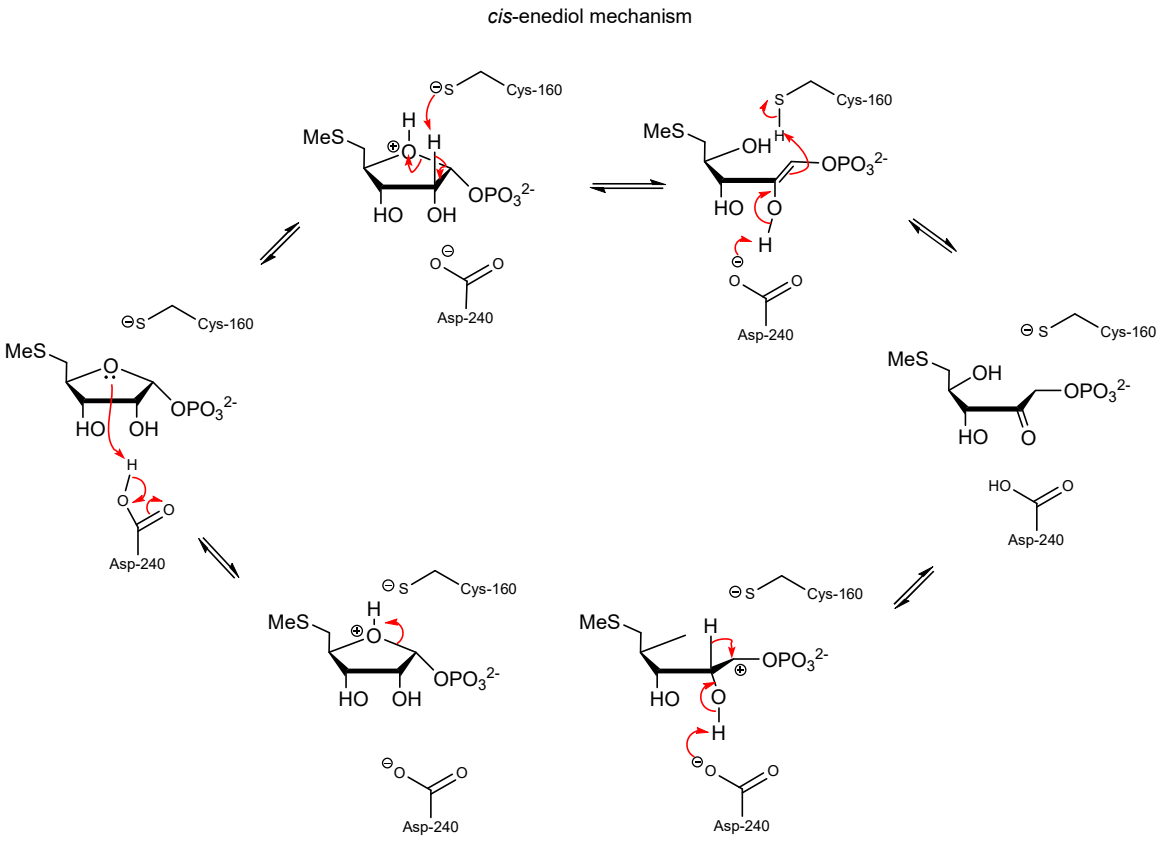


Figure 2: Schematic representation of MTRulP **50** bound active site of *B. subtilis* MTRPi, showing the interactions of **50** with the surrounding amino acids in the active site. The residues involved in binding of the phosphate group (orange), the backbone oxygens or hydroxyls (green) along with the hydrophobic interactions with methylthio group (magenta) is displayed. Image reproduced and modified with permission from Matsumura *et al*.[20]

Enzymology studies carried out on MTRPi have shown that it does not require metal ions for catalysis, analogous to enzymes that are believed to catalyse a *cis*-enediol intermediate isomerisation mechanism. However, NMR studies performed on MTRPi reactions carried out in the presence of D_2O have shown that deuterium was not incorporated into the product, MTRulP **50**,[24] resembling observations found in enzymes involved in a hydride transfer mechanism, as a *cis*-enediol mechanism should promote the deuterium exchange with the solvent. Therefore, attempts were made to obtain a crystal structure of the MTRPi complexed with the product, **50**. The MTRulP **50** bound structure of the *B. subtilis* MTRPi protein was solved by Matsumura in 2008 (PDB: 2YVK). Close analysis of the active site of MTRPi identified the positively charged residues: Arg51, Arg94 and Lys 251 (Figure 2; residues in orange) participating in hydrogen bonding to the phosphate group along with Gln199, to form the enzyme-substrate complex. The methylthio group of **50** was observed to be surrounded by a hydrophobic pocket formed by the conserved residues Pro54, Thr167, Ala158 and Ala162 (Figure 2; residues in magenta). Despite the product-bound crystal structure of MTRPi, these data did not shed any light on the mechanism. Two highly conserved residues, Asp240 and Cys160 (Figure 2; residues in green), were identified in

the active site as crucial for enzyme catalysis. Asp240 was suggested to act as a proton donor/acceptor and Cys160 was suggested to become deprotonated and act as a base for catalysis in *cis*-enediol mechanism or stabilise a transient positive charge on the intermediate during 1, 2 hydride shift (Scheme 4).[20,24] A small number of residues adjacent to these catalytic residues were also found to be conserved (highlighted by dark blue in Figure 1).

The co-crystal structure of the MTRPi enzyme bound to ring opened product MTRulP **50** did not provide any evidence supporting either mechanism. However if a crystal structure of the MTRPi with a ring closed substrate was obtained, it may provide evidence to support one of the potential mechanisms, as such a structure might demonstrate the active site in a 'tighter' configuration, and primed to isomerise the substrate.



Scheme 4: Possible mechanisms of *B. subtilis* MTRPi reaction.[20]

The following section will discuss previous crystallisation studies carried out on 5-FDRPi in an effort to probe its mechanism.

4.4 Previous studies on *Streptomyces cattleya* 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (5-FDRPi)

4.4.1 Site directed mutagenesis of 5-FDRPi

Sequence analysis of the enzymes closely related to 5-FDRPi; the MTRPi of *B. subtilis* (35%) and yeast Ypr118w (26%)[17], reveal several conserved residues which may be involved in catalysis. In the 5-FDRPi, Cys177 with Asn178, along with Asp257 and four adjacent residues were found to be conserved. The crystal structures of the MTRPi enzymes from both *B. subtilis* and yeast were both observed as dimers with the two active sites closely resemble each other. The active sites were found to be located in a pocket formed by α-helices held together by loops in their open conformation, and therefore it was postulated that the catalytic site of 5-FDRPi would be similar to that of *B. subtilis* MTRPi.[17,20]

In order to assess the importance of the Cys177 and Asp257 residues for catalysis, site-directed mutagenesis on 5-FDRPi was carried out. Cys177 was replaced with either an alanine, possessing an unreactive functional group or a serine, where -SH was replaced by -OH. Both mutants were found to be inactive demonstrating the importance of the cysteine residue at this position. [25]

Asp257 is postulated to be a proton donor/acceptor, and this was replaced by directed mutagenesis with either a glycine; or an asparagine where the carboxylate group is replaced with an amide. Both of these mutants were also found to be inactive. The highly conserved hydrophobic pocket within the active site restricts the side chain rotation of Asp257, which is believed to increase the pK_a of the carboxyl group from 3.8 to approximately 7.5.[20] This enables Asp257 to act as a general acid/base during catalysis, a function the amide group of asparagine would not be able to achieve, despite providing the required hydrogen bonding.[25]

4.4.2 Truncation mutants of the 5-FDRPi based on a secondary structure prediction

Many attempts have been made to crystallise native 5-FDRPi, involving the use of various screens, additives, incubation temperatures and seeding experiments, however any crystals obtained from these experiments were multi-layered and diffracted poorly. [25] Therefore the protein sequence of 5-FDRPi was analysed through a secondary structure prediction software (PROF[26]) to identify any disordered loops at the terminus of the protein,

as deletion of such regions can often improve crystallisation.[27,28] The PROF tool compares linear sequences of amino acids to a database of known secondary structures, and predicts the regions that are likely to fold either as α-helices, β-sheets or loop structures. Several disordered loop regions were identified in 5-FDRPi, out of which two regions were chosen from the N-terminus and two from the C-terminus. Truncated mutants were prepared and over-expressed in *E. coli*, however the 5-FDRPi over-expressed from all truncated mutants were found to be insoluble. Furthermore, systematic truncation mutants were also prepared of 5-FDRPi, where the peptide was sequentially shortened from both N terminus and C terminus, five residues at a time, and over-expressed in *E. coli*. The N-terminus shortened mutants were all found to produce insoluble protein whereas C-terminus mutants shortened by up to 25 residues were found to be partially soluble, demonstrating the importance of the N terminus in protein folding and solubility. A number of the C terminus truncated mutants of 5-FDRPi which produced soluble protein were over-expressed in *E. coli* in an attempt to produce crystallised protein, however these over-expressed proteins were found to be unstable and degraded over time.[25]

4.4.3 A phosphonate analogue of the 5-FDRPi substrate

In an attempt to crystallise a cyclic sugar to explore the key residues for binding and probe the catalytic mechanism of FDRPi, a non-hydrolysable analogue of 5-FDRP **24** was prepared by replacing the linking oxygen of the phosphate ester with a methylene group (Figure 3). The synthetic route to the 5′-fluororibosephosphonate **55** was designed by Nasomjai starting from D-ribose **39**.[29] It was postulated that this phosphonate would bind the active site despite the absence of the oxygen, but could not be hydrolysed by the enzyme. It may act as an inhibitor and co-crystallisation ligand.

5-FDRP **24** 5'-fluororibosephosphonate **55**

Figure 3: The substrate 5-FDRP **24** and the non-hydrolysable 5′-fluororibosephosphonate **55** postulated to be an inert substrate.

4.5 Project aims

In order to gain an insight into the mechanism of 5-FDRPi an attempt was made to co-crystallise the 5-FDRPi with the inert 5′-fluororibosephosphonate **55** substrate.[25] 5-FDRPi has previously been cloned, over-expressed and purified successfully in *E. coli*, and as such the 5-FDRPi was again purified using the same protocol, and extensive crystallisation trials were performed. Furthermore, ITC and NMR binding assays were carried out to assess the binding affinity of **55** analogue with 5-FDRPi.

4.6 Results and discussion

4.6.1 Over-expression and purification of 5-FDRPi

The cloning and over-expression of 5-FDRPi in *E. coli* has been reported previously.[15] The genomic DNA of *S. cattleya* was obtained following the standard procedure described in Kieser *et al* (2000), [30] coupled with a phenol/chloroform DNA extraction.[31] The *SCATT 20080* gene was amplified by PCR using genomic DNA as a template, and a set of primers designed to complement the gene. The PCR product was first cloned into the vector pGEM-T (Promega), as direct cloning into pEHISTEV proved difficult. The recombinant vector pGEM-T-*SCATT20080* was then transformed[32] in to DH10β *E. coli*. The transformed *E. coli* were analysed by colony PCR and visualised by agarose gel electrophoresis for the presence of the pGEM-*SCATT 20080*. The colonies containing the recombinant plasmid were sequenced to confirm the accuracy of the inserted gene product. The correct recombinant plasmid was digested with the restriction endonucleases NcoI and XhoI, and the gene was subcloned into enzyme digested pEHISTEV. The recombinant plasmid, pEHISTEV-*SCATT20080,* was then transformed[32] into BL21-Gold (DE3) *E. coli* cells. The pEHISTEV vector was designed to add a cleavable N-terminal poly-histidine tag (His_6-tag) to the expressed protein. The His_6-tag can be cleaved by treating the protein with TEV protease. This protease acts on a cleavage site inserted between the expressed protein and the His_6-tag.[33]

The over-expression and purification procedure for 5-FDRPi was adapted from that used to isolate *S. coelicolor* 5-FDRPi (SCO3014 gene product).[15] The protocol used to purify *S. cattleya* 5-FDRPi was a modified version of the previously reported procedure.[25] BL21-Gold (DE3) *E. coli* cells containing pEHISTEV-*SCATT20080* were used to inoculate large cultures of LB supplemented with kanamycin. The 5-FDRPi expression was induced with

IPTG and the cultures were grown for 20 h at 25 ºC following which the cells were harvested and re-suspended in lysis buffer. The cells were lysed using one shot cell disrupter and the soluble fraction was loaded on to a Ni-NTA column pre equilibrated with wash buffer. The 5-FDRPi was eluted using a buffer containing high imidazole concentration and the fractions obtained were analysed by SDS-PAGE. The 5-FDRPi was visualised as a ~42 kDa protein in the elution fractions (Figure 4). Addition of 1 mM β-ME to all buffers was required to prevent the oxidation of the thiol groups of cysteine residues which would promote protein aggregation.

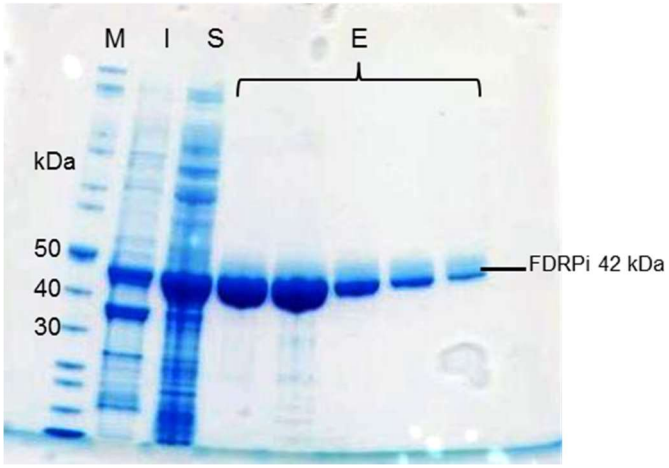


Figure 4: The SDS-PAGE gel electrophoresis of fractions obtained from Ni-NTA column. M: un-stained PageRuler protein ladder, S- soluble fraction, I- insoluble fraction, E- elutions.

The pooled fractions containing 5-FDRPi was incubated with His_6-tagged TEV protease in 1:20 ratio (TEV: 5-FDRPi) whilst dialysing against 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 mM β-M. The progress of the TEV protease mediated His_6-tag cleavage of 5-FDRPi was monitored by SDS-PAGE gel. Subsequently, a second Ni-NTA column was run to separate the His_6-tagged-TEV, uncleaved His_6-tagged 5-FDRPi and cleaved His_6-tags from the His_6-tag cleaved 5-FDRPi. The dialysed TEV digested 5-FDRPi fraction was loaded onto a Ni column and all fractions obtained were monitored by SDS-PAGE gel. The flow-through out of the Ni column was collected as the His_6-tag removed 5-FDRPi would not bind the column and elute immediately (Figure 5).

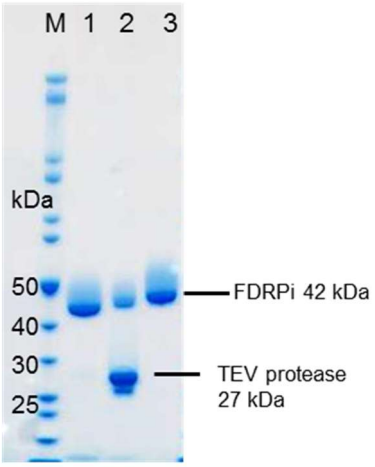


Figure 5: His_6-tag of 5-FDRPi was cleaved by incubating with purified TEV protease. M: un-stained PageRuler protein ladder (Thermo scientific), 1= 5-FDRPi post incubation with TEV, 2= TEV protease and 5-FDRPi, 3= 5-FDRPi prior to incubation with TEV.

The 5-FDRPi obtained was concentrated and loaded onto a Superdex 200 gel filtration column and eluted with buffer containing 20 mM Tris pH 7.5, 150 mM NaCl and 1 mM TCEP (Figure 6). TCEP is a considerably more stable reducing agent compared to more common reducing agents such as β-ME and dithiothreitol (DTT). In order to determine the molecular weight of the 5-FDRPi, gel filtration standard, which is a mixture of proteins of known molecular weight (Bio-Rad), was run using the same buffer as the 5-FDRPi in the same GF column. Based on the values obtained for the elution volume of each known protein, the logarithm of the molecular weights was plotted to a graph. Using the elution volume of 5-FDRPi, the molecular weight (log) of 5-FDRPi could be obtained to a moderate accuracy, and was observed to be around 82,300 Da. This suggests the 5-FDRPi may exist as a dimer (~84,000 Da) analogous to the *B. subtilis* MTRPi.

The SEC fractions containing 5-FDRPi were pooled and concentrated (10-20 mg/ml) for assays and crystallisation trials. The identity of the protein was confirmed as *S. cattleya* 5-FDRPi by MALDI-TOF analysis.

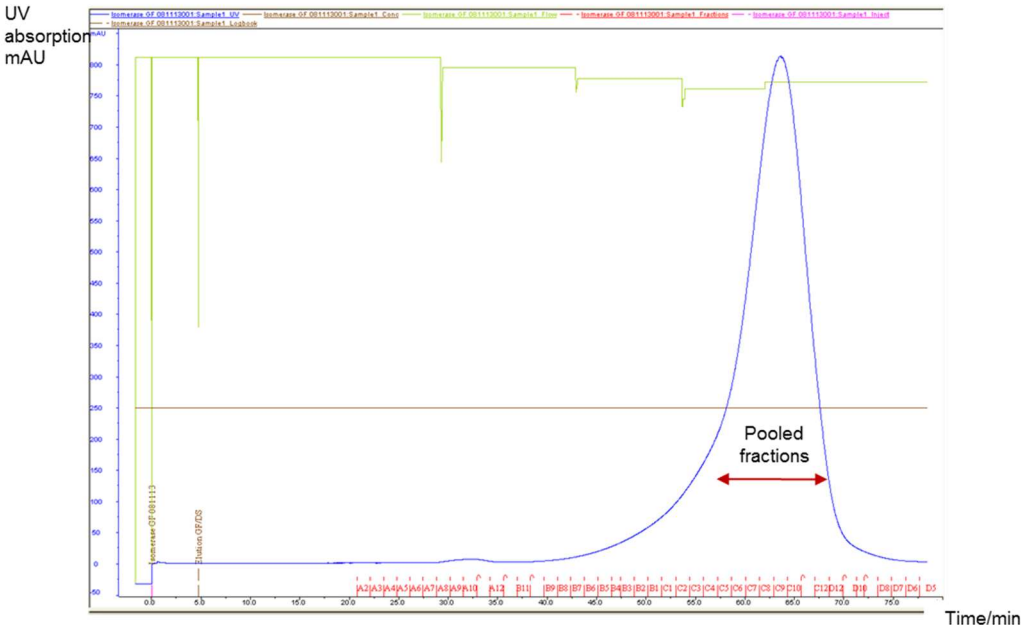
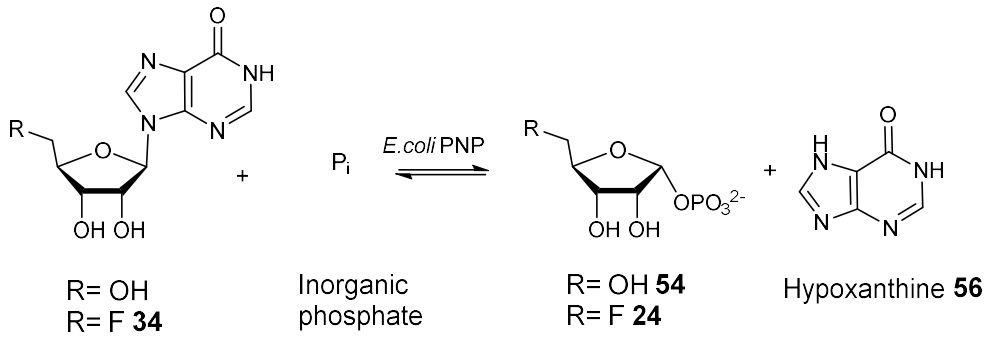


Figure 6: Size exclusion column chromatography of *S. cattleya* 5-FDRPi showing the UV absorption in blue, concentration of buffer in brown, flow rate in green and fractions in red.

4.6.2 NMR Assaying 5-FDRPi and small scale enzymatic preparation of 5-FDRP 24

The activity of the purified 5-FDRPi was assayed by ^{19}F-NMR.[14,15] Previously, the 5-FDRP **24** required for the 5-FDRPi assays was prepared enzymatically by treating synthetically prepared FDA **23**[34] with adenosine deaminase, to generate FDI **34**, this was then treated with recombinant *E. coli* PNPase (Sigma) to form **24**.[13,25] The natural substrate for this PNPase is inosine, which is converted to hypoxanthine **56** and ribose-1-phosphate **54** with the use of free phosphate (Scheme 5). Incubation of this enzyme with the unnatural substrate FDI **34**, mediated a slow, low yielding, reaction to produce **24** and xanthine which was then coupled to 5-FDRPi to generate **25**. The production of **25** was confirmed by GC-MS, however, the production of **24** from *E. coli* PNP was low yielding and inefficient, therefore an attempt was made to over-express the PNP (flB) from *S. cattleya*.



Scheme 5: *E. coli* PNP catalyses inosine/FDI **34** to ribose-1-phosphate **54** and hypoxanthine **56** in the presence of phosphate buffer at pH 7.5.

Over expression and purification of *S. cattleya* PNP proved to be difficult as the overexpressed PNP was found to be insoluble, however some partially purified protein could be obtained as an MBP-tagged fusion protein (to increase its solubility). Partially purified *S. cattleya* PNP was successfully coupled with the 5-FDRPi enzymatic assays consistent with the proposed biosynthetic pathway for *S. cattleya*.[15] However sufficient quantities of *S. cattleya* PNP could not be obtained to assay 5-FDRPi due to low quantities of soluble *S. cattleya* PNP obtained. Additionally an immobilised bacterial PNP enzyme, (kindly provided by GSK) was used to perform the PNP reaction starting from synthetic FDA **23**. This PNP proved to be much more efficient than *E. coli* PNP and was subsequently used to produce 5-FDRP **24** from **23**. A solution of **23** (1 mg) was incubated with immobilised PNP in 20 mM phosphate pH 7.5 with shaking at 40 rpm for 2 h. The reaction was filtered and the supernatant was incubated with purified 5-FDRPi (120 mM) at 37 ºC for 3 h and analysed by ^{19}F-NMR. This demonstrated the conversion of **24** to 5-FDRulP **25** in addition to some unreacted **23**. Therefore the immobilised PNP was found to be more efficient than PNP used in previous experiments.

5-FDRP **24** has been produced as an intermediate in small quantities during these PNP mediated enzyme reactions (which are then coupled to 5-FDRPi). As such, **24** has not been isolated and characterised previously due to its low production. Using the immobilised PNP, small scale production of 5-FDRP **24** was achieved. FDA **23** (6mg) was incubated with immobilised PNP (GSK) in 20 mM phosphate buffer pH 7.8 in a total reaction volume of 5 ml. The reaction was incubated at 37º C shaking at 180 rpm over 5 days. The progress of the reaction was monitored by ^{19}F-NMR. The reaction was observed to be slow; however, after 5 days of incubation, almost 96% of **23** has been converted to **24** (Figure 7).

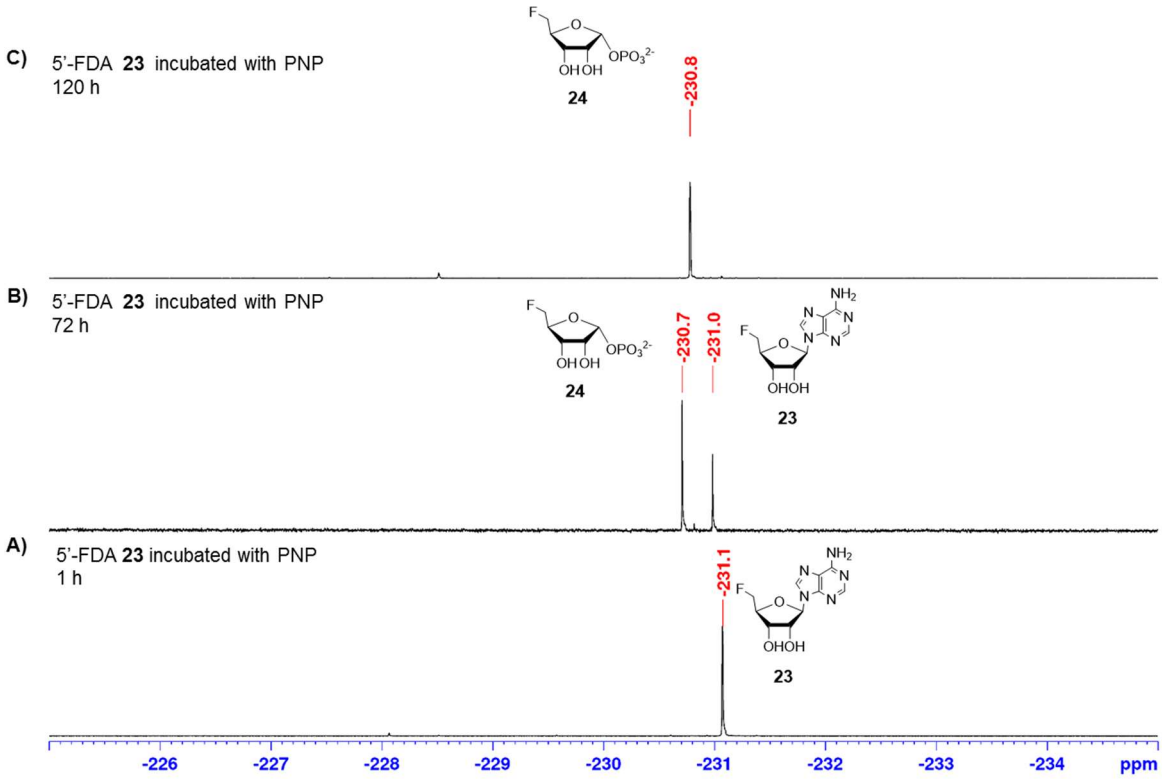


Figure 7: ^{19}F {^{1}H} NMR assays showing the conversion of 5′-FDA **23** to 5-FDRP **24** with immobilised PNP.

In order to separate the unreacted **23** from **24** several solid phase extraction (SPE) cartridges were trialled. Initially Discovery® DSC-SCX (strong cation exchange) was utilised in an effort to retain **23** and allow **24** to elute, along with the phosphate buffer. However no separation was observed using this SCX cartridge. Following this the Discovery® DSC-NH_2 (aminopropyl; weak anion exchanger) (Sigma) cartridge was used, however only ~20% of the 5-FDRP **24** was eluted with the remainder eluting with **23**. The next SPE cartridge trialled was the SOLA™WAX (weak anion exchanger) (Thermo Scientific). The pH of the PNP reaction product was lowered to pH 3.5 prior to loading on to the SOLA™WAX cartridge (Thermo scientific), which was then equilibrated with water-formic acid pH 3.5 (Figure 8; wash 1). The cartridge was then flushed with 100% methanol-formic acid pH 3.5 (Figure 8; wash 2 and 3) followed by 50% methanol- ammonium hydroxide pH 9 (Figure 8; elution) as described in Experimental 4.8.3. The fractions collected were freeze-dried and re-suspended in D_2O and analysed by ^{19}F {^{1}H} NMR. Wash 1 was observed to contain almost all of the unreacted 5′-FDA **23** and wash 2 was found to contain most of the **24** with a little **24** in elution 1. Wash 3 contained small quantities of both **23** and **24**. Wash 2 and elution 1 fractions were pooled and freeze-dried. This purification

procedure was found to be sufficient and **24** (~5mg) was obtained in phosphate buffer, and was characterised by ^{1}H and ^{13}C NMR (Figure 9). Furthermore 2D NMR analysis for 5-FDRP **24** was also performed as in Appendix 4.

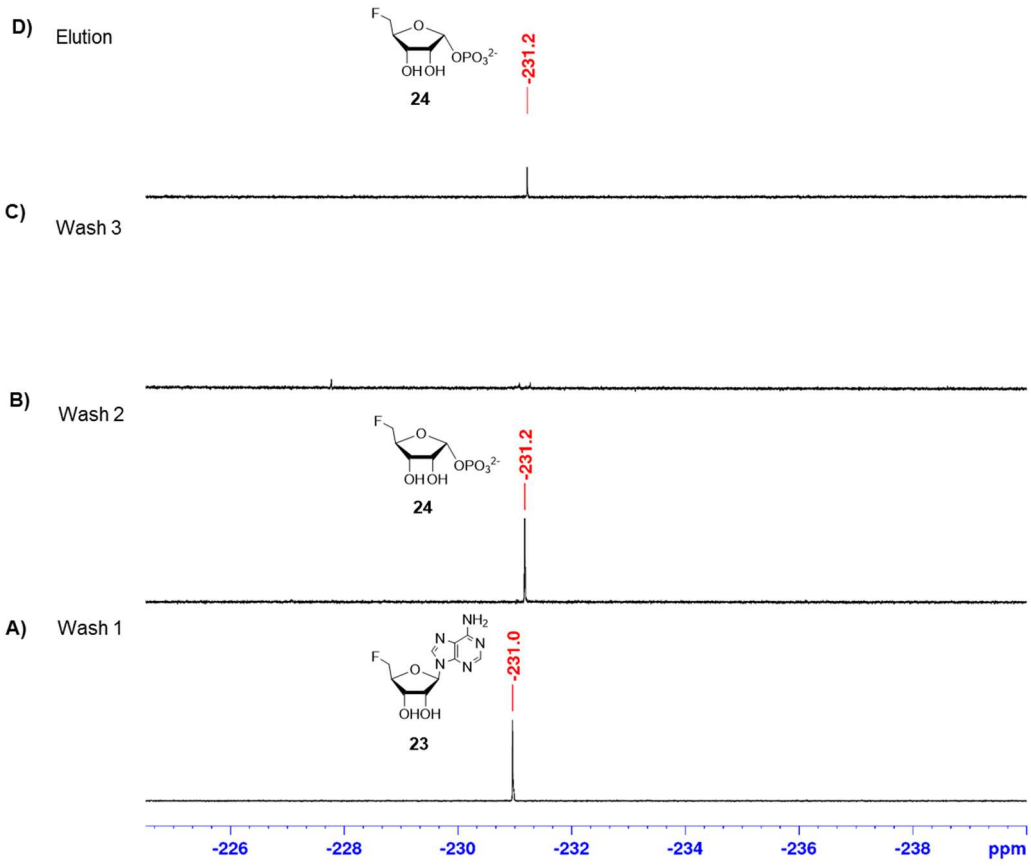


Figure 8: PNP reaction mixture was loaded to SOLA™WAX SPE cartridges and fractions collected were analysed by ^{19}F {^{1}H} NMR.

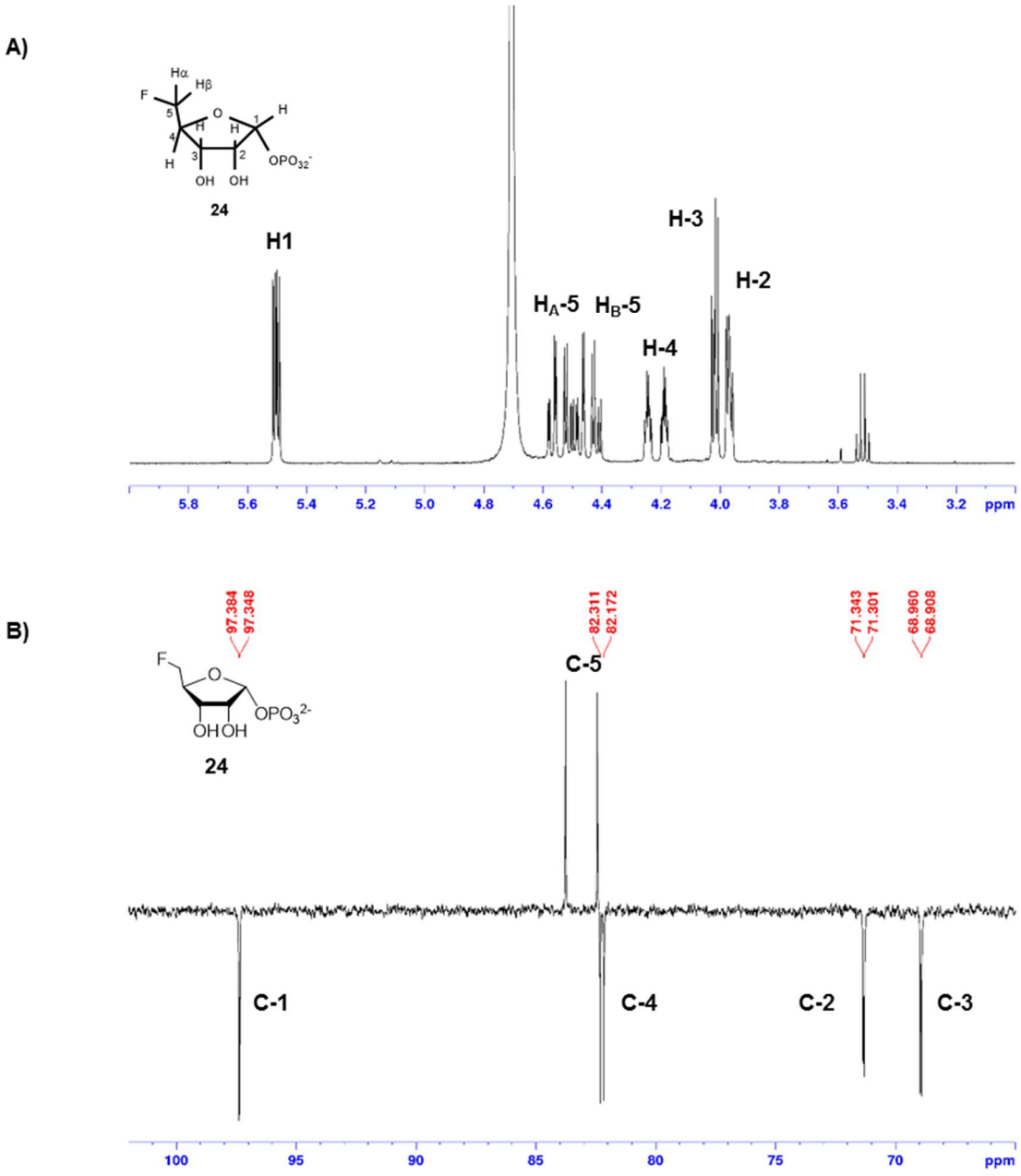


Figure 9: **A)** The ^{1}H NMR spectrum of 5-FDRP **24** in D_2O. **B)** The ^{13}C NMR spectrum of **24** in D_2O.

5-FDRP **24** produced *via* PNP was used in an isomerase assay (according to a previously reported procedure[25]). Thus **24** (1 mg) was incubated with purified 5-FDRPi (3 mg) at 37 ºC for 5 h. The production of **25** was observed as determined by ^{19}F {^{1}H} NMR (Figure 10).

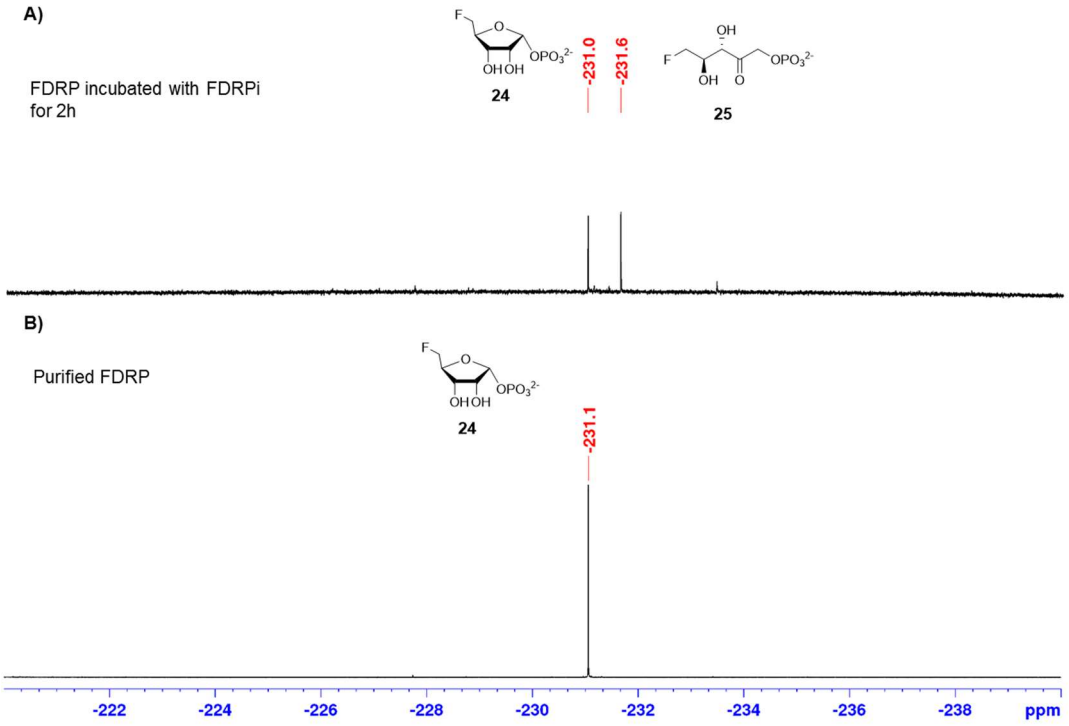


Figure 10: **A)** 5-FDRP **24** (~1 mg) produced *via* PNP was incubated with 5-FDRPi (3 mg) at 37 ºC for 5 h. **B)** The production of 5-FDRulP **24** was observed by ^{19}F {^{1}H} NMR. (These 5-FDRPi assays were carried out based on previous work[25]).

4.6.3 Crystallisation of 5-FDRPi

Attempts to crystallise native 5-FDRPi had previously proved unsuccessful, resulting in protein crystals with overlapping layers, which diffracted poorly.[25] Despite this lack of success, attempts were made to crystallise the native 5-FDRPi using a wide range of commercial screens and performing successive grid search optimisations to narrow the crystal forming conditions. Purified 5-FDRPi in 20 mM Tris pH 7.5, 150 mM NaCl and 1 mM TCEP was crystallised using the sitting drop vapor diffusion method in 96 well crystallisation plates. Routine screening crystal trials for 5-FDRPi were performed at 20 ºC and 4 ºC using four commercial screens; JSCG+, Wizard 1+2 (Molecular dimensions), PEGs I and II suite (Qiagen). All crystal trials were set up with an Arts Robbins Gryphon at protein to precipitant ratios of 2:1 and 1:1. Two 5-FDRPi concentrations at 14 mg/ml and 7.5 mg/ml were used to set up the initial screens as summarised in Table 1.

Table 1: Crystallisation conditions for 5-FDRPi used for initial screening and their outcomes.

Initial screen	5-FDRPi concentration 14 mg/ml	5-FDRPi concentration 7.5 mg/ml	Observation
JSCG+	4 ºC	4 ºC	No crystals were observed
	20 ºC	20 ºC	
Wizard 1+2	4 ºC	4 ºC	No crystals were observed
	20 ºC	20 ºC	Small crystals in few conditions (Table 2) in plate containing 14 mg/ml 5-FDRPi (E2)
PEGs I	4 ºC	4 ºC	No crystals were observed
	20 ºC	20 ºC	Small crystals in few conditions (Table 2) in plate containing 14 mg/ml 5-FDRPi (D10)
PEGs II	4 ºC	4 ºC	No crystals were observed
	20 ºC	20 ºC	Small crystals in few conditions (Table 2) in plate containing 14 mg/ml 5-FDRPi (G1)

Small rice-shaped, trapezohedron crystals (Figure 11A & B) were observed in three wells (Table 2) one in each screen, Wizard 1+2 (E12), PEGs I (D10) and PEGs II (G1) incubated at 20 ºC with 14 mg/ml 5-FDRPi. No crystals were observed in plates incubated at 4 ºC and plates set up with 7.5mg/ml 5-FDRPi. The three conditions were optimised, increasing the 5-FDRPi concentrations (16 mg/ml) and varying PEG concentrations, additive and buffer concentrations and pH as mentioned in Table 2. The outcome of the optimisations are

mentioned in Table 3. Few wells in D10 and G1 optimisations produced rice shaped, trapezohedron crystals as seen in Figure 11A and B. One condition in G1 optimisation (E1 well) produced thin sheet like crystals (Figure 11C) which had not been observed previously. Therefore, further optimisation screening of E1 well was set up using 16mg/ml 5-FDRPi (Experimental 4.8.5; Table 4).

The best crystals (Figure 11D) were obtained in 25% PEG 5K MME, 0.1 M Tris-HCl pH 7.6. These crystals were allowed to grow for 9 days before harvesting. The crystals grew as clustered bristles fanning outwards at one end. A small number of these bristles were broken off, picked and flash-frozen in a cryoprotectant containing precipitant, supplemented with 35% glycerol. The crystals were shot in-house at 100 K on a Rigaku 007 HFM rotating anode X-ray generator with a Saturn 944 CCD detector. A dataset of 236 frames of 0.5° oscillation with a crystal-to-detector distance of 60 mm and 240 sec exposure per frame was collected. The data were indexed and integrated using *MOSFLM* and scaled using *SCALA* to obtain a crystal structure up to 1.7 Å resolution with an R value of 0.25 (a measure of the quality of the atomic model obtained from the crystallographic data). The crystal structure was resolved by molecular replacement on MTRPi structure of *B. subtilis* by Dr Jesko Koehnke at University of St Andrews (X-ray diffraction data is presented in Appendix 2). The active site of the 5-FDRPi crystal structure was found to be empty. Subsequently, the same crystal forming conditions were used to set up co-crystallisation trials containing **55**. These co-crystallisations trials however proved to be unsuccessful.

Table 2: Crystal forming conditions of the commercial screens, Wizard 1+2 (Molecular dimensions), PEGs and PEGs II suite (Qiagen) were the basis to design optimisation screens.

Commercial Screen (14 mg/ml 5-FDRPi, 20 ºC)	Well	Contents of the well			Optimisation			
		Precipitant	Buffer	Salt	Precipitant	Buffer	Salt	5-FDRPi concentration
Wizard 1+2 14 mg/ml 5-FDRPi, 20 ºC	E12	30 % w/v PEG 400	0.1 M sodium cacodylate pH 6.5	None	30- 40% w/v PEG 400,	0.1 M sodium cacodylate pH 6.5 – 7.5	None	14 mg/ml or 16 mg/ml
PEGs I 14 mg/ml 5-FDRPi, 20 ºC	D10	20% w/v PEG 8000	0.1 M Tris-HCl pH 7.5	0.1 M $MgSO_4$	15- 20% w/v PEG 8000	0.1 M Tris-HCl pH 7.5-7.8	0- 0.3 M $MgSO_4$	14 mg/ml or 16 mg/ml
PEGs II 14 mg/ml 5-FDRPi, 20 ºC	G1	25% w/v PEG 5000 MME	0.1 M Tris-HCl pH 7.5	0.2 M $MgSO_4$	19-24% w/v PEG 5000 MME	0.1 M Tris-HCl pH 7.5-7.8	0- 0.3 M $MgSO_4$,	14 mg/ml or 16 mg/ml

Table 3: Summary of the 5-FDRPi crystallising trials obtained from initial optimisation conditions.

Screen E12 ;20 ºC	Screen D10 ;20 ºC	Screen G1;20 ºC
5-FDRPi concentration	5-FDRPi concentration	5-FDRPi concentration
16 mg/ml	16 mg/ml	16 mg/ml
Spherulites and precipitate.	Few conditions with rice shaped, trapezohedron crystals	Few conditions with rice shaped, trapezohedron crystals and one condition (E1) with thin sheet like crystals

The structure was solved with PHASER molecular replacement, using the *B. subtilis* MTRPi. The structure was refined iteratively through manual building in COOT, and an automated refinement was carried out using PHENIX.

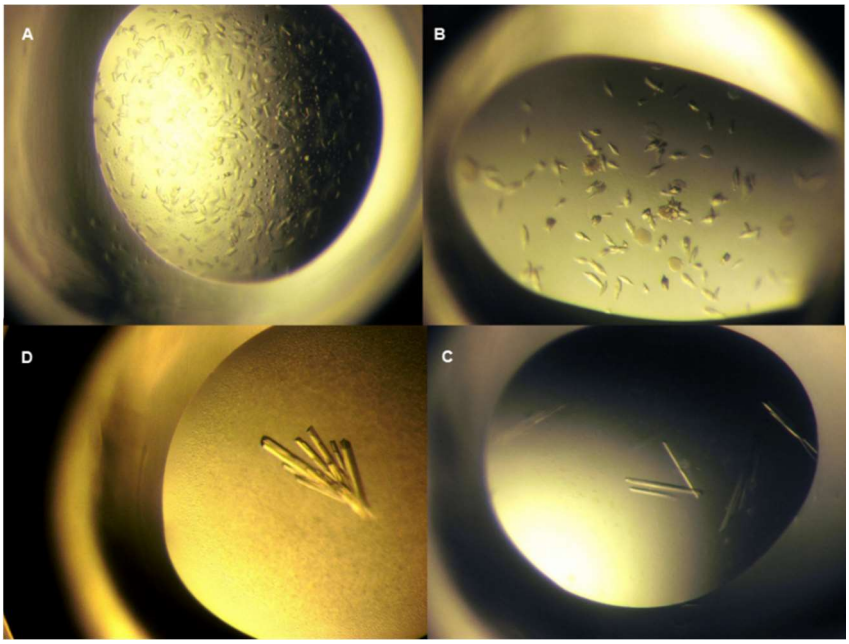


Figure 11: Crystals of 5-FDRPi **A)** Rice shaped crystals produced from initial optimisations **B)** Trapezohedron shaped crystals obtained during initial and secondary optimisation screens **C)** Thin sheet like crystals only obtained in E1 well of G1 optimisation screen **D)** Thick bristles of clustered crystals was obtained from E1 well optimisation screen.

Co-incubation with the 5′-fluororibosephosphonate **55** failed to produce any co-crystals. It was possible that the **55** does not bind to 5-FDRPi, despite a close structural relationship to the natural substrate, therefore binding was explored by conducting Isothermal titration calorimetry (ITC) studies.

4.6.4 Binding assays by Isothermal titration calorimetry (ITC)

In order to assess the binding affinity of the 5′-fluororibosephosphonate **55**, isothermal titration calorimetry was employed. This involved a series of independent experiments where **55** was titrated into a solution of 5-FDRPi in phosphate buffer. ITC analysis of 5-FDRPi with ribose-1-phosphate **54** was considered as a prospective positive control. Ribose-1-phosphate **54** is not commercially available and would require to be produced enzymatically. Incubating *E.coli* PNP with either adenosine **47** inosine in phosphate buffer similar to the enzymatic synthesis of 5-FDRP **24** produced ribose-1-phosphate **54**. However preparation of **54** in quantities required for ITC experiments proved difficult. In addition, the enzyme reaction required phosphate buffer which could not be completely removed from **54**. Any mismatch in buffer and salts between the ligand and enzyme would produce a high heat of dilution on ITC which potentially masks any heat produced upon binding.

5′-Fluororibosephosphonate **55** (1 mM) was titrated into 5-FDRPi (60 μM or 100 μM) and analysed by ITC. No thermal signal was observed, only the heat of dilution and background enthalpy of mixing could be recorded, as shown in Figure 12A. As a positive control tri-N-acetyl-glucosamine (tri-NAG) (1.5 mM) was titrated into chicken egg white lysozyme (150 μM) in phosphate buffer (Figure 12B). Each exothermic heat pulse (Figure 12B; upper panel) corresponds to individual injections (10 μ) of tri-NAG. Integrated heat data (Figure 12B; lower panel) constitute a differential binding curve that fit to a standard single binding site model to calculate the binding affinity (K_m).

This suggests 5′-Fluororibosephosphonate **55** may not bind to 5-FDRPi and highlights the importance of the bridging oxygen of **55** which must form important interactions within the active site. Replacing the oxygen with a methylene group may abolish any such interaction causing 55 to not bind to the active site.

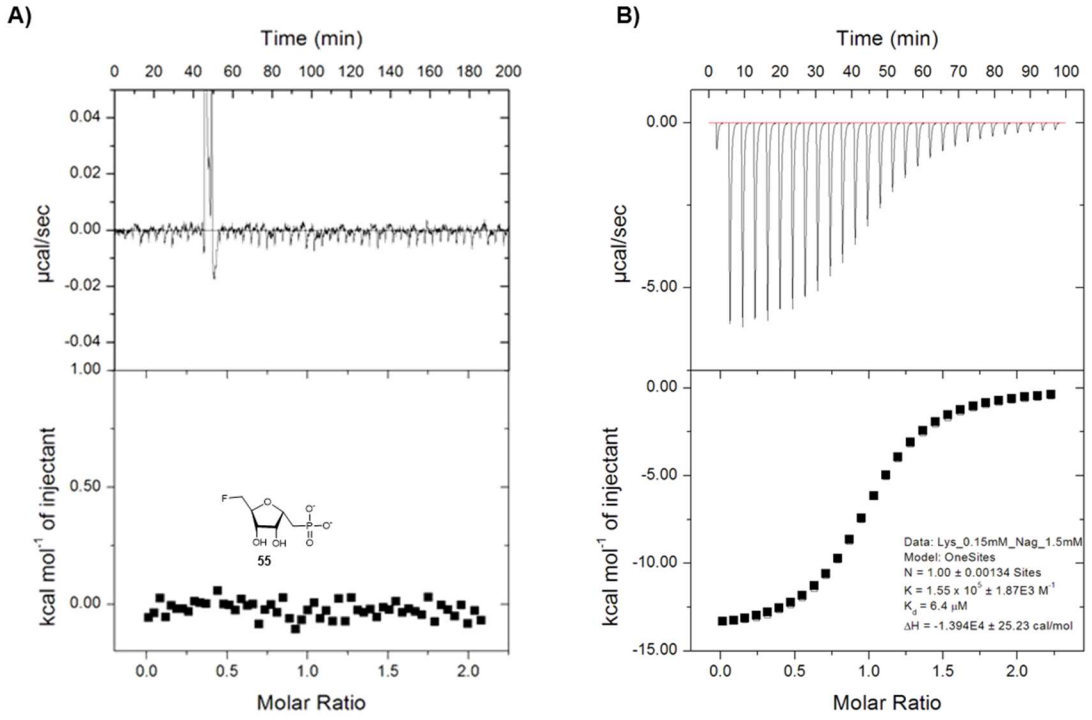


Figure 12: **A)** Isothermal titration calorimetric (ITC) determination of the binding affinity of 5′-fluororibosephosphonate **55** (1 mM) when titrated into FDRPi (60 µM). No binding was observed for **55**. **B)** Positive control for ITC: binding affinity of tri-N-acetyl-glucosamine (tri-NAG) (1.5 mM) titrated into chicken egg white lysozyme (150 µM).

4.6.5 Structural comparison of *B. subtilis* MTRPi and *S. cattleya* 5-FDRPi

X-ray structural analysis of apo-5-FDRPi indicated that it exists as a dimer. This was also supported by SEC chromatography (4.6.1). The crystal structure of apo-5-FDRPi was superimposed on to the crystal structure *B. subtilis* MTRPi and both were found to have a very similar monomeric arrangement (Figure 13A), despite only possessing a 35% aa sequence identity.[20] The tertiary structure of *B. subtilis* MTRPi was found to exist as a dimer, matching that of the preferred arrangement of 5-FDRPi. The structural alignment of the dimers of 5-FDRPi and *B. subtilis* MTRPi confirmed them to be conformationally very similar (Figure 13B). The amino acid sequence similarity was assessed in both *S. cattleya* 5-FDRPi and *B. subtilis* MTRPi by pairwise sequence alignment as the crystal structures of both were very similar despite low sequence identity. The 5-FDRPi and MTRPi were found to share ~48% amino acid sequence similarity which may contribute to the structural similarity (Figure 13C).

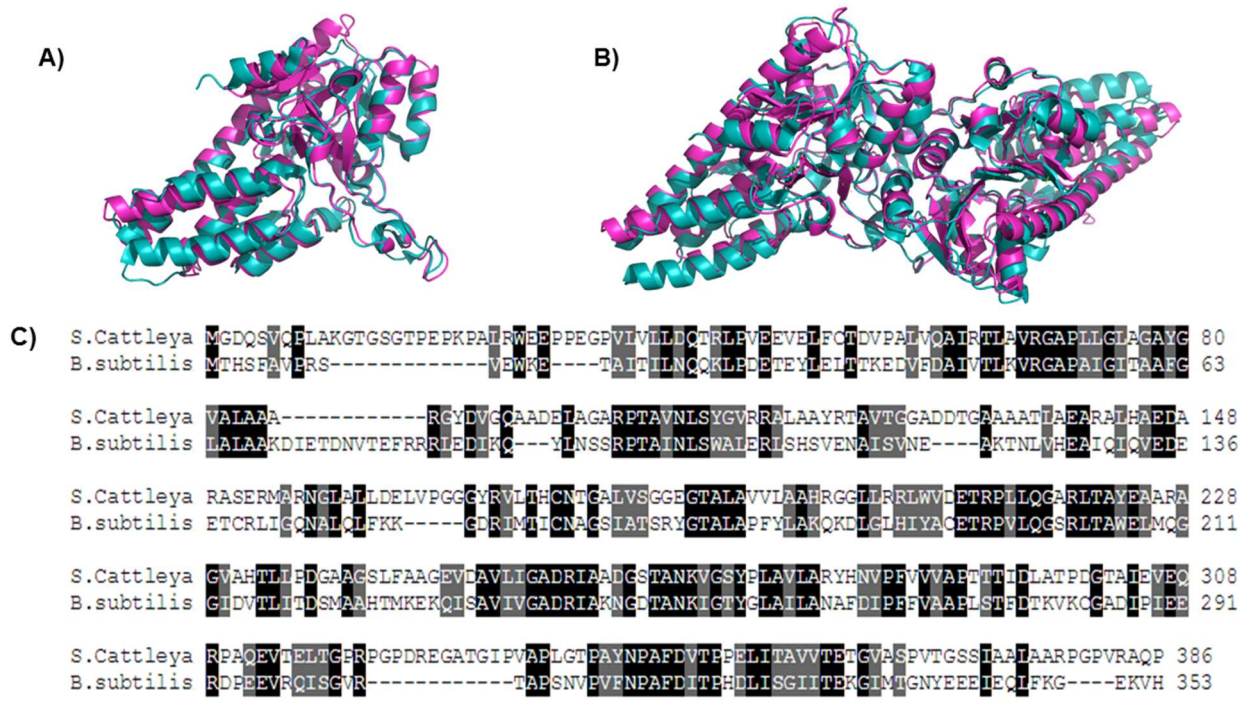


Figure 13: Superimposition of the structures of *S. cattleya* 5-FDRPi and *B. subtilis* MTRPi. **A)** Superimposition of monomeric 5-FDRPi of *S. cattleya* (magenta) and MTRPi of *B. subtilis* (teal). **B)** Superimposition of dimeric 5-FDRPi of *S. cattleya* (magenta) and MTRPi of *B. subtilis* (teal). **C)** Sequence similarity of both *S. cattleya* 5-FDRPi and *B. subtilis* MTRPi. (Residues that are identical are marked in black background, and amino acids that are similar among the sequences are given a grey background.

As the active site residues have been previously identified in the MTRulP **50** bound structure of *B. subtilis*, this MTRPi crystal structure was mapped on to the empty structure obtained from the *S. cattleya* 5-FDRPi, in order to observe any similarities of their active site formations (Figure 14A). All the amino acid residues which form hydrogen bonds with the phosphate group and backbone oxygens or hydroxyls of **50** were found to be conserved in 5-FDRPi (Figure 14A). Most of the residues providing the hydrophobic interactions with the methylthio group of **50** in *B. subtilis* MTRPi were also found to be conserved in the yeast MTRPi, except for Ala162, which is replaced with threonine in Ypr118w (PDB code: 1W2W).[20] Similarly Ala162 was now Thr179 in 5-FDRPi. In addition to this, Ala166 and Thr167 of *B. subtilis* MTRPi were found to be replaced by Val183 and S184 respectively in the 5-FDRPi structure. Close analysis of the active site of *B. subtilis* MTRPi revealed that the loop formed by the three amino acids Ala162, Ala166 and Thr167 is coiled twice, thus reducing the space available within the active site, conversely in 5-FDRPi the same loop now formed by Thr167, Val183 and S184 is more relaxed (Figure 14B). This reconfiguration

of the loop might be important in interacting with the terminal fluorine of the substrate, 5-FDRP **24**.

The mapped arrangement of the active site of 5-FDRPi is very similar to the active site of *B. subtilis* MTRPi, consistent with the high level of residue similarity, and both enzymes may reasonably operate *via* a similar mechanism.

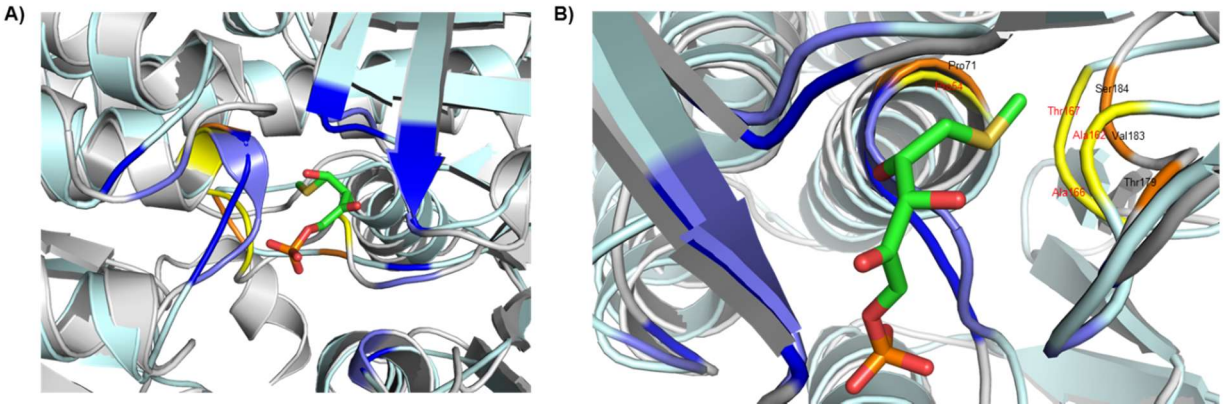


Figure 14: **A)** Superimposition of predicted active site residues for *S. cattleya* 5-FDRPi using *B. subtilis* MTRPi–MTRulP **50** bound structure as reference. 5-FDRPi in cyan, MTRPi in grey, highly conserved active site residues interacting around the phosphate group and the backbone oxygens of MTRPi (light blue) and possibly around 5-FDRPi (dark blue). The residues interacting with methylthio group of MTRulP **50** (yellow). Same residues mapped on to 5-FDRPi (orange) **B)** Close up of the arrangement of residues interacting with the terminal methylthio group of *B. subtilis* MTRPi (yellow residues annotated in red). The same residues mapped on to *S. cattleya* 5-FDRPi which would bind 5-FDRulP **25** with a terminal fluorine (orange residues annotated in black)

4.7 Conclusions

The objective of this aspect of the research was to obtain an X-ray crystal structure of the *S. cattleya* 5-FDRPi with the phosphonate analogue of 5-FDRP **24** bound to the active site, in an effort to provide a structure of a ring closed substrate and perhaps some insight into the mechanism of this enzyme. A crystal was obtained which diffracted to provide a structure of 1.7 Å resolution. The structure obtained of 5-FDRPi had an empty active site. In order to assess if the **55** would bind to the 5-FDRPi, binding studies were performed by ITC and NMR. Both binding studies suggested 5-FDRPi does not bind **55** and therefore subsequent co-crystallisation attempts were terminated. Comparative crystal structure

analysis of the *S. cattleya* apo-5-FDRPi and *B. subtilis* MTRPi bound to MTRP **30** revealed a similar tertiary structure and similar arrangement within the active site, which suggests that the two enzymes function *via* a similar mechanism. However this remains to be established unambiguously for 5-FDRPi.

4.8 Experimental

4.8.1 Over-expression and purification of 5-FDRPi

BL21-Gold (DE3) *E. coli* cells were transformed[32] with pEHISTEV-*SCATT20080*, which has been cloned previously (peptide sequence of the 5-FDRPi is mentioned in Appendix 3).[15] The over-expression and purification procedure for 5-FDRPi was adapted from that used to isolate *S. coelicolor* 5-FDRPi (SCO3014 gene product).[15] The protocol used to purify *S. cattleya* 5-FDRPi was a modified version of the previously reported procedure.[25] A single colony of *E. coli* was picked and used to inoculate 10 ml of LB supplemented with 1 % glucose and 50 µg/ml kanamycin. This culture was incubated at 37 ºC, shaking at 180 rpm for 12 h and used to inoculate 2 L baffled flasks filled with 400 ml of auto-induction media prepared using a mixture of ZY media (372 ml), 1 mM MgSO4 (final concentration), 5052 solution (10 ml of 50X), NPS (20 ml of 20X) and kanamycin (100 µg/ml) according to a method described by Studier.[35] The cultures were grown at 20 ºC, shaking at 180 rpm, for 36 h following which the cells were harvested by centrifugation at 3200 rpm for 30 min. The cell pellet was re-suspended in lysis buffer (20 mM Tris-HCl pH 7.5 and 150 mM NaCl; 250 ml) supplemented with Deoxyribonuclease I from bovine pancreas (Sigma; 0.1 mg/ml) and protease inhibitor cocktail (1 x Complete Mini–EDTA free®, Roche). Then the cells were lysed using a one shot cell disrupter (Constant systems™, UK) at 30 kPsi, 4 ºC. The lysate was centrifuged (15000 rpm, 20 min) to remove the cell debris and the soluble fraction was filtered (0.45 µm syringe filter) prior to loading onto a prepacked Ni-Sepharose High Performance His-Trap (GE Healthcare; 5 ml) column pre equilibrated with wash buffer containing 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM β-ME and 40 mM imidazole pH 7.5. The column was washed with wash buffer several times and Ni-bound 5-FDRPi was eluted using elution buffer containing 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM β-ME and 400 mM imidazole pH 7.5. An SDS-PAGE gel was run to analyse the fractions obtained from the Ni-NTA column. The eluted fractions containing FDRPi were pooled and diluted with lysis buffer and TEV protease (recombinant protein expressed in *E. coli* and kindly donated from Professor J. Naismith research group) at 1:20 ratio (TEV: protein) was added This protein cocktail was dialysed against 5 L of 20 mM Tris-HCl pH 7.5, 150 mM NaCl at 4 ºC for 16 h. The progress of the TEV cleavage was monitored by SDS-PAGE gel and

compared against uncut 5-FDRPi. The dialysed 5-FDRPi and TEV cocktail was loaded onto a pre-packed Ni-Sepharose High Performance His-Trap (GE Healthcare; 5 ml) in order to separate the 5-FDRPi (His_6-tag removed) from TEV and cleaved His_6-tags. The flow-through from this Ni column was collected and analysed by SDS-PAGE to reveal contain non His_6-tagged 5-FDRPi. The flow-through was pooled and concentrated to ~7 ml using a centrifugal filter unit (30000 Da, GE Healthcare) and loaded to a Superdex S-200 (HR 16/60;GE Healthcare) gel filtration column connected to a FPLC (ÅKTA express; Amersham Biosciences), equilibrated with 20 mM Tris pH 7.5, 150 mM NaCl and 1 mM TCEP. The fractions containing 5-FDRPi were pooled and concentrated using centrifugal filter unit (30000 Da, GE Healthcare) to approximately 18 mg/ml for assays and crystallisation. For ITC studies, 5-FDRPi was dialysed in to 20 mM phosphate pH 7.5, 150 mM NaCl and concentrated using centrifugal filter unit (30000 Da, GE Healthcare). The protein concentration was measured by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc.) and the identity of the protein was confirmed by MALDI-TOF.

4.8.2 NMR assays to test for activity of 5-FDRPi

A solution of synthetic FDA **23** (2 mg) in 20 mM phosphate buffer pH 7.5 (1 ml) was incubated with immobilised PNP beads at 37 ºC, shaking at 80 rpm for 12 h. The beads were filtered and the supernatant was freeze-dried and re-suspended in D_2O and analysed by ^{19}F $\{^1H\}$ NMR. The reaction was then incubated with 5-FDRPi (3 mg) for 6 h at 37 ºC. The reaction was heat treated at 95 ºC for 10 min to precipitate the protein and the solution was centrifuged at 13000 rpm for 10 min. The supernatant was removed and freeze-dried and re-suspended in D_2O and analysed by ^{19}F $\{^1H\}$ NMR.

The 5-FDRP **24** produced by PNP biotransformation, was used in an isomerase reaction to assay the purified enzyme. 5-FDRP **24** (1 mg) was incubated with purified FDRPi (3 mg) in 20 mM phosphate pH 7.5 at 37 ºC for 6 h. This reaction was heat treated at 95 ºC for 10 min followed by centrifugation (13000 rpm for 10 min) and the supernatant was mixed with D_2O and analysed by ^{19}F $\{^1H\}$ NMR.

The ^{19}F $\{^1H\}$ NMR spectra was recorded on a Bruker AVIII-HD 500 MHz with a SmartProbe BBFO+ and all experiments were conducted at 298K.

4.8.3 Small scale enzymatic preparation and purification of 5-FDRP 24

A solution of synthetic 5′-FDA **23** (6 mg) was prepared in 20 mM phosphate buffer pH 7.5 (4 ml) and incubated with immobilised PNP (GSK) at 37 ºC shaking at 80 rpm over 5-6 days. The progress of the reaction was monitored by ^{19}F-NMR. The sample was filtered and (the pH of which lowered to 3.5 using formic acid. A weak anion exchange cartridge (SOLA WAX, Thermo Scientific) was conditioned with methanol and equilibrated with water (the pH of which lowered to 3.5 with formic acid). The sample was loaded and the column was washed with water (pH 3.5). Another wash was performed with methanol pH lowered to 3.5 using formic acid. The column was then eluted with 50% methanol (the pH of which was increased to 9 using ammonium hydroxide). All fractions were analysed by ^{19}F {^{1}H} NMR and the two fractions containing 5-FDRP **24** (wash 2 and elution 1) were pooled, pH neutralised and freeze-dried. The freeze-dried sample was analysed by ^{1}H and ^{13}C NMR to confirm the purity of **24**.

4.8.4 Isothermal titration calorimetry (ITC) mediated binding assays

ITC was performed using MicroCal VP-ITC calorimeter (Malvern). The 5-FDRPi was dialysed against 20 mM phosphate buffer pH 7.5 and concentrated to around 10 mg/ml. The same dialysis buffer was used to equilibrate the instrument. Dialysis buffer was used to prepare a stock solution of 5′-fluororibosephosphonate **55** (5 mM) and this stock solution was used to prepare 1 ml of 1 mM solutions. The 5-FDRPi was diluted in dialysis buffer to prepare a 2 ml solution at a concentration of 60 µM or 100 µM. All solutions were pH corrected to ± 0.05 and degassed at 25 ºC. The ligand was titrated into the 5-FDRPi solution at 25 ºC/ 15 ºC. The first injection volume was 2 µl, and was followed by multiple injections of 5 µl. Data were analysed by MicroCal Origin software.

4.8.5 Crystallisation of 5-FDRPi

All crystal trials were set up with an Arts Robbins Gryphon at protein to precipitant ratios of 2:1 and 1:1. Initial screenings were carried out using commercial screens in a 96-well format (JSCG+ and Wizard 1+2; Molecular dimensions and PEGs I and II suite; Qiagen) prepared with 14 mg/ml and 7.5 mg/ml of purified 5-FDRPi using a sitting drop vapour diffusion technique at 20 ºC and 4 ºC (Table 1). Conditions that produced crystals were further optimised by varying the PEG concentrations, additive, buffer concentrations and the pH as mentioned in Table 2. Selected conditions which produced crystals were set up again with high FDRPi concentration (16 mg/ml) from which E1 well condition of G1 optimisation was chosen and further optimised (Table 4).

Part of the crystals formed were retrieved and flash-frozen in cryoprotectant containing precipitant supplemented with 35% glycerol. The crystals were shot in-house at 100 K on a Rigaku 007 HFM rotating anode X-ray generator with a Saturn 944 CCD detector. A dataset of 236 frames of 0.5° oscillation with a crystal-to-detector distance of 60 mm and 240 sec exposure per frame was collected.

Table 4: Further optimisation was carried out on G1 well (PEGs II screen).

Optimisation Screen ; Well in which crystals were observed	**Contents of the well**	**Concentration range for optimisation**
G1 screen; E1	24% w/v PEG 5000 MME and 0.1 M Tris-HCl pH 7.6	24-27% w/v PEG 5000 MME, 0- 0.2 M $MgSO_4$ and 0.1 M Tris-HCl pH 7.5-7.7.

4.9 Chapter 4 References

1 E. Albers, *IUBMB Life*, 2009, **61**, 1132–1142.

2 M. Sauter, *PLANT Physiol.*, 2004, **136**, 4061–4071.

3 N. M. Kredich and G. M. Tomkins, *J. Biol. Chem.*, 1966, **241**, 4955–4965.

4 A. F. Neuwald, B. R. Krishnan, I. Brikun, S. Kulakauskas, K. Suziedelis, T. Tomcsanyi, T. S. Leyh and D. E. Berg, *J. Bacteriol.*, 1992, **174**, 415–425.

5 Y. Dai, T. C. Pochapsky and R. H. Abeles, *Biochemistry*, 2001, **40**, 6379–6387.

6 J. W. Wray and R. H. Abeles, *J. Biol. Chem.*, 1995, **270**, 3147–3153.

7 A. Sekowska and A. Danchin, *BMC Microbiol.*, 2002, **2**, 1–8.

8 S. V. Rieder and I. Rose, *J. Biol. Chem.*, 1959, **234**, 1007–1010.

9 J. M. Berrisford, A. M. Hounslow, J. Akerboom, W. R. Hagen, S. J. J. Brouns, J. van der Oost, I. A. Murray, G. Michael Blackburn, J. P. Waltho, D. W. Rice and P. J. Baker, *J. Mol. Biol.*, 2006, **358**, 1353–1366.

10 Y. J. Topper, *J. Biol. Chem.*, 1957, **225**, 419–425.

11 A. Lavie, K. N. Allen, G. A. Petsko and D. Ringe, *Biochemistry*, 1994, **33**, 5469–5480.

12 T. D. Fenn, D. Ringe and G. A. Petsko, *Biochemistry*, 2004, **43**, 6464–6474.

13 M. Onega, *PhD thesis, Studies and application of the enzymes of fluorometabolite biosynthesis in Streptomyces cattleya*, University of St Andrews, 2009.

14 M. Onega, R. P. McGlinchey, H. Deng, J. T. G. Hamilton and D. O'Hagan, *Bioorg. Chem.*, 2007, **35**, 375–385.

15 H. Deng, S. M. Cross, R. P. McGlinchey, J. T. G. Hamilton and D. O'Hagan, *Chem. Biol.*, 2008, **15**, 1268–1276.

16 V. Barbe, M. Bouzon, S. Mangenot, B. Badet, J. Poulain, B. Segurens, D. Vallenet, P. Marliere and J. Weissenbach, *J. Bacteriol.*, 2011, **193**, 5055–5056.

17 M. Bumann, S. Djafarzadeh, A. E. Oberholzer, P. Bigler, M. Altmann, H. Trachsel and U. Baumann, *J. Biol. Chem.*, 2004, **279**, 37087–37094.

18 M. Land, L. Hauser, S.-R. Jun, I. Nookaew, M. R. Leuze, T.-H. Ahn, T. Karpinets, O. Lund, G. Kora, T. Wassenaar, S. Poudel and D. W. Ussery, *Funct. Integr. Genomics*, 2015.

19 B. L. Webb and C. G. Proud, *Int. J. Biochem. Cell Biol.*, 1997, **29**, 1127–1131.

20 H. Tamura, Y. Saito, H. Ashida, T. Inoue, Y. Kai, A. Yokota and H. Matsumura, *Protein Sci.*, 2007, **17**, 126–135.

21 E. B. Shirling and D. Gottlieb, *Int. J. Syst. Bacteriol.*, 1969, **19**, 391–512.

22 J. S. Kahan, F. M. Kahan, R. Goegelman, S. A. Currie, M. Jackson, E. O. Stapley, T. W. Miller, A. K. Miller, D. Hendlin, S. Mochales, S. Hernandez, H. B. Woodruff and J. Birnbaum, *J. Antibiot. (Tokyo).*, 1979, **32**, 1–12.

23 D. M. Klingeman, S. Utturkar, T.-Y. S. Lu, C. W. Schadt, D. A. Pelletier and S. D. Brown, *Genome Announc.*, 2015, **3**, e01344–15.

24 Y. Saito, H. Ashida, C. Kojima, H. Tamura, H. Matsumura, Y. Kai and A. Yokota, *Biosci. Biotechnol. Biochem.*, 2007, **71**, 2021–2028.

25 Jason Chan, *PhD thesis, Identification and enzyme studies of rare amino acid biosynthesis from Streptomyces cattleya*, University of St Andrews, 2014.

26 B. Rost and C. Sander, *J. Mol. Biol.*, 1993, **232**, 584–599.

27 P. Tompa, *Trends Biochem. Sci.*, 2002, **27**, 527–533.

28 V. N. Uversky, *Protein Sci.*, 2002, **11**, 739–756.

29 P. Nasomjai, D. O'Hagan and A. M. Z. Slawin, *Beilstein J. Org. Chem.*, 2009, **5**, 1–7.

30 T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood, *Practical Streptomyces Genetics*, John Innes Foundation, Norwich, 2000.

31 J. F. Sambrook and D. W. Russell, Cold Spring Harbor Laboratory Press, New York, 4th edn., 2001, 26–28.

32 H. E. Bergmans, I. M. van Die and W. P. Hoekstra, *J. Bacteriol.*, 1981, **146**, 564–570.

33 H. Liu and J. H. Naismith, *Protein Expr. Purif.*, 2009, **63**, 102–111.

34 T. D. Ashton and P. J. Scammells, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 3361–3363.

35 A. Sivashanmugam, V. Murray, C. Cui, Y. Zhang, J. Wang and Q. Li, *Protein Sci.*, 2009, **18**, 936–948.

36 H. J. Imker, A. A. Fedorov, E. V. Fedorov, S. C. Almo and J. A. Gerlt, *Biochemistry*, 2007, **46**, 4077–4089.

5. Thesis conclusion

5.1 Thesis conclusion and future work

This thesis is focused on enzymes involved in the fluorometabolite biosynthetic pathway in *S.cattleya* (Figure 1). The fluorinase enzyme catalyses the first step of this pathway between SAM **22** and inorganic fluoride to produce 5′-FDA **23** and L-methionine **27**. Fluorinase has found applications in PET due to its ability to catalyse C–^{18}F bond formation in the presence of [^{18}F]F$^-$ which can be used to incorporate ^{18}F into substrates for the synthesis of novel radiotracers. **Chapter 2** of this thesis describes the preparation of a user-friendly, shelf-stable, freeze-dried enzyme kit for PET, which couples fluorinase to a *Tv*NH enzyme allowing efficient production of [^{18}F]-FDR ([^{18}F]-**37**). The enzymatically synthesised [^{18}F]-**37** was then studied for its potential as a PET radiotracer in tumour induced mice relative to [^{18}F]FDG ([^{18}F]-**32**). [^{18}F] 5-FDR ([^{18}F]-**37**) was observed to have high early influx rates into the tumour and was stable for around 20 min, after which the efflux rate dominated the influx rate, thus reducing the image resolution. In comparison [^{18}F]-**32** uptake increased over the duration of the scan, consistent with the metabolic trapping of [^{18}F]-**32** due to intracellular phosphorylation. Overall, [^{18}F] 5-FDR ([^{18}F]-**37**) was found to provide good image contrast up to 20 min post injection. The stability of the freeze dried enzymes was also assessed over a period of 8 weeks. Shelf stability during storage is an important requirement for producing an enzyme kit. The freeze-dried enzymes were found to exhibit good stability with a modest loss of around 20% activity over 8 weeks when stored at -20 ºC.

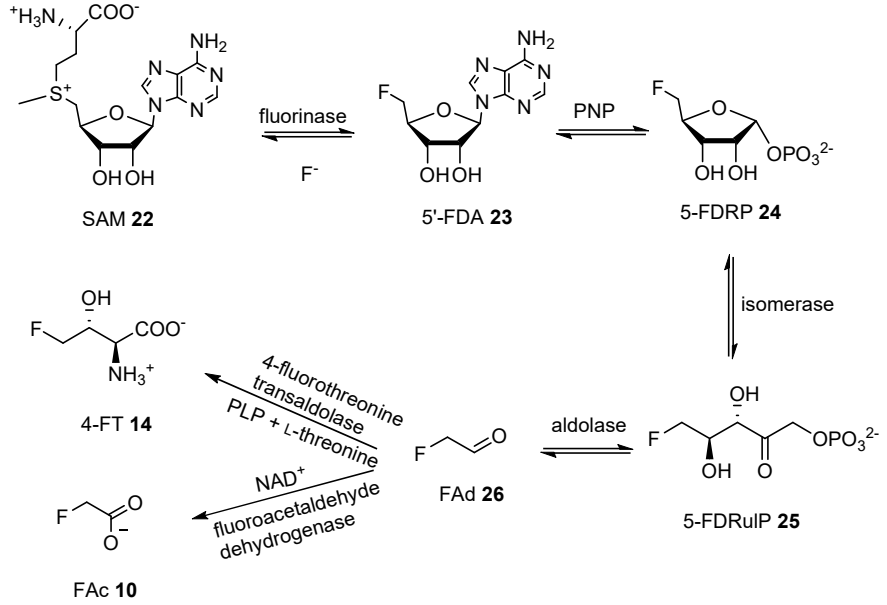


Figure 1: Scheme of the fluorometabolite biosynthetic pathway in *S.cattleya.*

As future work it would be beneficial to evaluate and optimise the protein purification methods used in the isolation of fluorinase and *Tv*NH by monitoring the specific activity and yield of each protein at every step of the purification.[1] The assessment of the specific activity of each enzyme from the cell pellet to the lysate, post affinity and size exclusion chromatography would allow for quantification of the efficiency of the purification method, which will aid in optimising the cost effectiveness of producing the enzyme kit. In addition, it would be valuable to record the activity loss over time for aqueous fluorinase and *Tv*NH stored at either -20 ºC or 4 ºC in relation to the freeze dried enzymes stored at the same temperatures, to investigate the significance of the freeze drying step on enzyme stability.

Following publication of the work above[2], Keng and co-workers reported the application of fluorinase immobilised on a functional polymer monolith to catalyse C-F bond formation in an aqueous environment. This methodology was used to produce [^{18}F]-FDA ([^{18}F]-**23**) from SAM with a 68% yield in 2.5 h. The yield obtained *via* polymer immobilised fluorinase was much lower than the yields reported (over 98%) by using freeze-dried fluorinase produce [^{18}F]-**23**. In addition, the reaction time reported to generate [^{18}F]-**23** using freeze-dried fluorinase was 20-30 min, significantly shorter than the 2.5 h reaction time reported for polymer immobilised fluorinase. Sergeev *et.al* demonstrated that the immobilised fluorinase could be recycled up to four times with an increase in yield up to 80%, however this also increased the reaction time considerably[3] demonstrating that using freeze-dried fluorinase still remains to be the most efficient method in generating [^{18}F]-**23** with high yields.

Most recent work on the applications of fluorinase explored the substrate tolerance of fluorinase to incorporate ^{18}F into peptides for PET.[4–6] Transhalogenation reactions were used to generate ^{18}F labelled 5'-fluoro-5'-deoxy-2-ethynyladenosine linked to a monomeric arginyl-glycyl-aspartic acid (RGD) peptide through a PEG linker at C-2 of the adenine base.[4,6] RGD peptides are known to bind upregulated $\alpha\nu\beta_3$ integrin motifs on the surface of cancer cells[7]. The ^{18}F labelled RGD peptide motifs were demonstrated to have high affinity to these integrins.[8–10] Furthermore, fluorinase was successful in ^{18}F labelling larger, more complex dendritic scaffolds carrying two to four cyclic RGD peptide motifs for clinical imaging.[5] [^{18}F]-FDR ([^{18}F]-**37**) has been explored as an efficient bioconjugation ligand for radiolabelling RGD peptides. FDR **37** was demonstrated to be a superior bioconjugation agent in comparison to FDG **32** as the fluorine positioning on the ribose ring promotes ring opening to form the reactive aldehyde from the sugar, allowing for oxime ligations to functionalised RGD peptides (Figure 2) to generate radiolabelled peptides. [11]

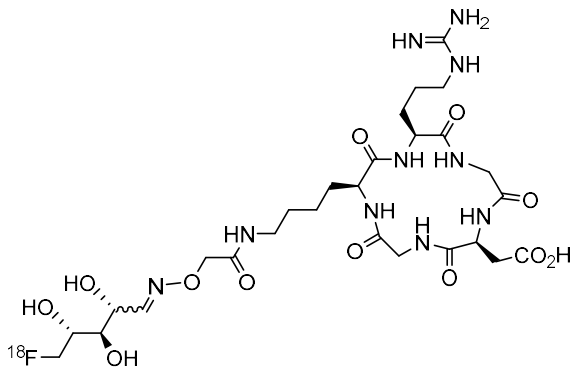


Figure 2: The aminooxy ether functionalised RGD peptide bioconjugated to [^{18}F] FDR **37** to produce radiolabelled RGD peptides for PET.[11]

The *S. cattleya* fluorinase was the only known fluorinase until recently, however four more fluorinases were identified by gene mining in the last four years. **Chapter 3** discusses the cloning, over-expression and purification of *Streptomyces* sp. MA37 fluorinase, along with a subsequent investigation of the kinetic parameters of all four fluorinases relative to the *S.cattleya* fluorinase. The structure of the *Streptomyces* sp. MA37 fluorinase was solved by X-ray crystallography and was observed to be almost identical to that of the *S.cattleya* fluorinase. Two other *flA* genes were identified from *N. brasiliensis* and *Actinoplanes* sp. N902-109 which were found to be fluorinases possessing higher catalytic efficiency relative to both *S. cattleya* and *Streptomyces* sp. MA37. Following the identification of these three fluorinases, an additional fluorinase gene was identified from a marine *Streptomyces xinghaiensis*. This organism was observed to be a fluoroacetate producer in sea salt dependent media.[12] The *flA4* emerged as the most efficient fluorinase with an almost 8 fold higher efficiency than the *N. brasiliensis* and *Actinoplanes* sp. N902-109 fluorinases and an almost 12 fold higher efficiency than *S.cattleya* fluorinase.[13]

Following the discovery of *Streptomyces* sp. MA37 fluorinase, Ang and co-workers have used this enzyme in directed evolution for improving the fluorination efficiency in the transhalogenation reaction to convert ClDA **28** to 5′-FDA **23** *via* SAM **22** intermediate for application in PET. Evolved FlA1 mutants A279Y and F213Y was capable of catalysing successful radiosynthesis of [^{18}F]FDA ([^{18}F]-**23**) with an overall radiochemical conversion over 3 fold higher than the wild type FlA1.[14] This mutant fluorinase has prospects for catalysing more efficient transhalogenation reactions to produce a variety of ^{18}F labelled radiotracers for PET imaging.

Chapter 4 describes the successful crystallisation of 5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (FDRPi), an aldose-ketose isomerase involved in the biosynthetic pathway of fluorometabolite biosynthesis in *S.cattleya*. The FDRPi was over-expressed, purified and assayed for activity. In an attempt to crystallise FDRPi with a cyclic sugar in order to explore the key residues for binding and to investigate the catalytic mechanism of FDRPi, a non-hydrolysable analogue of substrate, 5′-fluororibosephosphonate **55** was synthesised (by P. Nasomjai).[15] Binding studies performed *via* ITC showed 5′-fluororibosephosphonate **55** did not bind FDRPi. Therefore attempts to crystallise **55** bound FDRPi were discontinued, however apo- FDRPi crystals were obtained and the structure of FDRPi was resolved. The FDRPi protein structures was observed to be similar to that of *B.subtilis* MTRPi, the closest representative of 5-FDRPi of which the crystal structure has been previously resolved.[16] The active sites of both *B.subtilis* MTRPi and *S.cattleya* FDRPi possessed similar arrangements despite a loop rearrangement in *S.cattleya* FDRPi, which may be important in an interaction with the terminal fluorine of 5-FDRP **24**. Probing the catalytic mechanism of FDRPi is yet to be accomplished and may require the design of new inhibitors capable of binding FDRPi for crystallographic analysis.

Overall this thesis has evaluated the prospects of using the fluorinase to generate [^{18}F]-FDR ([^{18}F]-**37**) for PET imaging of tumours; characterised a more efficient fluorinase from *Streptomyces* sp. MA37 (through crystallography and kinetic evaluation) which may have applications in future PET studies; and explored the crystallisation of *S. cattleya* FDRPi revealing the arrangement of residues in the active site of FDRPi, which could be further explored for insight into its catalytic mechanism. The information obtained from the work carried out in this thesis has broadened the knowledge on the enzymes related to fluorometabolite biosynthesis in *S. cattleya* and *Streptomyces* sp. MA37.

5.2 Chapter 5 References

1 J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry*, New York, 5th edn., 2002.

2 S. Dall'Angelo, N. Bandaranayaka, A. D. Windhorst, D. J. Vugts, D. van der Born, M. Onega, L. F. Schweiger, M. Zanda and D. O'Hagan, *Nucl. Med. Biol.*, 2013, **40**, 464–470.

3 M. E. Sergeev, F. Morgia, M. R. Javed, M. Doi and P. Y. Keng, *J. Mol. Catal. B Enzym.*, 2013, **92**, 51–56.

4 S. Thompson, Q. Zhang, M. Onega, S. McMahon, I. Fleming, S. Ashworth, J. H. Naismith, J. Passchier and D. O'Hagan, *Angew. Chemie Int. Ed.*, 2014, **53**, 8913–8918.

5 S. Thompson, I. N. Fleming and D. O'Hagan, *Org. Biomol. Chem.*, 2016, **14**, 3120–3129.

6 S. Thompson, M. Onega, S. Ashworth, I. N. Fleming, J. Passchier and D. O'Hagan, *Chem. Commun.*, 2015, **51**, 13542–13545.

7 E. Ruoslahti, *Annu. Rev. Cell Dev. Biol.*, 1996, **12**, 697–715.

8 A. J. Beer, R. Haubner, I. Wolf, M. Goebel, S. Luderschmidt, M. Niemeyer, A.-L. Grosu, M.-J. Martinez, H. J. Wester, W. A. Weber and M. Schwaiger, *J. Nucl. Med.*, 2006, **47**, 763–769.

9 C. Yu, D. Pan, B. Mi, Y. Xu, L. Lang, G. Niu, M. Yang, W. Wan and X. Chen, *Eur. J. Nucl. Med. Mol. Imaging*, 2015, **42**, 2021–2028.

10 H. Chen, G. Niu, H. Wu and X. Chen, *Theranostics*, 2016, **6**, 78–92.

11 S. Dall'Angelo, Q. Zhang, I. N. Fleming, M. Piras, L. F. Schweiger, D. O'Hagan and M. Zanda, *Org. Biomol. Chem.*, 2013, **11**, 4551–4558.

12 S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan and H. Deng, *Org. Biomol. Chem.*, 2014, **12**, 4828–4831.

13 L. Ma, Y. Li, L. Meng, H. Deng, Y. Li, Q. Zhang and A. Diao, *RSC Adv.*, 2016, **6**, 27047–27051.

14 H. Sun, W. L. Yeo, Y. H. Lim, X. Chew, D. J. Smith, B. Xue, K. P. Chan, R. C. Robinson, E. G. Robins, H. Zhao and E. L. Ang, *Angew. Chemie Int. Ed.*, 2016, **55**, 14277–14280.

15 P. Nasomjai, D. O'Hagan and A. M. Z. Slawin, *Beilstein J. Org. Chem.*, 2009, **5**, 1–7.

16 H. Tamura, Y. Saito, H. Ashida, T. Inoue, Y. Kai, A. Yokota and H. Matsumura, *Protein Sci.*, 2007, **17**, 126–135.

Appendix 1

Crystallographic Data for *Streptomyces sp.* MA37 fluorinase (PDB: 5LMZ)

Crystallization experiments:

Method: 7% *w/v* PEG 6000, 0.5 M NaCl; Vapor diffusion sitting drop

Temperature = 293.0 K

Crystal data: cubic

Unit cell	
Length (Å)	Angle (º)
A = 125.92	α = 90
B = 125.92	β = 90
C =125.92	ϒ = 90

Symmetry:

Space Group: P 2_1 3

Diffraction:

Detector = CCD, Rigaku Saturn 944

Temperature = 100 K

Protocol: single wavelength

Data collection:

Resolution (high) =2.55 Å

Resolution (low) =18.17 Å

No. of reflections observed= 21982

Refinement:

R-Value work = 0.166

R-value free = 0.216

Appendix 2

Crystallographic Data for *Streptomyces cattleya* isomerase (FDRPi) (unpublished)

Crystallization experiments:

Method: 25% *w/v* PEG 5K MME, 0.1 M Tris-HCl pH 7.6; Vapor diffusion sitting drop

Temperature = 293.0 K

Crystal data: monoclinic

Unit cell	
Length (Å)	Angle (º)
A = 60.73	α = 90
B = 59.95	β = 108
C =95.84	ϒ = 90

Symmetry:

Space Group: P 2_1

Diffraction:

Detector = CCD, Rigaku Saturn 944

Temperature = 100 K

Protocol: single wavelength

Data collection:

Resolution (high) =1.67 Å

Resolution (low) =57.61 Å

No. of reflections observed= 75438

Refinement:

R-Value work = 0.185

R-value free =0.212

Appendix 3

fIA1 nucleotide sequence

Streptomyces sp. MA37 gene for fluorinase

GTGGCTGCAAACGGCTCTCAGCGCCCCATCATCGCCTTCATGAGCGACCTCGGCAC
CACGGACGACTCGGTGGCCCAGTGCAAGGGCCTCATGCACAGCATCTGCCCCGGCG
TCACGGTGGTGGACGTCTGCCACTCCATGACGCCGTGGGACGTCGAGGAGGGTGCG
CGCTACATCGTCGACCTCCCCCGCTTCTTCCCCGAGGGCACGGTGTTCGCCACACCA
CCTATCCCGCGACCGGCACCACGACCCGCTCCGTGGCCGTGCGGATCCGGCAGGC
GGCCAAGGGCGGCGCCCGCGGCCAGTGGGCCGGCTCCGGCGACGGCTTCGAGCG
GGCCGACGGCTCGTACATCTACATCGCGCCCAACAACGGCCTGCTGACCACCGTGC
TCGAGGAGCACGGATACATCGAGGCCTACGAGGTCACGTCCACCAAGGTCATCCCC
GCGAACCCCGAACCGACCTTCTACAGCCGCGAGATGGTGGCCATCCCCTCCGCCCA
CCTCGCCGCCGGCTTCCCGCTGGCCGAGGTCGGCCGCCGGCTCGACGACAGCGAG
ATCGTCCGCTTCCACCGGCCCGCGGTCGAGATCTCCGGGGAGGCGCTCTCCGGCGT
CGTCACCGCGATCGACCACCCCTTCGGCAACATCTGGACCAACATCCACCGCACGG
ACCTGGAGAAGGCCGGCATCGGCCAGGGGAAGCACCTGAAGATCATCCTCGACGAC
GTCCTGCCGTTCGAGGCCCCGCTCACCCCGACCTTCGCGGACGCCGGGGCAATCG
GCAACATCGCCTTCTACCTCAACAGCCGCGGCTACCTGTCCCTCGCCCGCAACGCC
GCCAGCCTCGCCTACCCCTACAACCTCAAGGCCGGTCTGAAGGTCCGCGTGGAAGC
GCGCTGA

FIA1 protein sequence

Streptomyces sp. MA37 peptide sequence for fluorinase

GenBank: CDH39444.1

MW= 32 kDa

MAANGSQRPIIAFMSDLGTTDDSVAQCKGLMHSICPGVTVVDVCHSMTPWDVEEGARYI
VDLPRFFPEGTVFATTTYPATGTTTRSVAVRIRQAAKGGARGQWAGSGDGFERADGSYI
YIAPNNGLLTTVLEEHGYIEAYEVTSTKVIPANPEPTFYSREMVAIPSAHLAAGFPLAEVG
RRLDDSEIVRFHRPAVEISGEALSGVVTAIDHPFGNIWTNIHRTDLEKAGIGQGKHLKIILD
DVLPFEAPLTPTFADAGAIGNIAFYLNSRGYLSLARNAASLAYPYNLKAGLKVRVEAR

5-deoxy-5-fluoro-D-ribose 1-phosphate isomerase (FDRPi) peptide sequence

Streptomyces cattleya peptide sequence for FDRPi

GenBank: AEW94379.1

MW= 39 kDa

MGDQSVQPLAKGTGSGTPEPKPALRWEEPPEGPVLVLLDQTRLPVEEVELFCTDVPAL
VQAIRTLAVRGAPLLGLAGAYGVALAAARGYDVGQAADELAGARPTAVNLSYGVRRALA
AYRTAVTGGADDTGAAAATLAEARALHAEDARASERMARNGLALLDELVPGGGYRVLT
HCNTGALVSGGEGTALAVVLAAHRGGLLRRLWVDETRPLLQGARLTAYEAARAGVAHT
LLPDGAAGSLFAAGEVDAVLIGADRIAADGSTANKVGSYPLAVLARYHNVPFVVVAPTTTI
DLATPDGTAIEVEQRPAQEVTELTGPRPGPDREGATGIPVAPLGTPAYNPAFDVTPPELI
TAVVTETGVASPVTGSSIAALAARPGPVRAQP

Appendix 4

2D NMR analysis for FDRPi

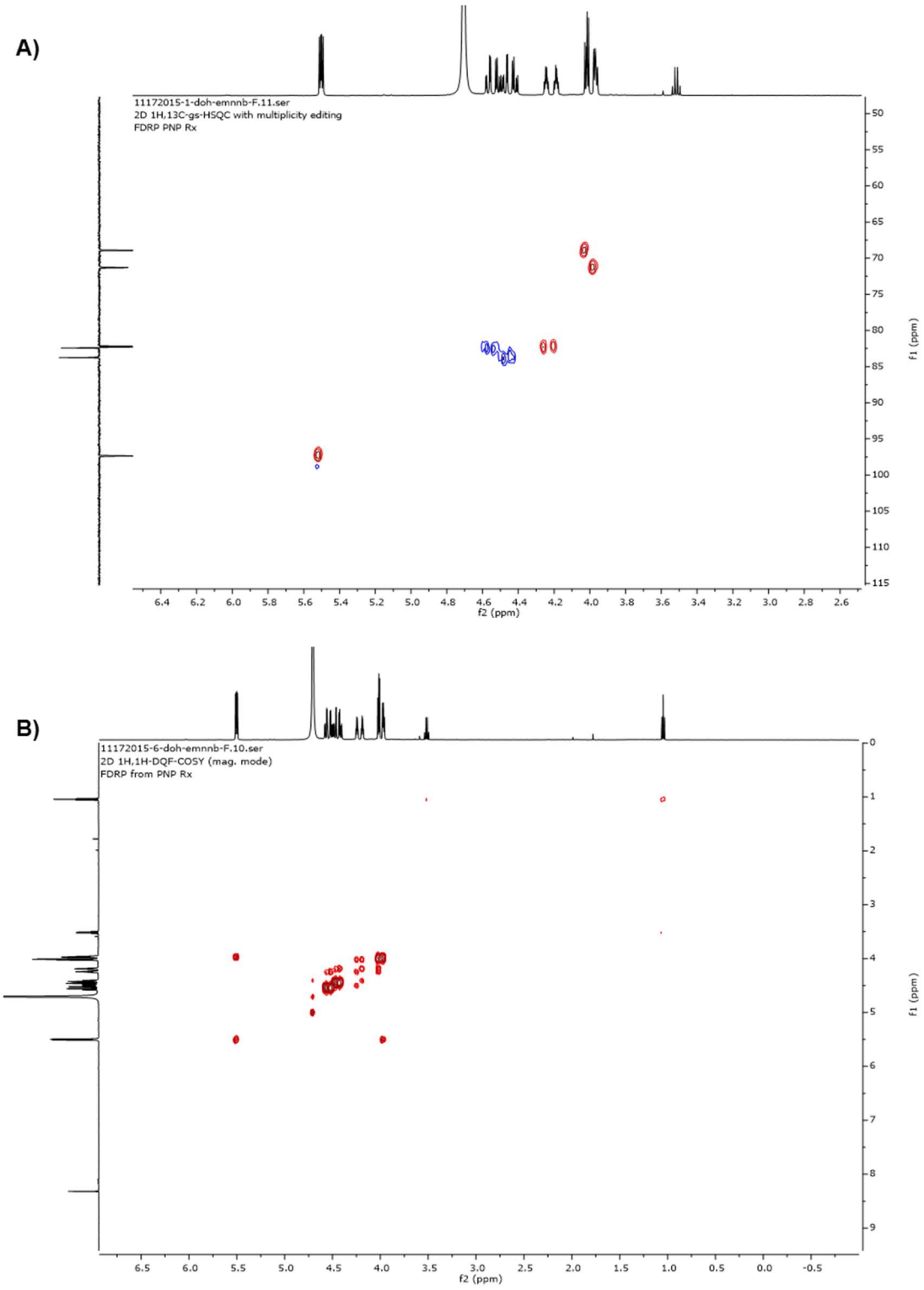


Figure 1: **A)** The ^{1}H, ^{13}C-HSQC 2D spectrum of 5-FDRP **24** in D_2O. **B)** The ^{1}H,^{1}H–COSY 2D spectrum of 5-FDRP **24** in D_2O.