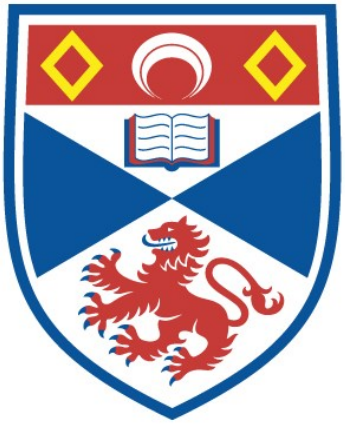
IDENTIFICATION AND METABOLISM STUDIES OF FLUOROMETABOLITES FROM DIFFERENT STREPTOMYCES

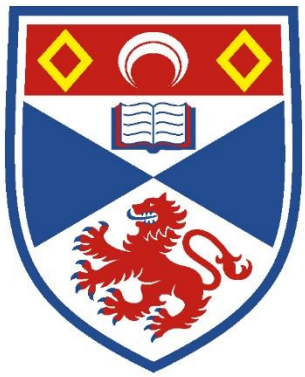
Axel Bartholomé

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



2017

Identification and metabolism studies of fluorometabolites from different *Streptomyces*



University of
St Andrews

Axel Bartholomé

Supervisor: Prof. David O'Hagan

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

March 2017

Declarations

1. Candidate's declarations:

I, Axel Bartholomé, hereby certify that this thesis, which is approximately 40,000 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 September, 2013 and as a candidate for the degree of Doctor of Philosophy in September, 2014; the higher study for which this is a record was carried out in the University of St Andrews between 2013 and 2017.

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

Date signature of candidate signature of supervisor

Abstract

To date, only five fluorinated natural products have been identified. These were isolated from both plants and bacteria. The bacterium *Streptomyces cattleya* has the ability to biosynthesise fluoroacetate and 4-fluoro-L-threonine. The first enzyme discovered to be capable of catalysing a C-F bond from fluoride ion, the fluorinase, was identified from *S. cattleya* in 2002 and is involved in the first step in the biosynthesis of fluorometabolites. The complete metabolic pathway of fluoroacetate and 4-fluoro-L-threonine in *S. cattleya* was elucidated utilising a variety of different techniques.

Recently, genome studies revealed the presence of four new fluorinase enzymes from different bacterial species. Cultures of one of these species, named *Streptomyces* sp. MA37, showed the production of new unidentified fluorometabolites. Over-expression of the *FdrC* gene from *Streptomyces* sp. MA37 was performed, and enzymatic assays of the FdrC enzyme allowed the conversion of 5-fluoro-5-deoxy-ribose to (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid. Identification of (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid as a new fluorometabolite was then confirmed by synthetic synthesis.

Nucleocidin, an antibiotic containing fluorine, was isolated in 1957 from the soil bacterium, *Streptomyces calvus*. Since its isolation, attempts at re-establishing nucleocidin producing cultures have proven unsuccessful. The biosynthesis of nucleocidin involves a C-F bond-forming enzyme unique to *Streptomyces calvus*. Production of a commercial strain from Pfizer was established and isotopic labelling studies with different labelled glycerols were completed. Pulse feeding experiments with (2*R*)-[1-$^{2}H_{2}$]-glycerol, (2*S*)-[1-$^{2}H_{2}$]-glycerol, glycerol-1,1,2,3,3-d_5 and [2-^{13}C]-glycerol proved to be successful. Concomitantly, synthesis of highly pure putative substrates for the fluorinating enzyme was carried out. Unfortunately, cell-free extract experiments were achieved, but results from these were not conclusive.

Acknowledgements

First of all, I would like to thank my supervisor Professor David O'Hagan, for giving me the opportunity to work as a research student within his group at the University of St Andrews. I would like to express my gratitude for his invaluable advices and patience over the past years, and really thank him for his constant encouragement.

My thanks also go to the past and present members of the DOH group for their constant availability and help. I am very thankful to have been taught organic chemistry by Dr. Stephen Thompson. I am very grateful to Dr. Nouchali Bandaranayaka and Dr. Long Ma who introduced me to biology. I am indebted to Dr. Nawaf Al-Maharik for his advice in chemistry, particularly about safety. I owe special thanks to Dr. Phillip Lowe for proof reading this work. I would also like to thank Tanya, Ricardo, Davide, Rudy, Rodrigo, Flavio, Leonardo, Tony, Kevin, Fatah, Bouchra, Fang and Maria for creating such a good atmosphere in the lab.

Thank you to the people who have made this research possible in the Department. My thanks go to Mrs Caroline Horsburgh for mass spectrometry analyses, Mrs Melanja Smith and Dr. Thomas Lebl for running the NMR facility, as well as Prof. Alexandra Slawin and Dr. David Cordes for X-ray crystal structure determination.

I would like to thank our contacts in Pfizer, Dr. Alessandra Eustaquio, Dr. Jeffrey Janso and Dr. Usa Reilly who have been very helpful by giving *Streptomyces calvus* strains dated from 1956. I owe thanks to Dr. Thomas Shepherd from the James Hutton Institute in Dundee and Dr. Hai Deng from the University of Aberdeen for their collaborations.

I am blessed by the company of many good friends from varied backgrounds back in France and I would like to particularly share my gratitude to a small association named "FIF". I am also very grateful to my family who supported me through my studies in St Andrews. I thank them from the bottom of my heart: Mamy, Didier, Herve, Thomas, Marie, Gandalf, Guapa and

Panda. I am especially grateful to my mother for always being there for me since the beginning. I made it through over the last 4 years thanks to her care.

I wish to acknowledge the French company named Maestria for believing in me and waiting for the end of my PhD to hire me. I consider myself as a very lucky individual for my professional development and I am really looking forward to start working for this company.

Finally, I would like to dedicate my thesis to my grandfather, who passed away during my time in St Andrews. He always encouraged me to start new project, such as this thesis, and work hard and not to give up easily. All the precious interesting conversations with him will forever remain in my mind.

Abbreviations

Ac	acetyl
Acetyl CoA	acetyl coenzyme A
AcOH	acetic acid
ACN	acetonitrile
ACP	acyl carrier protein
AgF	silver fluoride
ATP	adenosine triphosphate
B_{12}	cobalamine
b.p.	boiling point
BAIB	bis(acetoxy)iodobenzene
$BF_3.OEt_2$	borontrifluoride-diethyl ether
$BnEt_3NCl$	benzyltriethylammonium chloride
br	broad
BSTFA	*N*,*O*-bis(trimethylsilyl)trifluoro acetamide
Bz	benzoyl
BzCl	benzoyl chloride
c	concentration
cat.	catalytic
CFE	cell free extract
ClDA	5'-chloro-5'-deoxyadenosine
ClDR	5-chloro-5-deoxyribose
ClDRP	5-chloro-5-deoxyribose-1-phosphate
COSY	correlation spectroscopy
d	doublet
Da	dalton
DAST	(diethylamino)sulfur trifluoride
DCA	dichloroacetic acid
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
dd	doublet of doublets
ddd	doublet of doublet of doublets
dddd	doublet of doublet of doublet of doublets
Deoxofluor®	bis(2-methoxyethyl)aminosulfur trifluoride
DMA	dimethylacetamide
DMAP	dimethylaminopyridine
DMF	dimethylformamide
DMP	Dess-Martin periodinane
DMSO	dimethylsulfoxide
DMSO-d_6	hexadeuteriodimethylsulfoxide
dt	doublet of triplets
E. coli	*Escherichia coli*
eq.	equivalents
ES	electrospray
ESI	electrospray ionisation
Et	ethyl
Et_2O	diethyl ether
Et_3N	triethylamine
EtOAc	ethyl acetate
FAc	fluoroacetate
FAD	flavine adenine dinucleotide (oxidised form)
$FADH_2$	flavine adenine dinucleotide (reduced form)

FDA	5'-fluoro-5'-deoxyadenosine
FDR	5-fluoro-5-deoxyribose
FDRP	5-fluoro-5-deoxyribose-1-phosphate
FHPA	(2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid
FRL	5-fluoro-5-deoxy-lactone
FTIR	Fourier transform infrared
g	grams
h	hour
HCl	hydrochloric acid
HMBC	heteronuclear multiple bond correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence spectroscopy
IBX	2-iodoxybenzoic acid
IMP	inosine monophosphate
^{i}PrOH	isopropanol
IR	infrared
J	coupling constant
k	kilo, 10^3
k_{cat}	turnover number
K_M	Michaelis Menten constant
LC-MS	liquid chromatography-mass spectrometry
$LiAlD_4$	lithium aluminium deuteride
$LiAlH_4$	lithium aluminium hydride
m.p.	melting point
m	multiplet
M	molar
M	molecular ion
Me	methyl
MeOH	methanol
MTBD	7-methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene
MS	mass spectrometry
min	minutes
n-	normal-
n-BuLi	*n*-butyllithium
NAD^+	nicotinamide adenine dinucleotide (oxidised form)
NADH	nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidised form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NBS	*N*-bromosuccinimide
NFSI	*N*-fluorobenzenesulfonimide
NMR	nuclear magnetic resonance
OD	optical density
PE	petroleum ether
Ph	phenyl
PNP	purine nucleoside phosphorylase
ppm	parts per million
PTSA	*para*-toluenesulfonic acid
PyFluor	2-pyridinesulfonyl fluoride

q	quartet
quant.	quantitative
r.t.	room temperature
R_f	retention factor
RP	reverse phase
rpm	revolutions per minute
s	singlet
S.	Streptomyces
SAH	*S*-adenosylhomocysteine
SAM	*S*-adenosyl-L-methionine
Selectfluor	1-chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane
t	triplet
TBAF	tetrabutylammonium fluoride
TBDPS	*tert*-butyldiphenylsilyl
TBDPSCl	*tert*-butyl diphenylchlorosilane
TBS	*tert*-butyldimethylsilyl
TBSCl	*tert*-butyldimehtylsilyl chloride
TEMPO	(2,2,6,6-tetramethylpiperidin-1-yl)oxidanyl
tert-	tertiary-
Tf	trifluoromethanesulfonyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMS	trimethylsilyl
TMSCl	trimethylsilyl chloride
TMSOTf	trimethylsilyl trifluoromethanesulfonate
t_R	retention time
Tris	tris(hydroxymethyl)aminomethane
Ts	tosyl, *para*-toluenesulfonyl
TsCl	tosyl chloride, *para*-toluenesulfonyl chloride
UV	ultraviolet
v/v	volume per volume
ν_{max}	maximal absorption
W	watts
w/v	weight per volume
δ	chemical shift

Contents

1. Introduction

1.1. Metabolism

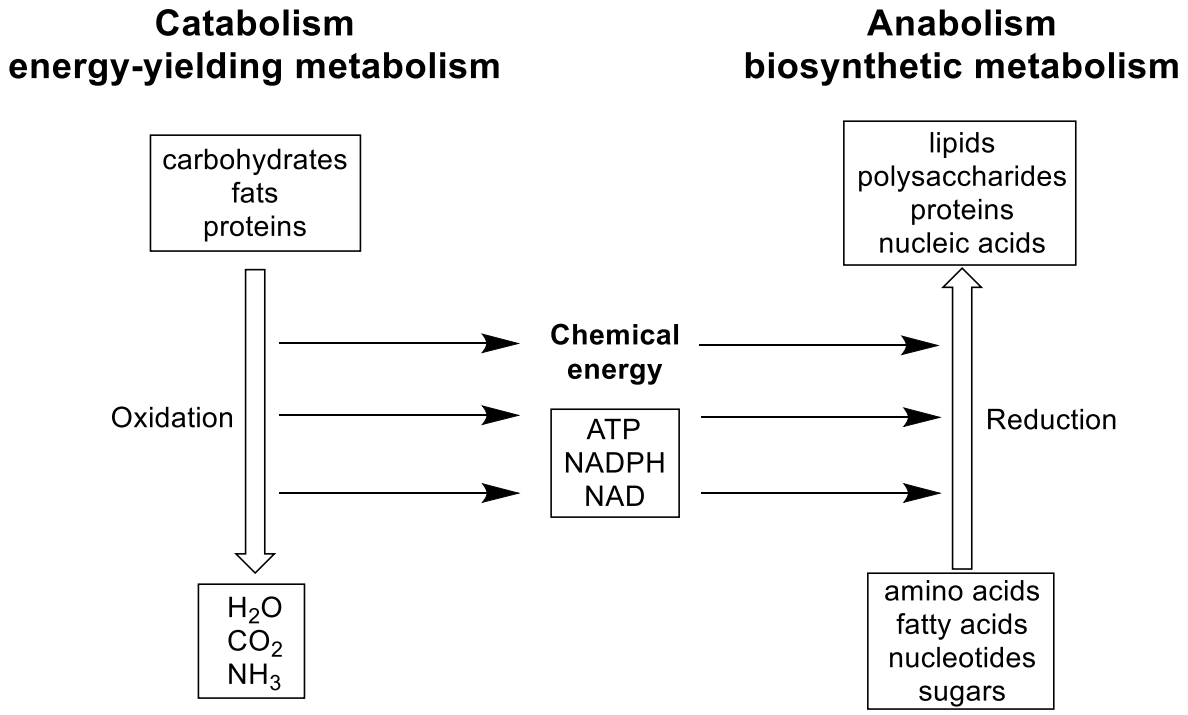
1.1.1. Catabolism and anabolism

The word metabolism comes from the Greek "*metabolē*", which means "change". Metabolism refers to all chemical reactions involved in maintaining the living state of the cells within an organism. Biochemical transformations, which are enzymatically catalysed, allow cells to: grow, duplicate, maintain their structures and adapt to the environment. Enzymes are a crucial component of cellular metabolism, as they are responsible for mediating essential biotransformations, within the cell. They catalyse kinetically favoured reactions, which require energy, by coupling them to spontaneous reactions that release energy.[1] The chemical reactions of metabolism can be organised into metabolic pathways and the regulation of these different pathways can also be carried out by enzymes, often in response to external signals. Some of these metabolic pathways are interconnected, leading to elaborate complexity.[2]

Metabolism is heavily reliant on the availability of nutrients, which are broken down in order to provide the energy essential to run cellular processes.[1] However, the types of nutrients that can be digested vary from one living organism to another. Despite this, a large variety of nutrients are capable of being utilised by most organisms, such as: animals, plants, and bacteria, and these are derived from three types of compounds: amino acids, carbohydrates and lipids. Metabolism is also involved in the elimination of nitrogenous waste such as ammonia, urea and uric acid.[3]

Each biochemical reaction involved in metabolism can be divided into two categories: catabolism and anabolism. Catabolism involves the breakdown of large organic molecules to obtain energy and small components needed by anabolic reactions. These catabolic reactions

are very different for each organism, and depend on many factors such as their energy and carbon sources.[4] Anabolism comprises the biosynthesis of all compounds required by the cell. In order to do this, the cell uses the energy released along with the small precursors made during catabolism. Anabolism contains three major steps; the first involves the biosynthesis of precursors e.g. amino acids, nucleotides, *etc*. The second usually involves the activation by ATP of these precursors and finally the last step involves the assembly of these activated precursors into more complex molecules such as proteins, lipids, polysaccharides and nucleic acids.[5]



Scheme 1. Scheme showing interconnection between catabolism and anabolism.

1.1.2. Primary and secondary metabolites

Metabolites are small molecules that are considered to be either intermediates or products of cell metabolism. Metabolites possess various functions and can be divided into two classes: primary and secondary metabolites. Primary metabolites are fundamental to all cells e.g. DNA, proteins, lipids, *etc*. They are universal and responsible for the vital functions within in a cell

and as such are involved in the generation of energy, growth, cellular reproduction, breakdown of waste products and synthesis of cellular components.[6]

Secondary metabolites are not necessary for the survival of the cell, and are often specific to a particular subspecies or even individual organism. They include polyketides, terpenoids, alkaloids, *etc*. and thus fulfil a variety of different roles within the cell, acting as: antibiotics, toxins, pigments, *etc*.[6] Secondary metabolites are derived from the assembly of building blocks produced by primary metabolism. Primary and secondary metabolisms are closely connected. However, the metabolic pathways of secondary metabolites differ from organism to organism. Consequently, the range of natural products produced in nature is very diverse and these molecules contribute to a wide structural diversity.[7]

1.2. Soil bacteria

1.2.1. A large diversity

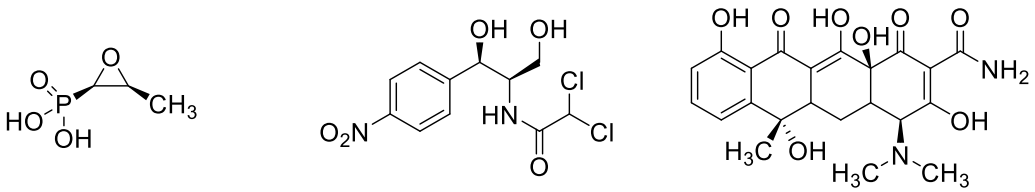
Bacteria are single-celled organisms, typically 1 µm in length. This small size however, is compensated for by the vast number of bacteria in the soil; approximately 40 million bacterial cells can be found in one gram of soil.[8] Moreover, soil is a habitat for the largest distribution of the world's biodiversity. Bacteria are able to reproduce very quickly by binary fission and consume organic matter from their environment in order to make nutrients necessary for this. Such nutrients are also taken up by plants or different organisms. Despite having a comprehensive affiliation, the links between soil bacteria and other organisms, such as plants, are complex and often difficult to interpret.

Soil bacteria can be divided into four different groups, the first of which represents the majority; the decomposer bacteria. They are named as such because they consume simple carbon compounds, for instance, most decomposers can breakdown pesticides and pollutants. The second group are the mutualists, because they form an intimate partnership with plants. Pathogens compose the third group of soil bacteria and they are responsible for the formation

of a gall, which is an abnormal plant growth. The last group contains bacteria that consume compounds of nitrogen, sulfur or hydrogen instead of carbon in order to obtain their energy. All bacteria from these four groups execute important processes related to water dynamics and nutrient cycling.[9,10]

1.2.2. *Streptomyces* bacteria

Most antibiotics in clinical use are isolated from Actinomycetes.[11] The abundance of microbial antibiotics from this source can be explained by the limited presence of specific nutrients in soil. The competition between soil bacteria for these nutrients increases the production of their secondary metabolites, which include antibiotics.[12] Therefore antibiotic production is used as a defence strategy by bacteria in order to ensure survival in a competitive environment. It has been argued that secondary metabolites act upon specific receptors in competing organisms.[13] Among soil bacteria the largest antibiotic-producing genus are the Streptomycetes.[14] *Streptomyces* are Gram-positive bacteria that grow in various environments, and are characterised by a rich secondary metabolism. Moreover, the antimicrobial secondary metabolites produced by *Streptomyces* tend to be of low molecular weight with complex chemical structures. The production of antibiotics is specific for each species, depending upon their external environments.[15] Production of a given metabolite usually begins in response to a lack of or presence of a particular nutrient in the culture medium and tends to occur over a narrow time window, prior to events such as sporulation. Some examples of antibiotics discovered from different *Streptomyces* bacteria are shown below in Scheme 2.



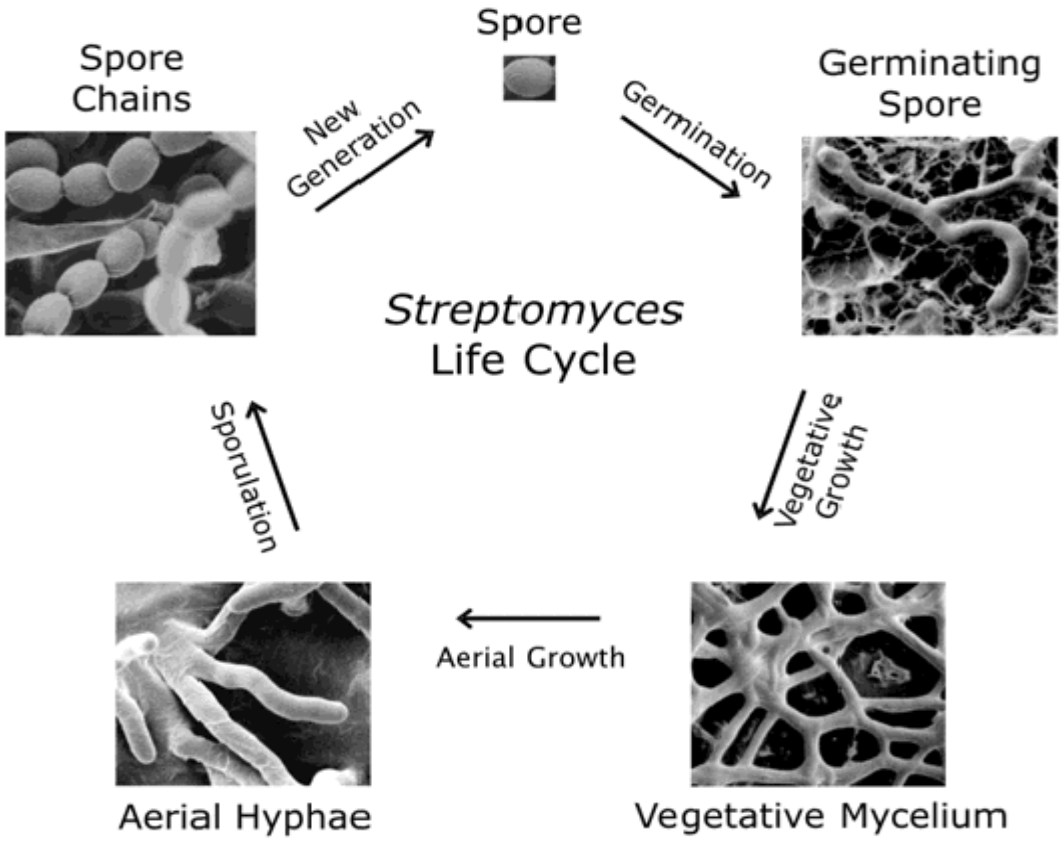
fosfomycin 1
(*Streptomyces fradiae*)

chloramphenicol 2
(*Streptomyces venezuelae*)

tetracycline 3
(*Streptomyces aureofaciens*)

Figure 1. Antibiotics isolated from *Streptomyces* bacteria.

Streptomyces bacteria form spores with very limited mobility. After spore germination, vegetative hyphae grow and then aerial hyphae with new spores are formed as coiled threads at the terminus. Finally, the layers of hyphae can differentiate into chains of spores. Each different stage of the *Streptomyces*' life cycle are described in Scheme 2.[15] Due to the limited availability of nutrients *Streptomyces* secrete antimicrobial secondary metabolites in order to protect their resources and to eliminate competing organisms.[12]



Scheme 2. *Streptomyces* life cycle.

1.2.2.1. *Streptomyces cattleya*

The soil bacterium *S. cattleya*, discovered in 1976, has the ability to produce the antibiotic thienamycin **4** (Figure 2). *S. cattleya* presents a recognizable phenotype, as it appears with purple spores as shown below in Figure 3. Thienamycin **4** was the first of a series of naturally occurring *β*-lactam antibiotics to be found with a carbapenem ring system.[16] This antibiotic presented with good activity against both Gram-positive and Gram-negative bacteria. Its mode of action is similar to that of penicillin, operating *via* disruption of cell wall biosynthesis. Thienamycin **4** was isolated using extensive purification procedures and its structure was elucidated in 1979, however, isolation of this antibiotic proved to be very difficult due to its instability.[16]

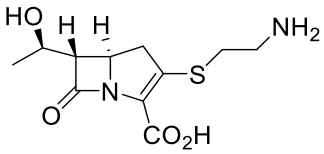


Figure 2. Structure of thienamycin **4**.

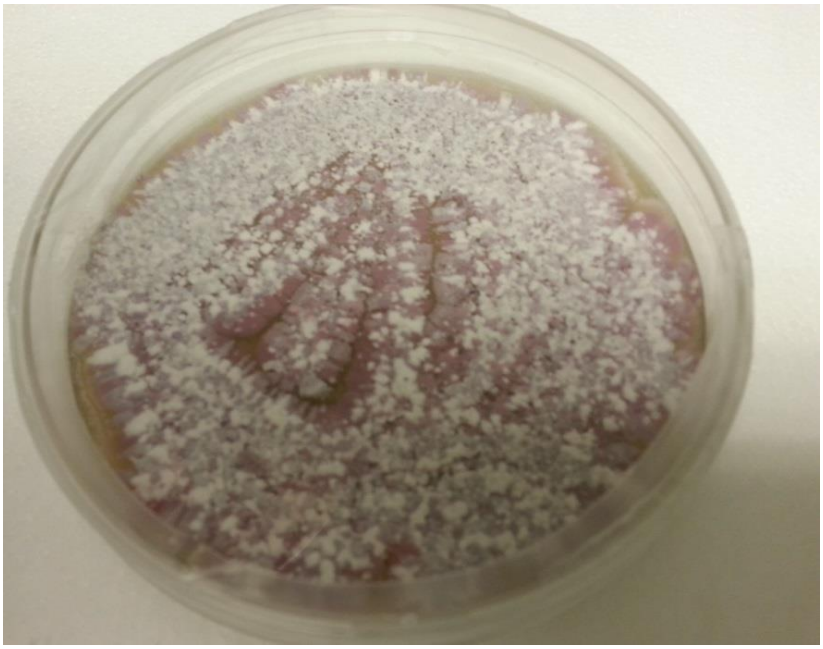
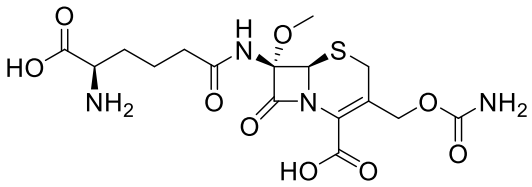
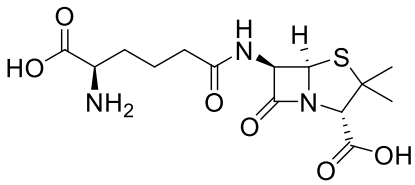


Figure 3. Photography of *Streptomyces cattleya*.

Thienamycin **4** is not the only *β*-lactam antibiotic produced by *S. cattleya*, indeed this soil bacterium has the ability to biosynthesise two other *β*-lactam antibiotics namely cephamycin C **5** and penicillin N **6** (Figure 4).[17] Cephamycin C **5** is a very efficient antibiotic against anaerobic microbes.



cephamycin C 5
(*Streptomyces cattleya*)



penicillin N 6
(*Streptomyces cattleya*)

Figure 4. Antibiotics isolated from *Streptomyces cattleya*.

1.2.2.2. *Streptomyces calvus*

In the 1950s a soil bacterium called *S. calvus* was isolated in India. Unlike *S. cattleya*, *S. calvus* displays a “bald” phenotype deficient in the formation of aerial mycelium and spores, as shown below in Figure 5. It is well established that the inability to sporulate can decrease the production of secondary metabolites in *Streptomycetes*. Despite a lack of aerial mycelium, this soil bacterium is able to produce two related antibiotics, adiposin 1 **7** and adiposin 2 **8** (Figure 6), which are α-glucosidase inhibitors.[18]

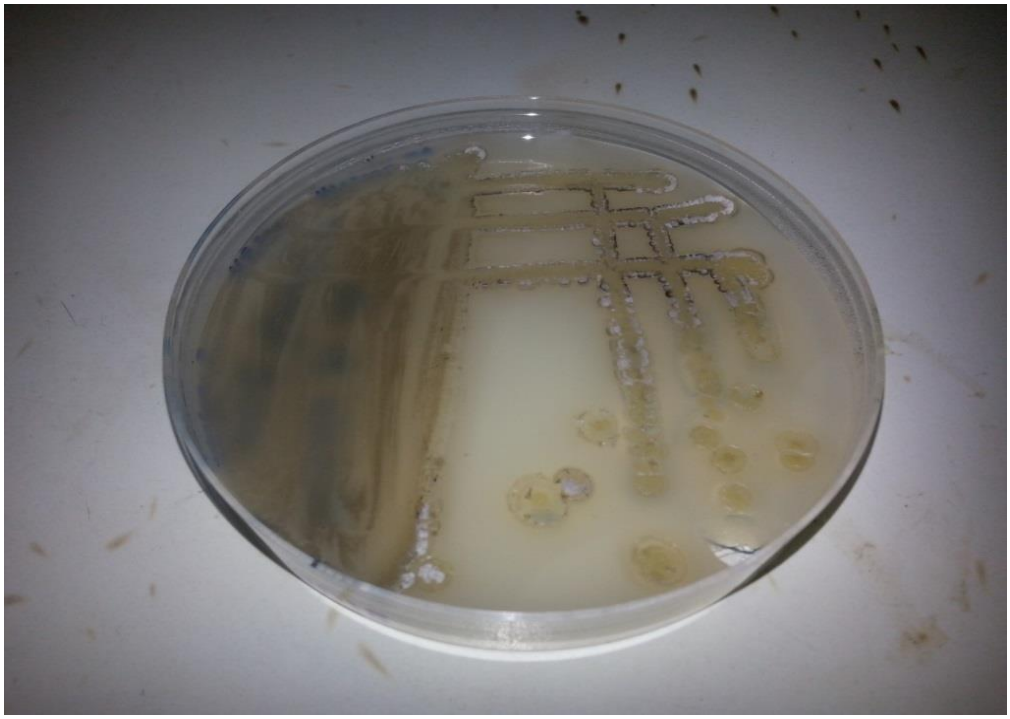
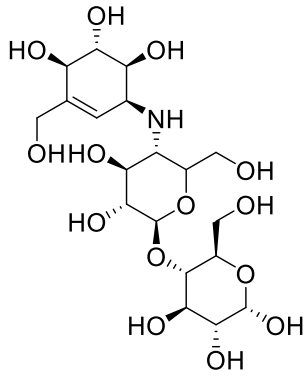
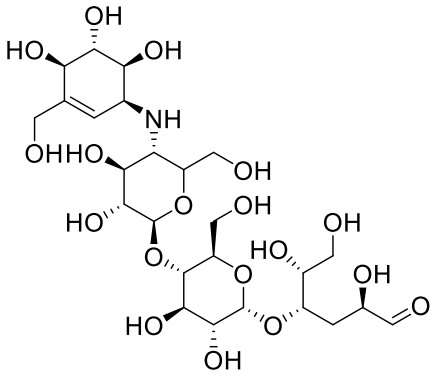


Figure 5. Photography of *Streptomyces calvus*.



adiposin 1 7
(*Streptomyces calvus*)



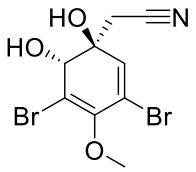
adiposin 2 8
(*Streptomyces calvus*)

Figure 6. Antibiotics isolated from *Streptomyces calvus*.

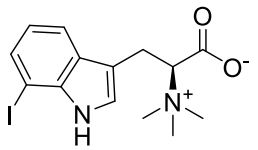
1.3. Halogenated natural products

1.3.1. Brominated, iodinated and chlorinated natural products

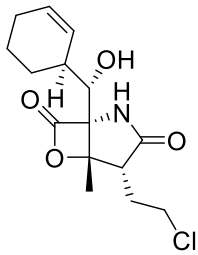
Over 4000 halogenated natural products have been isolated and structurally characterised.[19] Most natural organohalogens contain chlorine or bromine atoms, with iodinated and fluorinated much rarer by comparison; with just over 100 iodinated natural products isolated so far.[19] Natural products containing halogen atoms exhibit a wide range of important bioactive roles.[20] Moreover, these natural products are produced by a variety of organisms such as plants, soil and marine bacteria, mammals, *etc*. Some examples of brominated, chlorinated and iodinated natural products with very different bioactivities are shown below in Figure 7. Fluorinated natural products will be described later in this chapter.



aeroplysinin-1 9
(Marine sponge)



plakohypaphorine 10
(Caribbean sponge)



salinosporamide A 11
(*Salinispora tropica*)

Figure 7. Examples of halogenated natural products.

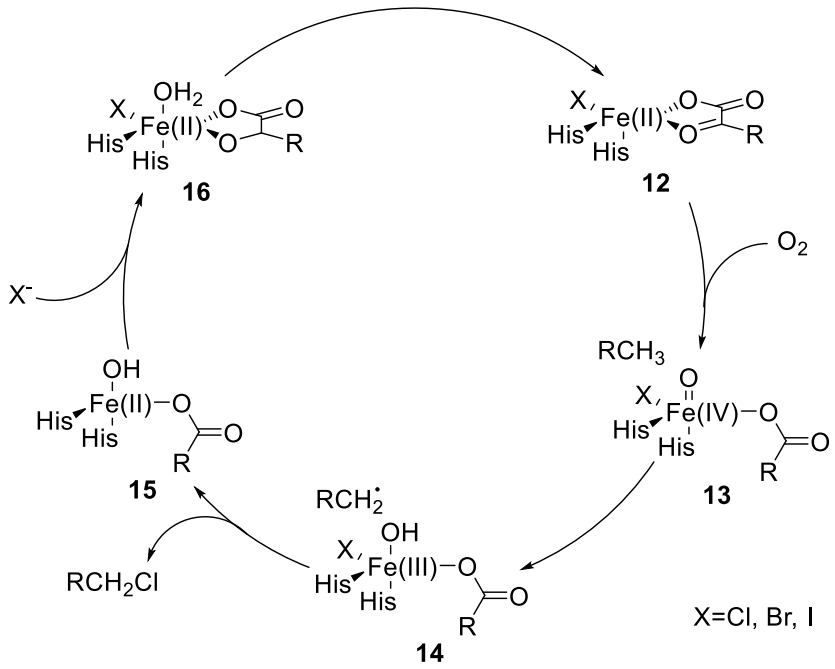
1.3.2. Biological halogenation with chlorine, bromine and iodine

Biological halogenation is achieved by three different mechanisms, *via*; a halide anion (X^-), a cationic halonium ion species (X^+) or a radical ($X^\cdot$).[21] However most biological halogenations use either oxidative or radical reactions to generate formal electrophilic (X^+) or radical ($X^\cdot$) species from the corresponding chloride, bromide or iodide ions.

The main mechanism of biological halogenation is *via* the formation of a reactive hypohalite species through oxidation of the halide.[21] This strategy allows the formation of carbon-halogen

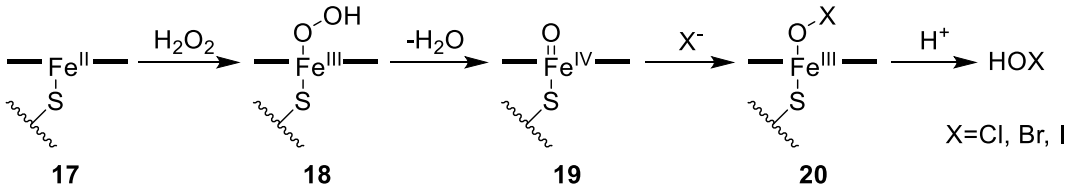
bonds between an electron-rich carbon centre and an electron-deficient hypohalite. Numerous enzymes are responsible for this type of biological halogenation such as haloperoxidases and halogenases.[21] In this section only biological halogenation concerning bromine, chlorine and iodine are detailed. Biological fluorination will be discussed in further details in the next section.

A large number of iron-dependent enzymes were discovered from various organisms, which are able to biocatalyse radical halogenation.[22] These enzymes are called non-haem-iron halogenases, and they are able to promote halogen-carbon bond formation with an unactivated sp^3 carbon, such as a methyl group. They catalyse radical halogenation only in the presence of molecular oxygen, α-ketoglutarate and a chloride or bromide ion, which initially affords the formation of halo-Fe(IV)-oxo species **13**. The next step proceeds *via* the abstraction of a hydrogen radical from the substrate to produce an Fe(III) species **14**. Finally, the last step in this halogenation mechanism is the formation of halogen-carbon bond, promoting the reduction of the iron centre which returns to the ferrous oxidation state, as shown below in Scheme 3.[23]



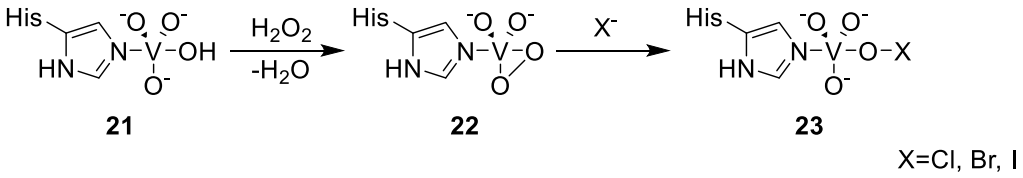
Scheme 3. Catalytic cycle of Fe-dependent halogenases.

Despite the existence of these non-haem-iron halogenases, which catalyse radical halogenation, most biological halogenation is achieved *via* cationic halonium ions (X^+). The first halogenating enzyme to be isolated was a heme-iron haloperoxidase from the fungus *Caldariomyces fumago*.[24] The mechanism of this enzyme was subsequently elucidated and is shown below in Scheme 4. During the first step peroxide binds to the resting Fe^{III}-porphyrin complex **17**, generating **18**. The intermediate Fe^{IV}-oxo species **19** is then formed by water elimination, and is intercepted by halide to form Fe^{III}-hypopalite species **20**. Finally, the halogenation step can be achieved by two different routes; it can either be carried out directly from the Fe^{III}-hypohalite species or hypohalous acid can be released from **20** and the halogenation reaction takes place away from the Fe centre.[25]



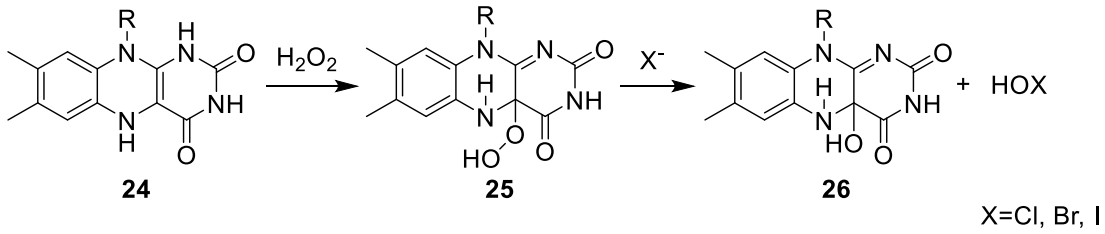
Scheme 4. Mechanism of heme-iron haloperoxidase.

Studies of biological halogenation in marine environments have revealed the presence of a second subclass of haloperoxidase enzymes, which utilise a vanadium co-factor.[26] The mode of action of these vanadium haloperoxidase, shown below in Scheme 5, is very similar to that of the iron haloperoxidase mechanism. Indeed, it begins with the binding of peroxide to the vanadium centre to form an oxo-peroxo-V(V) species **22**. This species **22** is an activated intermediate that is intercepted by halide in order to generate the hypohalite ion **23**. Finally the nucleophilic substrate reacts with the hypohalite ion **23** to establish the halogen-carbon bond.[27]



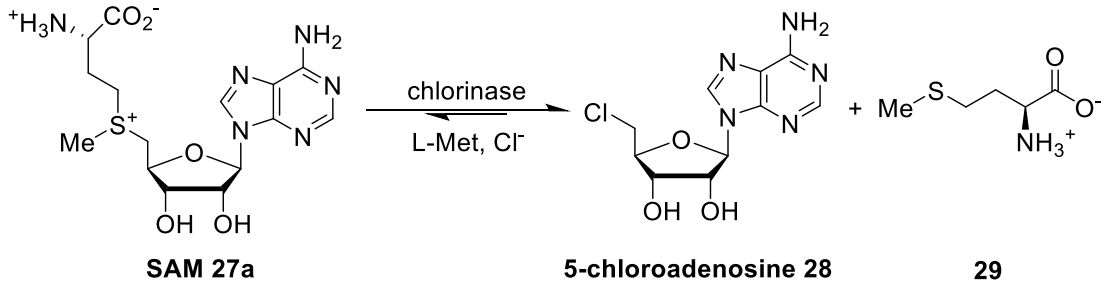
Scheme 5. Formation of the hypohalite **23** by vanadium haloperoxidase.

Most recently, another type of electrophilic enzymatic halogenation was discovered using flavin as a co-factor. During flavin catalysed halogenation, $FADH_2$ **24** reacts with peroxide to form an oxidised flavin intermediate **25**, as shown below in Scheme 6. The halide ion is then oxidised by this flavin intermediate **25** to give hypohalous acid, which consequently reacts with the substrate.[28]



Scheme 6. Mechanism of flavin catalysed halogenation.

The last known process of biological halogenation is *via* a nucleophilic halide (X^-).[21] One example of this mechanism was recently discovered from *Salinispora tropica.* This SAM-dependent chlorinase is responsible for the chlorination step during the biosynthesis of the natural product salinosporamide A **11,** which was described in Figure 7. This enzyme transfers a chloride anion to the 5'-carbon atom of the *S*-adenosylmethionine **27a** substrate to form 5-chloroadenosine **28**, as shown in Scheme 7. The first step of this mechanism involves the desolvatation of the chloride by replacement of hydrogen bonds to water with hydrogen bonds to amino acid residues within the enzyme active site. The desolvated chloride then attacks the 5'-carbon atom of the sulfonium centre, displacing methionine **29**, to generate 5'-chloroadenosine **28** in a S_N2 type reaction.[29]



Scheme 7. Formation of a chlorine-carbon bond mediated by the chlorinase.

1.3.3. Fluorination techniques

1.3.3.1. Fluorine: properties, effects and applications

Elemental fluorine was prepared for the first time in 1886 by Henri Moissan, who was subsequently granted the Nobel Prize in 1906 for his work. Fluorine is an exceptional atom amongst the halogens in several respects. Firstly, it is the smallest of the halogens (its van der Waal's radius is close to that of hydrogen and the hydroxyl group) and it is the most electronegative of all of the elements. Therefore the fluorine-carbon bond is highly polarised, which provides it with a substantial ionic character. It is very unlikely to form a positive fluorinium ion. The only natural isotope of fluorine is fluorine-19 with a spin quantum number (I) of ½, the same as hydrogen. This property is useful for NMR spectroscopy, which allows the detection of fluorine in biological systems. Physical properties of fluorine and hydrogen atoms are compared below in Figure 8.

X	Bond dissociation energy (kcal/mol)	Bond length (Å)	Van der Waal's radius (Å)	Electronegativity (Pauling scale)
F	106	1.38	1.47	4.0
H	104	1.09	1.20	2.2

Figure 8. Comparison of physical properties of fluorine and hydrogen.

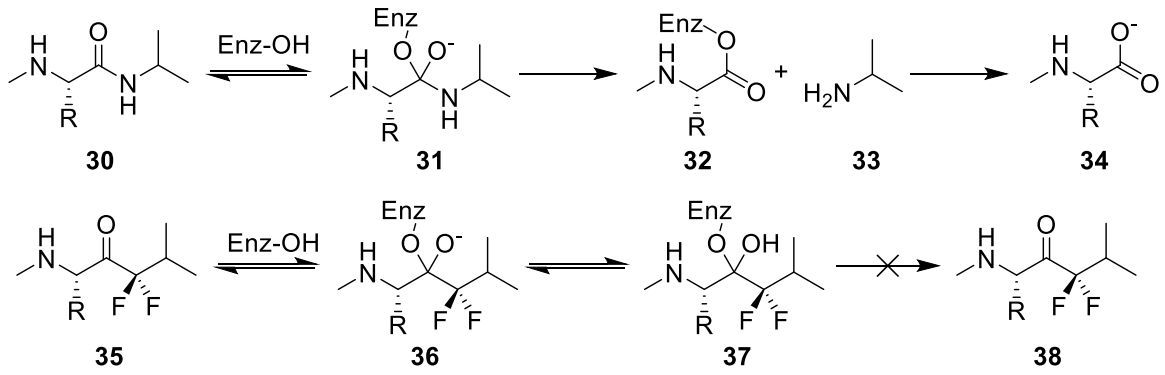
All of the properties discussed previously explain why substitution of a hydrogen or hydroxyl group with a fluorine atom in a molecule can be highly profitable. This type of substitution does not induce significant conformational changes, because the bond lengths and van der Waal's radius are very similar. However, it generates impactful electronic effects because of fluorine's high electronegativity. Fluorine can also form a weak hydrogen bond by acting as a hydrogen bond acceptor.[30] Aromatic and olefinic fluorine atoms are weaker hydrogen bond acceptors than aliphatic fluorine atoms. Therefore, in order to investigate substrate-enzyme interactions, fluorine atom replacing hydroxyl group is a method widely used.[31]

The electronegativity of fluorine can change the pKa of functional groups depending on the position of the fluorine atom. For instance, acetic acid has a pKa of 4.8 and trifluoroacetic acid has a pKa of 0.2. Introduction of a fluorine atom close to carboxyl or phosphate group tends to increase their acidity, but its close proximity decreases the basicity of amino groups. All these changes can provoke significant consequences in substrate/enzyme binding, particularly when binding efficiency is sensitive to acidity.

By definition, a fluorine atom is not a good leaving group because of the high bond dissociation energy of the fluorine-carbon bond. However, its ease of hydration favours hydrogen fluoride elimination (HF) under specific conditions. For instance when a fluorine atom is located vicinal to an acidic hydrogen, elimination of hydrogen fluoride can occur under both acidic and basic

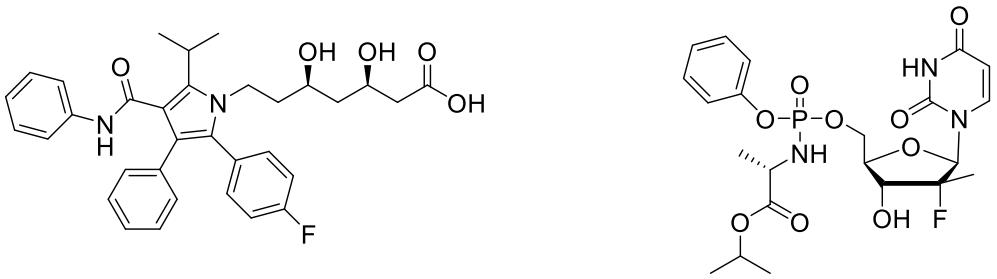
conditions. This strategy has been significantly investigated in the production of enzyme inhibitors known as suicide substrates.[32]

For instance, methyl ketone substrates have proven to be inhibitors of the acetylcholinesterase enzyme. Fluorinated ketone analogues have shown better efficiency as inhibitors through formation of a stable hemiketal with the active site serine residue.[33] Fluoro ketone substrate analogues are also inhibitors of protease enzymes when replacing NH with a CF_2 moiety. Fluorinated amide mimetic **35** can undergo the same nucleophilic attack from the enzyme; however, cleavage cannot occur as it does with the amide, see scheme 8.



Scheme 8. Example of an enzyme inhibition by a fluorinated analogue.

The introduction of a fluorine atom or trifluoromethyl group into a molecule tends to change its lipophilicity. Fluorination often increases the lipophilicity, which makes the organocompound more fat soluble. However, this is not general; for instance, introduction of a fluorine atom in *α* position to carbonyl compounds lowers the lipophilicity.[34] These properties of the fluorine atom and the effects caused by hydrogen-to-fluorine substitutions reinforce why fluorine is considered a low-risk, and potential high impact substituent. Appropriate introduction of a fluorine atom during drug optimisation can improve efficacy. Nowadays around one fifth of all drugs in the market contain at least one fluorine atom. Indeed, the structures of two blockbuster drugs contain fluorine substituents as shown below in Figure 9.

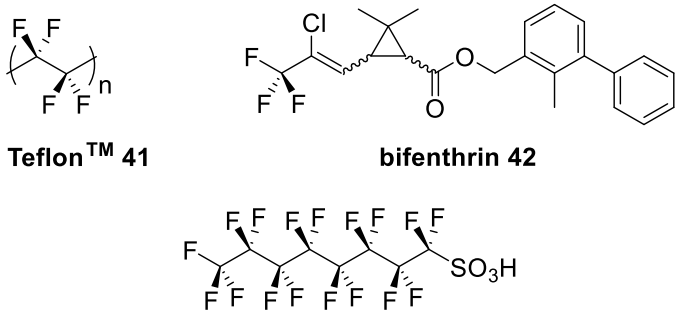


atorvastatine 39 **sofosbuvir 40**

Figure 9. Blockbuster drugs containing fluorine substituents.

Atorvastatine **39**, also known as Lipitor, is a lipid lowering agent which is used for prevention of cardiovascular disease. It is an inhibitor of HMG-CoA reductase which is an enzyme in the liver responsible for the production of cholesterol.[35] Sofosbuvir **40** is a drug blockbuster which is used for the treatment of hepatitis C. This prodrug **40** is metabolised to an active antiviral agent that acts as an inhibitor of viral RNA synthesis.[36]

Fluorine substituents are not only used in medicinal chemistry but are also found in numerous other applications. Indeed, production of fluorinated organic compounds for industrial application began in the 1930s with the preparation of chlorofluorocarbons as refrigerants.[37] The major turning point for the development of industrial fluoroorganic chemistry was the Manhattan project during the 1940s. This project was launched in the context of War World II for the production of weapons. From the 1950s more civilian applications of industrial fluoroorganic chemistry were developed such as in pharmaceuticals. Nowadays the fluorochemical industry is involved in a wide range of different fields of application, such as: fluoropolymers, agrochemicals, fluorosurfactants, liquid crystals, *etc*. Some examples of these are shown below in the Figure 10.



Teflon™ 41 **bifenthrin 42**

perfluorooctanesulfonic acid 43

Figure 10. Examples of fluorinated compounds used in industry.

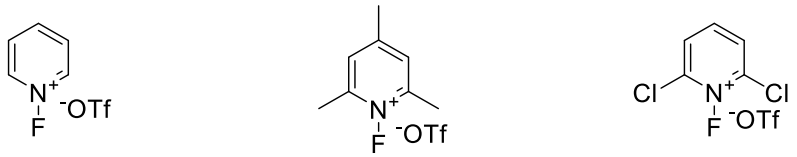
Polytetrafluoroethylene **41**, also known as Teflon™, is a fluoropolymer which is used for many different applications. It was discovered accidentally in 1938 by Roy Plunkett.[38] Most applications of Teflon involve wiring in aerospace and computer applications; however it can also be used in cookware such as coating non-stick frying pans. Bifenthrin **42** is an insecticide from the family of pyrethroids, which are harmless to humans, but are toxic to insects.[39] This particular insecticide is primarily used against the red imported fire ant by acting on the insect's neuromuscular system.[40] Perfluorooctanesulfonic acid **43** is a fluorosurfactant which is utilised across various industries, such as: commercial aviation (hydraulic fluid), fire retardant flames industry, semiconductor industry, *etc.*[41,42]

1.3.3.2. Fluorination reactions

Despite the fact that the fluorine atom is widely used and considered as a crucial substituent for industrial applications, it is still very difficult to introduce fluorine into organic compounds. Fluorine gas was utilised as the first technique for fluorination and is still used today despite safe handling challenges. Due to the rigorous safety considerations of this technique, many novel fluorinating reagents have been developed in the last half-century. Organic reactions that create a C-F bond can be carried out *via* two different processes: electrophilic and nucleophilic fluorination.[43]

1.3.3.2.1. Electrophilic fluorination

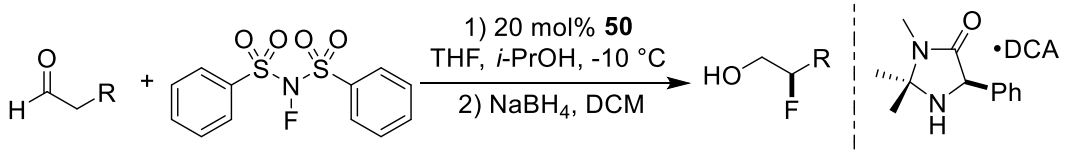
Cationic fluorine species cannot be formed because of the high electronegativity of fluorine, however, numerous species that contain an electrophilic source of fluorine have been recently developed. Reactivity of these electrophilic fluorinating reagents is based on a nitrogen-fluorine bond. The first type of these reagents to become wide-spread on the market was the *N*-fluoropyridinium salts e.g. **44**-**46** of Umemoto.[44] These reagents can react with most nucleophilic substrates and can be tuned by modifying the pyridinium ring with different functional groups as shown below in Figure 11.



44 **45** **46**

Figure 11. Structures of different *N*-fluoropyridinium salts.

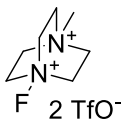
A second type of electrophilic fluorinating reagent is the *N*-fluorosulfonamides. The most widely used reagent of this class is *N*-fluorobenzenesulfonimide **48**, also known as NFSI. An enantioselective organocatalytic electrophilic fluorination of aldehydes with NFSI was developed by MacMillan and co-workers, as shown below in Scheme 9.[45] The aldehyde **47** reacts with the chiral amine **50** forming a chiral enamine, which then reacts with NFSI **48** forming α-fluoroaldehyde which was immediately reduced to fluoroalcohol **49** with $NaBH_4$ because of its inherent instability.



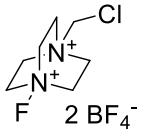
47 **48** **49** **50**

Scheme 9. Enantioselective fluorination by MacMillan.

The most widely used electrophilic fluorinating reagent Selectfluor **52** was developed by Banks, along with many of its derivatives (F-TEDA-X).[46] Selectfluor **52** is a commercially available reagent which is very stable and widely used in organic chemistry for numerous applications.[47] The oxidation potential of these F-TEDA-X reagents can be tuned by modification at the nitrogen substitution using electron-withdrawing substituents. Some structures of this type of electrophilic fluorinating reagents are shown below in Figure 12.



F-TEDA-OTf 51

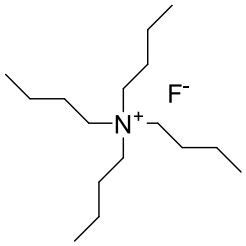


Selectfluor 52

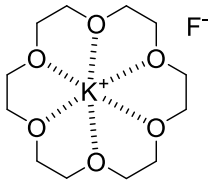
Figure 12. Structures of electrophilic fluorinating reagents.

1.3.3.2.2. Nucleophilic fluorination

Fluoride ion was the first source of nucleophilic fluorine, and can be found as alkali fluorides such as: NaF, KF, LiF and CsF. LiF is the least reactive of these alkali fluorides, because of the decrease in both nucleophilicity and solubility observed when ionic strength increases. In order to increase the solubility of these fluorinating reagents in organic solvents crown ethers can be used in combination with alkali metal fluorides, such as KF-18-crown-6 **54**.[48] However, a fluoride ion can form very strong hydrogen bonds which lowers its nucleophilicity, whereas its high basicity can often lead to elimination reactions. Therefore, additional fluorinating reagents based around the fluoride ion were developed, such as tetrabutylammonium fluoride **53** (TBAF).[49] The tetrabutylammonium counter ion increases the solubility of the fluoride ion, which in turn increases the reactivity of TBAF **53** as a fluorinating reagent. All of these nucleophilic fluorinating reagents can react with activated alcohols such as mesylates, tosylates, triflates, halides, *etc*.



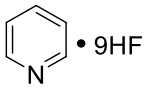
tetrabutylammonium fluoride 53



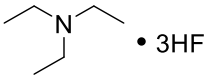
KF-18-crown-6 54

Figure 13. Structures of TBAF **53** and KF-18-crown-6 **54**.

Another important source of nucleophilic fluorine is based on anhydrous hydrogen fluoride (HF), a toxic colourless gas which necessitates rigorous safety considerations. This explains why substantial efforts were made to form stable hydrogen fluoride liquid reagents, such as HF-pyridine **55** and triethylamine tris(hydrogen fluoride) **56** ($Et_3N{\cdot}3HF$), as shown below in Figure 14. HF-pyridine **55** is an acidic nucleophilic fluorinating reagent which was synthesised by Olah in 1973 as a 70% *w/w* solution of HF in pyridine.[50] Triethylamine tris(hydrogen fluoride) **56** is a neutral reagent which is more stable to elevated temperatures than HF-pyridine **55**.



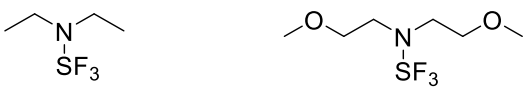
HF-pyridine 55



TREAT-HF 56

Figure 14. Structures of HF·pyridine **55** and triethylamine tris(hydrogen fluoride) **56**.

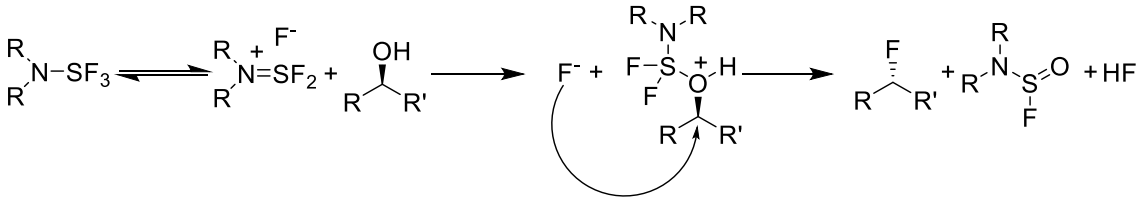
In the 1970's, due to the harsh conditions employed during the use of anhydrous hydrogen fluoride, safer nucleophilic fluorinating reagents were developed.[51] These reagents are easier to handle and sulfur-based, such as diethylaminosulfur trifluoride (DAST) **57** and Deoxofluor® **58** as shown below in Figure 15. These three nucleophilic fluorinating reagents are used for the deoxofluorination of alcohols.



DAST 57 **Deoxofluor® 58**

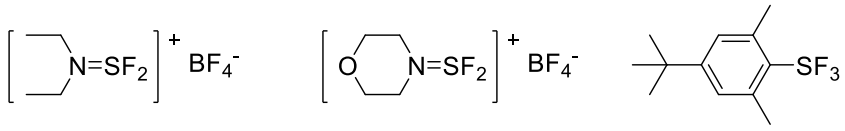
Figure 15. Structures of DAST **57** and Deoxofluor® **58**.

Both DAST **57** and Deoxofluor® **58** exist in equilibrium between their neutral and activated form with a fluoride counter ion. The lone pair of the alcohol attacks the sulfur atom of the activated form of DAST **57** or Deoxofluor® **58**, which causes the liberation of fluoride ion. This fluoride ion can then attack the electrophilic centre of the recently formed activated alcohol in order to give the fluorinated product.



Scheme 10. General mechanism of fluorination of alcohol with DAST **57** or Deoxofluor® **58**.

More recently, other sulfur-based fluorinating reagents such as XtalFluor-E® **59**, XtalFluor-M® **60** and Fluolead™ **61** were developed in 2010.[52,53] XtalFluor-E® **59** and XtalFluor-M® **60** are more stable and easier to handle than DAST **57** and Deoxofluor® **58**. Moreover, these reagents are safer because they do not generate free-HF. Fluolead™ **61** also proved to have a very high thermal stability. This reagent **61** also showed a resistance to aqueous hydrolysis which is a significant advantage in fluorine chemistry.



XtalFluor-E® 59 **XtalFluor-M® 60** **Fluolead™ 61**

Figure 16. Structures of XtalFluor-E® **59**, XtalFluor-M® **60** and Fluolead™ **61**.

1.3.3.3. Biological fluorination

Despite progress in the development of new fluorinating reagents, most require harsh conditions to react. This has led to a greater interest concerning fluorination methods found in nature. In contrast to biological halogenation with bromine, iodine and chlorine, which can be performed by electrophilic, nucleophilic or radical reactions, biological fluorination is a very rare phenomenon in nature. One reason for this is that fluorination in nature can be achieved only *via* a nucleophilic reaction, as fluoride cannot be oxidised by haloperoxidases to generate radical and cationic fluorine. Indeed, fluorine's redox potential is too high for coupling to known oxidations in living cells, such as haloperoxidase enzymes, as shown below in Figure 17.[21]

	I	Br	Cl	H_2O_2	F
Redox potential (V)	+0.54	+1.08	+1.36	+1.77	+2.87

Figure 17. Redox potentials of halogens and peroxide.

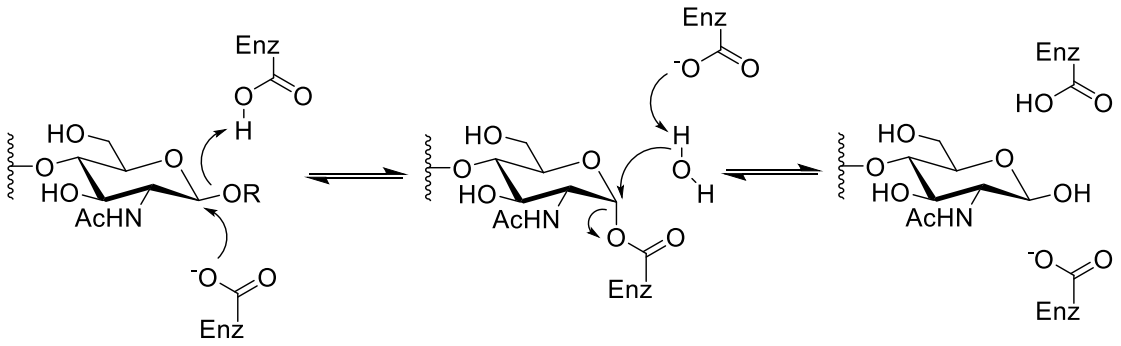
The level of fluoride in sea water is very low, at approximately 1.3 ppm. On the other hand, fluorine is the most abundant of the halogens in the Earth's crust.[54] Nevertheless, most fluorine in nature exists as insoluble fluoride minerals, such as fluorite (CaF_2) and cryolite (Na_3AlF_6) (Figure 18). Therefore, these fluoride minerals cannot be used by living organisms. Moreover, fluoride ion is a poor nucleophile in aqueous systems because it is highly hydrated and consequently this hydration shell reduces its nucleophilicity. As mentioned previously, the fluorine atom is very small so it forms numerous and strong hydrogen bonds with water molecules. Thus fluorine biochemistry has hardly evolved in nature. Despite all these disadvantages, a small number of discoveries have been made concerning biological fluorination.

cryolite **fluorite**

Figure 18. Photographies of cryolite and fluorite.

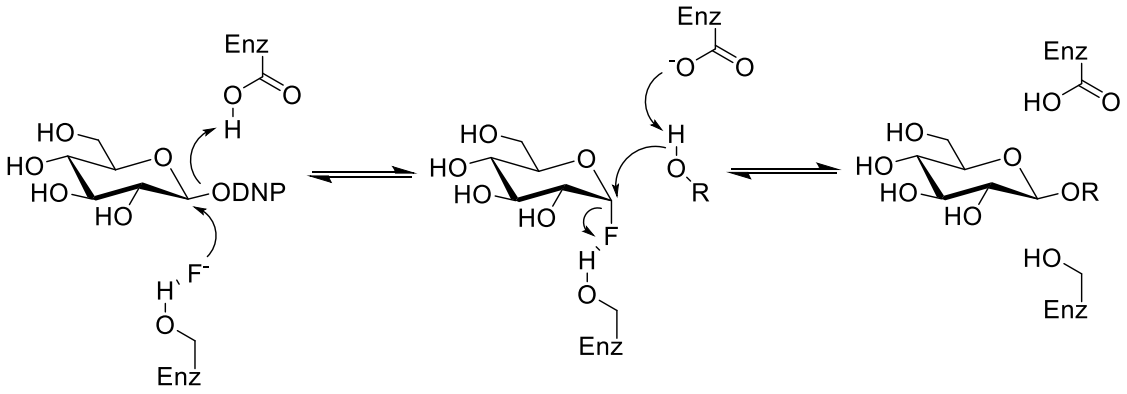
1.3.3.3.1. Glycosidase mutants

The first two examples for the enzymatic formation of a carbon-fluorine bond were reported with two different modified glycosidases. Glycosidase enzymes are responsible for the hydrolysis of glycosidic bonds in complex sugars. These enzymes are widespread in nature and are involved in numerous functions, such as: degradation of biomass, pathogenesis mechanisms, anti-bacterial defence strategies, *etc*. The mechanism of these glycosidase enzymes involve two steps; firstly a carboxylate group from an amino acid residue in the active site behaves as a nucleophile and attacks the anomeric centre generating a glycosyl enzyme intermediate. In the second step another carboxylate group acts as a base and deprotonates water, affording hydrolysis of the intermediate as shown below in Scheme 11.



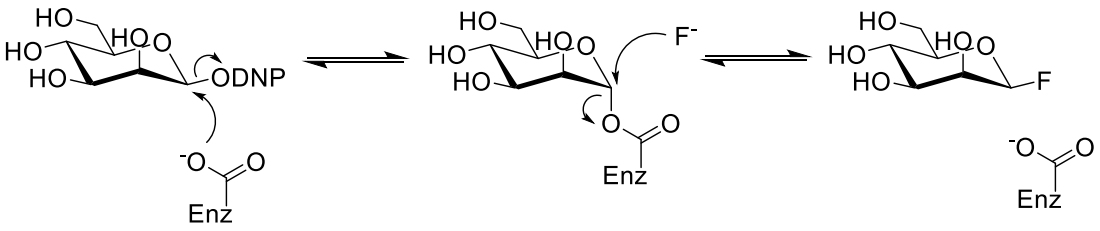
Scheme 11. General mechanism of glycosidase.

Examples of these two enzymes were engineered by site directed mutagenesis from two different bacteria, *Agrobacterium* sp. and *Cellulomonas fimi*.[55,56] These mutations replaced the glutamate residue within the active site of the glycosidases with alternate residues such as glycine, serine or alanine. The first enzyme, *β*-glucosidase from *Agrobacterium* sp., was engineered with a serine residue in place of the native glutamate residue within the active site. It was found that in the presence of inorganic fluoride and 2,4-dinitrophenyl *β*-glycoside, the engineered enzyme showed very high activity.[55,56] A proposed mechanism suggested that fluoride attacks the anomeric alcohol and transglycosylation occurs *via* transitory formation of glucosyl fluoride as depicted in Scheme 12.



Scheme 12: Proposed mechanism for transglycosylation by *Agrobacterium* sp. *β*-glucosidase.

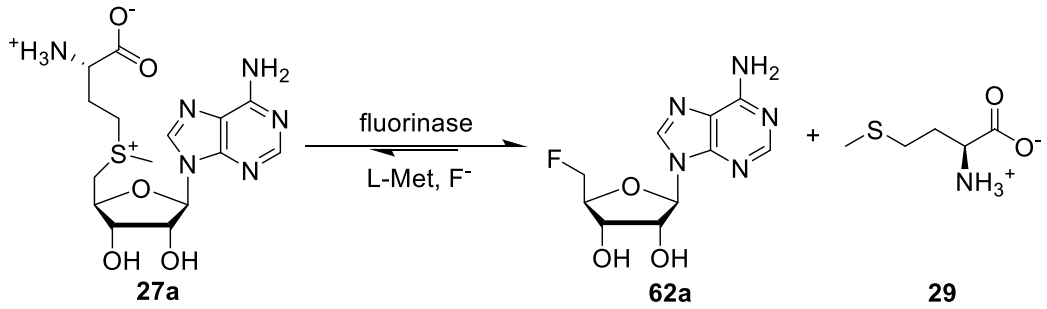
A second range of mutants was prepared from *Cellulomonas fimi* by substitution of the glutamate residue in the active site by an alanine residue. Due to the absence of the carboxylate group which should act as a base, deglycosylation did not take place causing the glycosyl enzyme intermediate to accumulate in the media. It was also discovered that in the presence of high concentrations [~2M] of inorganic fluoride the anomeric carbon can be attacked by fluoride anion giving *β*-mannosyl fluoride, as depicted in Scheme 13.



Scheme 13. Proposed mechanism for transglycosylation by *Cellulomonas fimi β*-mannosidase.

1.3.3.3.2. Fluorinase enzyme

Despite the fact that the enzymatic synthesis of a carbon-fluorine bond was first reported with mutant glycosidases, the fluorinating activity of these enzymes was not native. The first, and unique, native fluorinating enzyme was discovered in 2002 from a soil bacterium named *Streptomyces cattleya*.[57] This enzyme was given the common name "fluorinase" and *S*-adenosyl methionine **27a** was identified as the substrate. SAM **27a** is converted to 5'-fluorodeoxyadenosine **62a** (FDA) (Scheme 14), which is immediately turned over towards the biosynthesis of other fluorometabolites.



Scheme 14. Reaction of the fluorinase enzyme.

Following purification, the structure of the fluorinase enzyme was elucidated by X-ray diffraction.[58] The fluorinase exists as a dimer of trimers, containing three active sites with which the substrate SAM **27a** can bind. The locations of these active sites are rather unusual, in that they are located at the subunit interfaces. X-ray structures of the fluorinase with the substrate SAM **27a** and also the product FDA **62a** bound, revealed that the ribose rings of the bound substrate and product possess an unusually ring planar structure (Figure 19).[56]

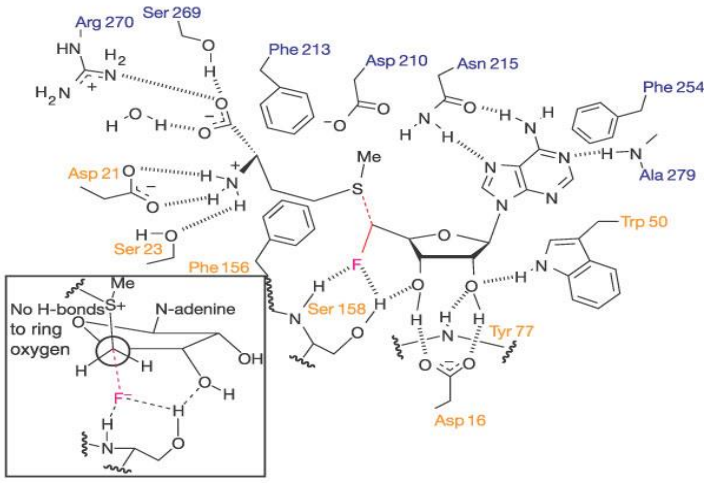
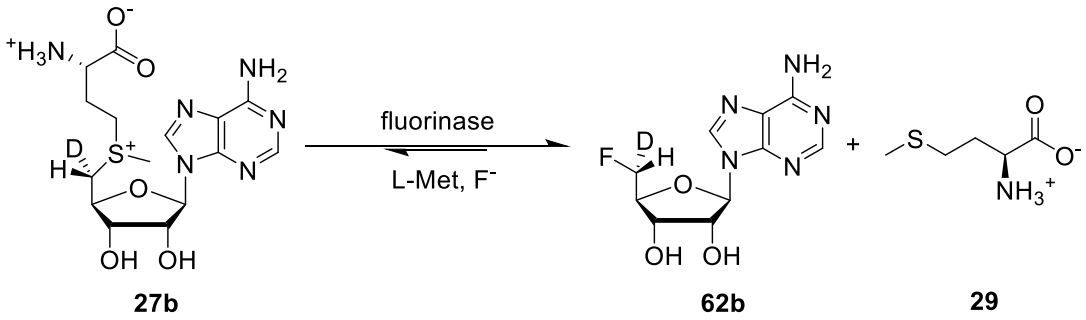


Figure 19. Representation of FDA **62a** and methionine **29** bound to the active site of fluorinase.

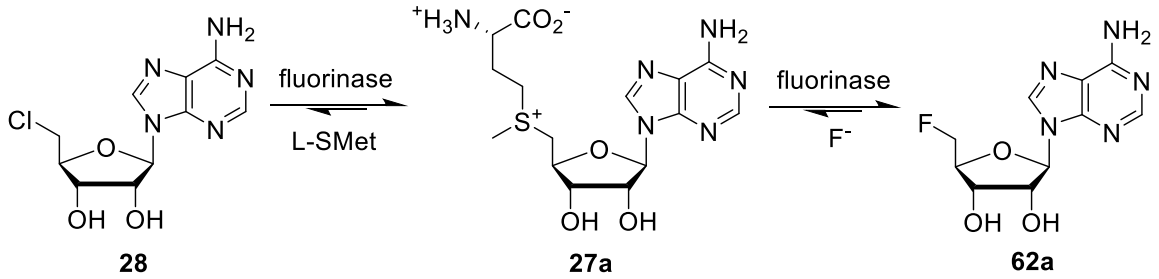
An experiment, probing the stereochemistry with SAM **27b** bearing deuterium at the *5'-pro-S* site, revealed that the fluorination reaction occurs with an inversion of configuration (Scheme 15), consistent with a S_N2 mechanism.[59] Crystallographic[58] and theoretical studies[60] indicated that the fluoride ion was desolvated within the enzyme, allowing hydrogen bonds with three amino acid residues within the active site to be established. Furthermore, there is an electrostatic interaction between the negatively charged fluoride ion and the positively charged sulfonium atom in SAM **27a**.[61] All of these factors contributed to stabilising the transition state, and consequently lowering the activation barrier.[60]



Scheme 15. The fluorination reaction with an inversion of configuration.

The fluorinase mediates the equilibrium between SAM **27a** and the products FDA **62a** and methionine **29**, favouring the products (Scheme 15). The enzyme has high substrate specificity, and as such only a small number of substrates have been shown to bind to its

active site. 5'-Chloro-5'-deoxyadenosine **28** (ClDA) is an example. The fluorinase also displayed chlorinase activity. Indeed, the enzyme catalysed the conversion of SAM **27a** to ClDA **28**, in the presence of inorganic chloride, in which the equilibrium lies heavily in favour of SAM **27a** (Scheme 16).[62] This observation has been particularly useful in the development of fluorinase mediated fluorination strategies, as ClDA derivatives are easier to synthesise than SAM derivatives.



28 **27a** **62a**

Scheme 16. Chlorinase activity of fluorinase.

1.4. Fluorometabolites

1.4.1. Fluoroacetate 63a

Fluoroacetate **63a** is the most common fluorinated natural product, and was first isolated by Marais in 1943 from a South African plant named *Dichapetalum cymosum*.[63,64] Due to the presence of fluoroacetate **63a** this small shrub was considered as a hazard to livestock in the Transvaal since the 1830s, and as such was known under the name of "gifblaar" which means "poison leaf".[65] This plant proved to be very difficult to exterminate, because of its persistent root system. During the spring, the young leaves are particularly toxic and contain approximately 2500 mg kg^{-1} dry wt of fluoroacetate **63a**.

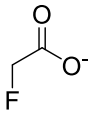


Figure 20. Structure of fluoroacetate **63a**.

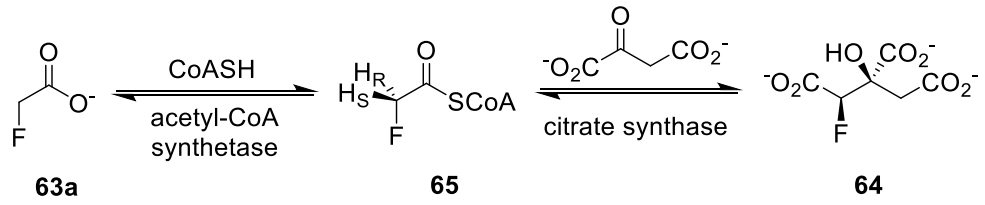
After the first isolation of fluoroacetate **63a**, this toxin was found to be produced by other species of *Dichapetalaceae* such as *D. braunii* in Tanzania.[66] Fluoroacetate producers have not only been found in the African continent. A Brazilian species called *Palicourea margravii*, which is a shrub, is able to biosynthesise this toxin in high concentration (5000 mg kg^{-1} dry wt) in its seeds. However, Australia remains the country with the highest number of plants producing fluoroacetate **63a**. Indeed more than 35 plant species were found in Western Australia with high concentrations of the toxin. *Gastrolobium bilobum* and *Oxylobium parviflorum* are both examples of Australian plants which accumulate high levels of fluoroacetate **63a**. Despite the fact that all of these fluoroacetate plant producers have different origins, they are all found within tropical weather regions. However, the biosynthetic pathway of fluoroacetate **63a** in plants has not been elucidated yet.

In 1986, the first microorganism able to produce fluoroacetate **63a** in the presence of inorganic fluoride was discovered.[67] This microorganism is a soil bacterium named *Streptomyces cattleya*. The production of this fluorinated natural product was detected fortuitously during the optimisation for the production of the antibiotic thienamycin **4**. The complete metabolic pathway of fluoroacetate **63a** in *S. cattleya* was subsequently fully elucidated and will be described later on.

1.4.2. Fluorocitrate 64

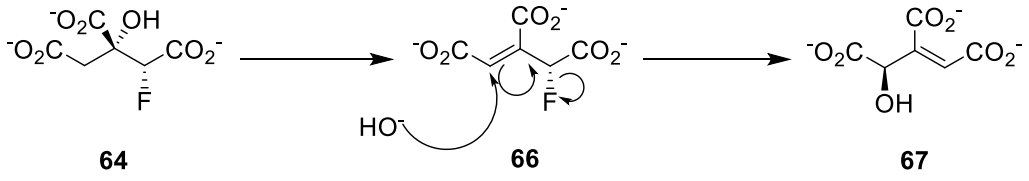
The toxicity of fluoroacetate **63a** is explained by the fact that it is absorbed into a living cell and then converted *in vivo* to fluorocitrate **64**. This metabolic conversion from fluoroacetate **63a** to fluorocitrate **64** was referred to as the "lethal synthesis" by Peters.[68] During this "lethal synthesis", fluoroacetate **63a** is converted to fluoroacetyl coenzyme A **65** by enzymes which usually accept acetate as their natural substrate.[69] Fluoroacetyl coenzyme A **65** subsequently enters into the Krebs cycle where a condensation reaction with oxaloacetate is catalysed by citrate synthase forming fluorocitrate **64**. This reaction is stereospecific since only the (2*R*, 3*R*)-fluorocitrate **64** is produced. The first step of this condensation is the selective abstraction

of the 2-*pro*-*S* hydrogen. The second step is the attack at the *Si*-face of the carbonyl group of oxaloacetate which brings about an overall inversion of configuration at the fluoromethyl group of fluoroacetyl coenzyme A **65**.[70]



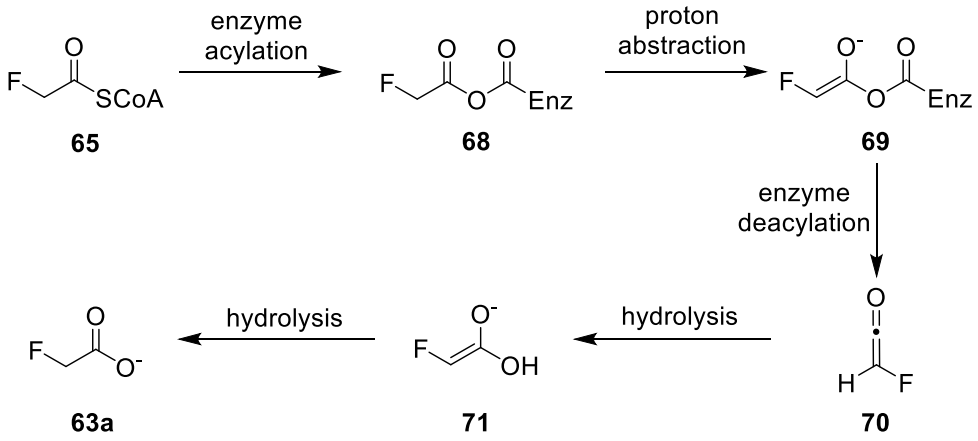
Scheme 17. Biosynthesis of fluorocitrate **64**.

The toxicity of fluorocitrate **64** can be partially attributed to the fact that it is a competitive inhibitor of the aconitase enzyme.[71] Aconitase is the subsequent enzyme after citrate synthase on the Krebs cycle, which usually catalyses the interconversion of citrate and isocitrate. The exact mechanism for the inhibition of aconitase by fluorocitrate **64** has not yet been fully elucidated, but a proposed mechanism was described after the reporting of an X-ray structure where 4-hydroxy-*trans*-aconitate remained tightly bound to the enzyme.[72] The first step begins by dehydration of fluorocitrate **64** to generate fluoro-*cis*-aconitate **66**. Hydrolysis of the double bond of fluoro-*cis*-aconitate **66** leads to the displacement of fluoride resulting in the production of 4-hydroxy-*trans*-aconitate **67** as shown below in Scheme 18. However, the toxicity of fluorocitrate **64** cannot solely be explained by inhibition of the aconitase enzyme. The major toxicity of fluorocitrate **64** is due to the fact it binds covalently to a citrate carrier protein inside the mitochondria. This binding causes the inhibition of tricarboxylic acid transport which leads to the death of the cell.[73,74]



Scheme 18. Proposed mechanism for the inhibition of the aconitase by fluorocitrate **64**.

To avoid autotoxicity from fluorocitrate **64**, plants and microorganisms have developed several mechanisms, one of which, among plants producing fluoroacetate **63a**, is to compartmentalise fluoroacetate **63a** in vacuoles or sacs remote from the mitochondria. It was also suggested that in some plants the citrate synthase enzyme has a lower affinity for fluoroacetyl coenzyme A **65**. The only known mechanism of defence from fluorocitrate **64** in microorganisms was discovered in *S. cattleya*. One gene from *S. cattleya*'s genome has the ability to encode for a fluoroacetyl-CoA thioesterase. This enzyme allows the hydrolysis of fluoroacetyl coenzyme A **65** as described below in Scheme 19.[75] After enzyme acylation, a proton is abstracted by a histidine residue forming the enolate **69**. After deacylation of the enolate **69**, a ketene intermediate **70** is formed, followed by hydrolysis affording fluoroacetate **63a**.

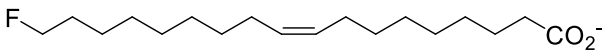


Scheme 19. Hydrolysis of fluoroacetyl coenzyme A **65** to fluoroacetate **63a**.

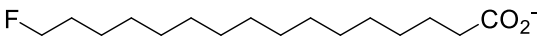
1.4.3. ω-Fluorofatty acids

In plants, such as the West African shrub *Dichapetalum toxicarium*, fluoroacetate can be further metabolised to toxic ω-fluorofattyacids. These fluorofatty acids are produced in the seeds, comprising about 3% of the seed oil and represents 80% of fluoroorganic compounds biosynthesised by the plant.[76,77] Further investigation using GC/MS technology revealed that plants can produce a large range of ω-fluorofattyacids, such as: ω-fluorooleic acid **72**, ω-fluoropalmitic acid **73**, ω-fluorocapric acid **74**, *etc*. All of these ω-fluorofattyacids isolated were found to be more toxic than fluoroacetate **63a** due to their higher lipid solubility, which

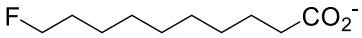
facilitates transport across cell membranes. It is thought that plants evolved to produce ω-fluorofattyacids as a mechanism of defence against predators such as herbivore animals.



ω-fluorooleic acid 72



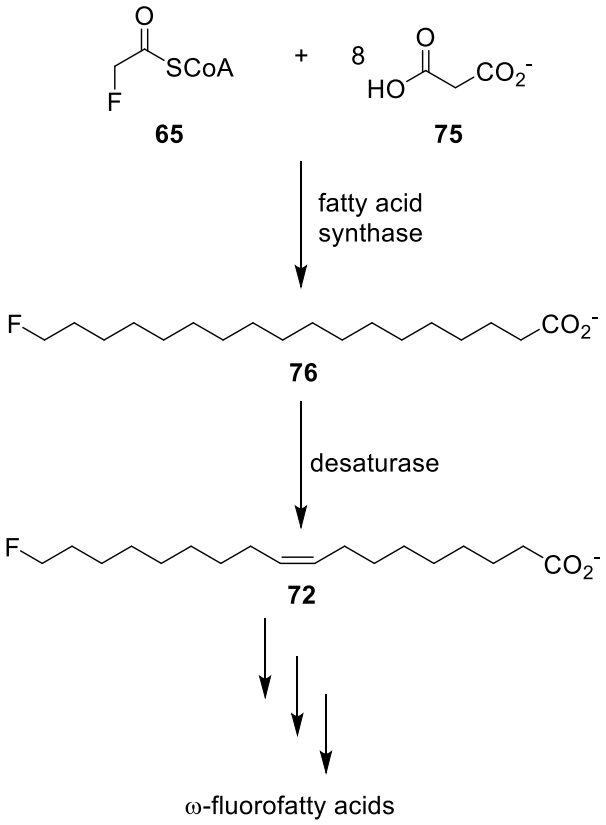
ω-fluoropalmitic acid 73



ω-fluorocapric acid 74

Figure 21. Examples of natural ω-fluorofattyacids.

The metabolic pathway of these ω-fluorofattyacids is identical to the conventional biosynthesis of fatty acids. Due to the high structural similarity to the natural substrate of the fatty acid synthase enzyme, fluoroacetyl coenzyme A **65** undergoes a condensation with malonyl-acyl carrier protein.[78] In the biosynthesis of ω-fluorooleic acid **72**, conversion from fluorostearoyl-ACP to ω-fluorooleic-ACP is catalysed by a nonspecific stearoyl desaturase. ω-Fluorooleic acid **72** is generated by a final hydrolysis as described in Scheme 20. All fluorofatty acids isolated contained the fluorine atom at the ω-position. This can be explained by the fact that the enzymes involved in the fatty acid biosynthesis have very narrow substrate specificity, only accepts fluoroacetyl coenzyme A **65** as a starter unit and as such fluorinated malonyl-CoA is not accepted.



Scheme 20. Biosynthesis of ω-fluorooleic acid **72**.

1.4.4. 4-Fluoro-L-threonine 77a

As described previously, the soil bacterium *Streptomyces cattleya* has the ability to produce fluoroacetate **63a**, along with another fluorinated natural product, 4-fluoro-L-threonine **77a**. They are produced in a 2:1 ratio. Indeed, during optimisation for the production of the *β*-lactam antibiotic thienamycin **4** by *S. cattleya*, the fluorinated 4-fluoro-L-threonine **77a**, was identified and isolated in 1986.[67] This was discovered fortuitously from a culture contaminated with inorganic fluoride in the medium. The structure shown in Figure 22 of this single stereoisomer was confirmed by asymmetric synthesis. The biological activity of fluorothreonine **77a** was investigated and revealed it to exhibit antimicrobial activity against a variety of bacteria most probably due to its ability to act as an antimetabolite of L-threonine. Following this finding, 25 species from the genus *Streptomyces* were cultured in the presence of inorganic fluoride ion. Over half of these species were able to uptake fluoride ion, however no fluorinated natural products were detected by ^{19}F-NMR.[79]

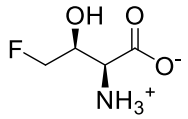


Figure 22. Structure of 4-fluoro-L-threonine **77a**.

1.4.5. Nucleocidin 78a

In 1957 the soil bacterium *Streptomyces calvus* was isolated in India. Interestingly, *S. calvus* has the ability to biosynthesise nucleocidin **78a**, the only fluorinated nucleoside discovered.[80] Nucleocidin **78a** is a nucleoside sulfonamide antibiotic containing a fluorine atom located at the 4'-carbon of the adenosine unit, as shown in Figure 23. Nucleocidin **78a** exhibits a wide range of antibacterial activity against gram-positive and gram-negative bacteria, including some pathogenic bacteria, in addition to its efficiency against trypanosomes.[80] However, due to its high toxicity, nucleocidin **78a** never became a clinical antibiotic. This fluorinated natural product is unique because it is the only fluorosugar derivative discovered in nature.

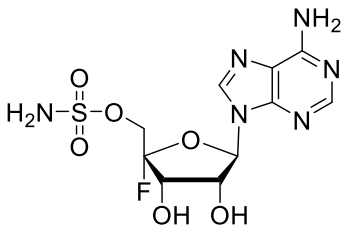


Figure 23. Structure of nucleocidin **78a**.

Nucleocidin **78a**, along with the analogous nucleoside antibiotics ascamycin and dealanylascamycin,[81] constitute a class of antibiotics that share a common structure, each containing a sulfonamide group at the ribose-C5' position. This class of antibiotic is known to inhibit protein biosynthesis in human cells by interfering with the ribosome-dependent translation process.[82] However, the biosynthetic mechanism of nucleocidin **78a** for inhibiting protein biosynthesis has yet to be elucidated.

Several structures were initially proposed for nucleocidin **78a**, but it was not until 1969 that it was recognised to contain a fluorine atom.[83] This confusion lay in the unusual coupling

patterns in the ^{1}H-NMR spectrum, which are due to ^{19}F-^{1}H couplings. Assignment of the D-ribo configuration was confirmed by the total synthesis of nucleocidin **78a** in 1976.[84] Nucleocidin **78a** appeared to be the first example of a furanose sugar bearing a functional substituent at C-4. The fact that the fluorine atom was attached to the C-4 of the ribose ring suggested that the biosynthesis of nucleocidin **78a** does not involve the fluorinase enzyme using SAM **27a**, but rather, has a unique fluorination enzyme. However, the biosynthesis of nucleocidin **78a** is, as of yet, completely unknown and thus requires further investigation.

Microbial large-scale production of nucleocidin **78a** was not established due to very low titres (1-5 mg/L) and poor reproducibility. Since 1957, re-isolation of nucleocidin **78a** from publically available culture collections has been unsuccessful.

1.5. Studies of metabolic pathways

1.5.1. Different techniques

The first technique commonly used for the elucidation of a metabolic pathway involves the use of putative precursors containing isotopic labels such as: ^{13}C, ^{2}H, ^{3}H, ^{15}N, *etc*. Radioisotopes can be detected by scintillation counting, however, a high degree of radiochemical purity is important to avoid errors from cross-contamination. Moreover, this technique does not reveal much in the way of regiochemical information, unless chemical degradation studies are employed. Stable isotopes have been used much more frequently due to the increasing sensitivity of analytical techniques available, such as GC/MS.

Isotope	Relative natural abundance, %	Half-life
^{2}H	0.015	stable
^{3}H	< 0.001	12 years
^{13}C	1.1	stable
^{14}C	< 0.001	5700 years
^{18}O	0.2	stable
^{15}N	0.37	stable
^{32}P	< 0.001	14 days

Figure 24. Properties of isotopes.

^{13}C-NMR is used to detect incorporation of an administered ^{13}C labelled putative precursor, since heavy atom induced shifts. The development of ^{2}H-NMR has also allowed the selective detection of incorporated deuterium into the metabolite under investigation. Deuterium labelling is approximately 60 times more sensitive than ^{13}C labelling, however ^{2}H spectra often present with poor resolution since ^{1}H vibrates at higher frequency (400 MHz) than ^{2}H. Another reason that explains its lower sensitivity is the quadrupolar nature of deuterium nucleus.

The advantage in the study of the biosynthesis of fluorinated natural products is that isotope incorporations can be detected by ^{19}F-NMR, without the need to purify the labelled fluorometabolites. However, in the presence of incorporated ^{2}H or ^{13}C the fluorine signals are shifted due to a heavy isotope effect. Incorporation of deuterium geminal to a fluorine atom causes an isotope shift of approximately 0.6 ppm to a lower frequency compared to the reference fluorine signal. This effect is additive, therefore two geminal deuterium atoms induce a more pronounced shift of about 1.2 ppm compared to that of the original fluorine signal. The shift induced by a deuterium atom bound geminal to the fluorine is considerably smaller (0.15-0.35 ppm according to the angle).

The use of labelled intermediates is also exploited with cell-free extracts. Cell-free extract experiments allow the study of co-factor requirements and substrate specificity. Indeed, putative intermediates, which can be chemically synthesised, are added to the cells in order to identify if they are involved in the biosynthesis of the desired metabolites.

Mutations to the genome of a particular organism of interest can be induced by numerous processes such as the treatment of cells with: UV-light, X-ray beams or chemical mutagens. These mutations can cause the accumulation of intermediates by blocking certain stages of the metabolic pathway. These mutations are induced at the genetic level by altering or knocking out gene enzymes.

1.5.2. Metabolic pathway of fluoroacetate and fluorothreonine in *S. cattleya*

1.5.2.1. Examples of feeding experiments in *Streptomyces cattleya*

The bacterium *Streptomyces cattleya* has been cultured in a medium with 20% D_2O. This produced fluorometabolites with a range of deuterium isotopomers. The $^{19}F\{H\}$-NMR spectrum of the supernatant of a resting cell experiment revealed each shift induced by deuterium incorporation, as shown in Figure 25 (Figure taken from M. R. Amin's thesis).[85] This spectrum revealed the production of two labelled fluoroacetates. The signal of the monodeuterated fluoroacetate is shifted upfield (α-shift) by approximately 0.6 ppm. Due to the additive nature of deuterium incorporation, the signal of the dideuterated fluoroacetate is shifted upfield (2x α-shift) by ~1.2 ppm. Considering fluorothreonine, the first upfield shifted peak resulted from a β-shift of around 0.2 ppm due to the presence of a deuterium atom at the C-3 position of fluorothreonine. The second upfield shifted peak is shifted about 0.6 ppm compared to the reference signal, due to the α-shift. The third upfield shifted peak of about 0.8 ppm corresponds to double label incorporation at positions the C-3 and C-4, which results from a combined α + β-shift. The last peak at the highest field is due to a 2x β-shift of about 1.2 ppm.

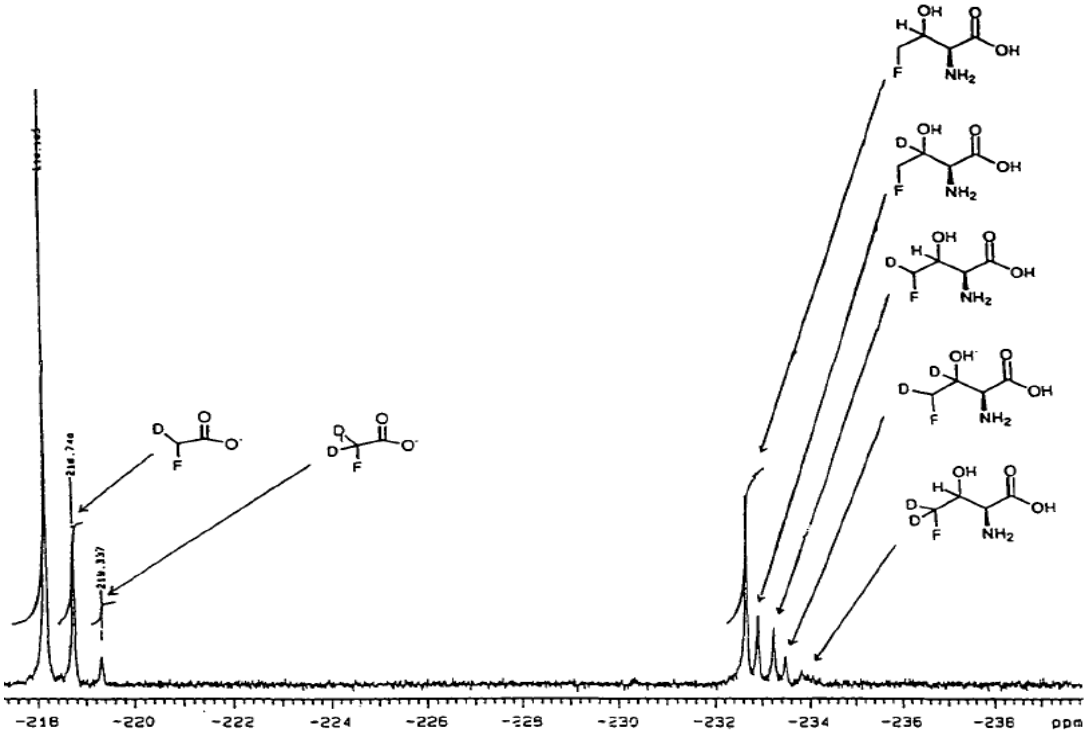


Figure 25. $^{19}F\{^{1}H\}$-NMR spectrum of the supernatant from resting cells of *S. cattleya* incubated with 20% of D_2O (M. R. Amin's thesis, University of Durham, 1996).

Feeding experiments using ^{2}H- and ^{13}C- metabolic precursors such as glycolate, glycine, serine, pyruvate, succinate, *etc,* suggested that only one fluorination enzyme is involved in the biosynthesis of both fluoroacetate **63a** and fluorothreonine **77a**.[86-88] Indeed both fluorinated natural products showed similar incorporations of the labelled precursors. Feeding experiments with synthetic [2,2-$^{2}H_2$]fluoroacetaldehyde proved that fluoroacetaldehyde **79** is the branching point in the metabolic pathway of fluoroacetate **63a** and fluorothreonine **77a**.[89]

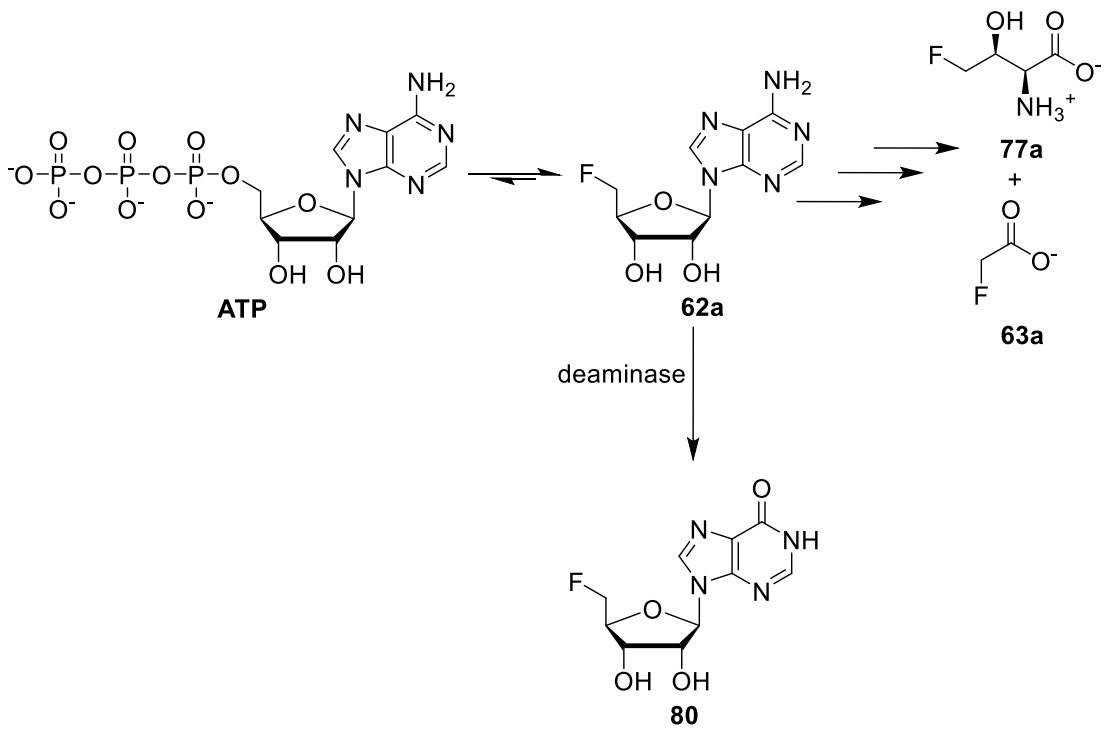
1.5.2.2. Detection and purification of the fluorinase

Cell free extract experiments were carried out using cells resuspended in a specific TAPS buffer (100 mM, pH 7.8) supplemented with potassium bicarbonate, magnesium chloride and potassium fluoride. Nucleotides such as ATP, UTP, GTP and CTP were administered to cell-free extracts and those experiments were incubated at 37 °C. After incubation of ATP in these cell-free extracts, four different fluorinated compounds were observed by ^{19}F-NMR: fluoroacetate **65**, fluorothreonine **77a**, 5'-fluoro-5'-deoxyadenosine **62a** (FDA) and 5'-fluoro-

5'-deoxyinosine **80** (FDI).[90] Control experiments were performed on cell-free extracts incubated without the presence of potassium bicarbonate, magnesium chloride, potassium fluoride or ATP. Such experiments were only successful in producing fluoroacetate **63a**, fluorothreonine **77a**, 5'-fluoro-5'-deoxyadenosine **62a** (FDA) and 5'-fluoro-5'-deoxyinosine **80** (FDI) in the presence of KF, ATP and $MgCl_2$. Magnesium chloride is very important because it stabilises the structure of ATP. At first, the triphosphate group of ATP was considered as a potential leaving group which could be attacked by a fluoride ion to form the product of the enzyme reaction named FDA **62a**.

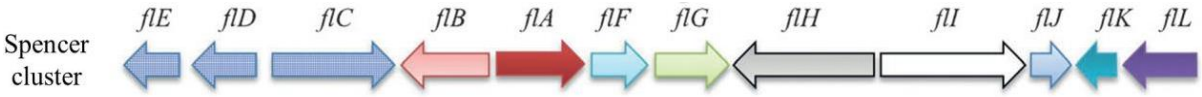
However, after further investigations, it was established that addition of the amino acid L-methionine **29** enhanced the production of the fluorinated organocompounds observed previously. This observation suggested that ATP might not be the substrate for the fluorinating enzyme. Indeed, SAM **27a** was then explored as the substrate for the fluorination reaction, as it is derived from ATP and L-methionine **29** by the action of SAM synthase. A cell-free extract experiment with SAM **27a** in place of ATP produced the same result. This experiment suggested that SAM **27a** is the natural substrate for enzymatic fluorination as shown in Scheme 21.

At first, FDI **80** was thought to be an intermediate in the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a**. After further investigation, it was proven that FDI **80** is not an intermediate but a metabolic shunt product of FDA **62a**. This side-product **80** is generated by the action of a deaminase enzyme which is known to be involved in purine degradation.



Scheme 21. Cell-free extract experiments with ATP.

The fluorinase was purified to homogeneity in a four step protocol (ammonium sulphate addition; hydrophobic column; gel-filtration chromatography; anion exchange chromatography) to purify the fluorinase enzyme with the molecular mass of approximately 180-190 kDa.[90] Consequently, over-expression of the fluorinase for X-ray crystallography studies became a priority. From analysis of the wild-type protein, a partial amino acid sequence was acquired by Edman degradation. This was used by Dr. Spencer from the University of Cambridge to design PCR primers and the fluorinase gene was then located and cloned into *Escherichia coli* in order to over-express this enzyme.[58] This method allowed for efficient purification of the enzyme.

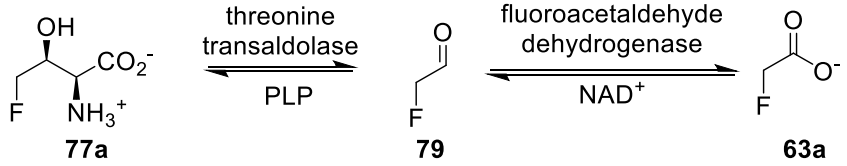


ORF	Length (amino acids)	Annotated function
flA	299	Fluorinase
flB	299	Purine nucleoside phosphorylase
flC	397	MFS permease
flD	216	Dehalogenase/Phosphatase
flE	222	DNA binding regulatory protein
flF	185	DNA binding regulatory protein
flG	234	DNA binding regulatory protein
flH	467	Na^+/H^+ antiporter
flI	489	Homocysteine lyase
flJ	131	DNA binding regulatory protein
flK	139	Thioesterase/Acyltransferase
flL	225	DNA binding regulatory protein

Figure 26. The Spencer cluster.

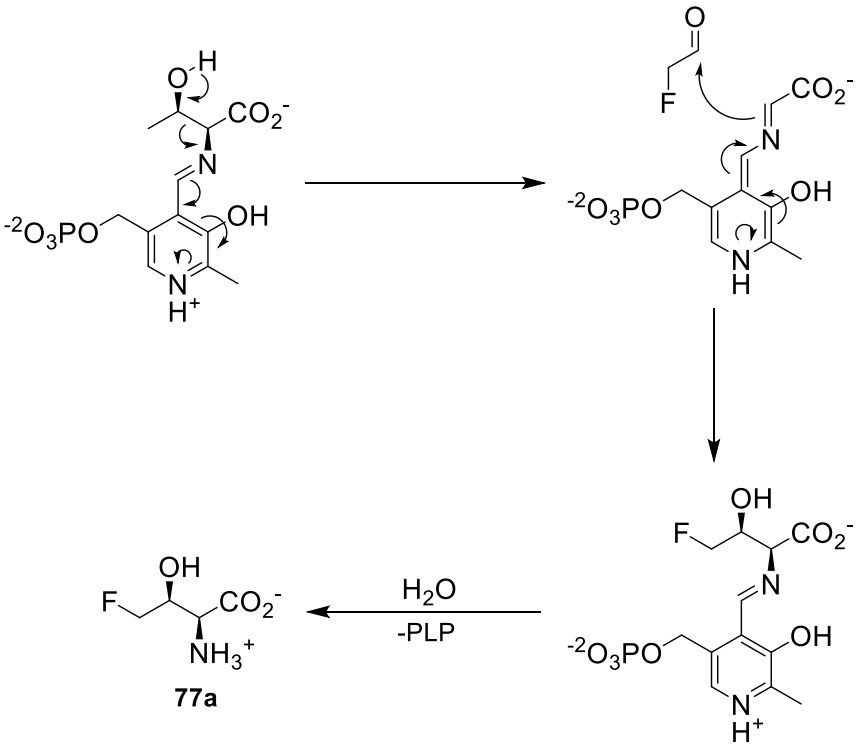
1.5.2.3. Metabolic pathway of fluorometabolites in *S. cattleya*

An aldehyde dehydrogenase, which is responsible for the oxidation from fluoroacetaldehyde **79** to fluoroacetate **63a**, was purified from *S. cattleya* homogenates.[91] This NAD^+-dependent enzyme is ten-fold more specific for fluoroacetaldehyde **79** over acetaldehyde. Upon incubation of a cell-free extract of *S. cattleya* with fluoroacetaldehyde **79**, L-threonine and pyridoxal 5'-phosphate (PLP), production of fluorothreonine **77a** was enhanced. When the assay was carried out in the absence of L-threonine and PLP, there was no detection of fluorothreonine **77a** by ^{19}F-NMR. These experiments allowed for the identification and purification of a PLP-dependent fluorothreonine transaldolase.



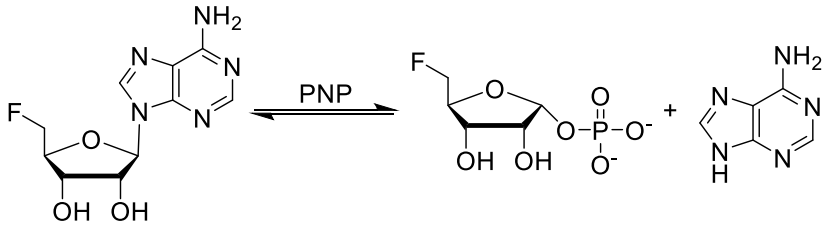
Scheme 22. Metabolism of fluoroacetaldehyde **79**.

The Schiff base was proposed to react with fluoroacetaldehyde **79** in an aldol-type reaction to afford the product fluorothreonine **77a** as shown in Scheme 23.[92] This PLP enzyme was found to be different from other related enzymes because glycine was not accepted as a substrate.



Scheme 23. Proposed mechanism for the enzymatic formation of fluorothreonine **77a** by threonine aldolase.

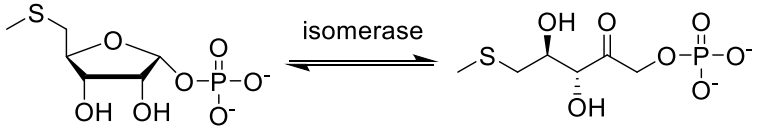
After investigation of the fluorinase enzyme, the fate of FDA **62a** was explored. Due to its similar structure to adenosine, enzymes that metabolise adenosine were examined. One of these enzymes was a purine nucleoside phosphorylase (PNP), which was incubated with FDA **62a** resulting in the displacement of the adenine base by phosphate at the anomeric carbon. This enzymatic assay revealed the conversion of FDA **62a** to 5-deoxy-5-fluoro-α,D-ribose 1-phosphate **81** (FDRP), as illustrated in Scheme 24. FDRP **83** was subsequently incubated in cell-free extract experiments of *S. cattleya* and production of fluoroacetate **63a** was observed, proving that FDRP **81** is a common intermediate in the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a**.[93]



62a **81** **82**

Scheme 24. The PNP reaction.

5-*S*-Methyl-5-thioribose-1-phosphate **83** (MTRP) is a well-established intermediate of the methionine salvage pathway and was found to be a close analogue to FDRP **81**. This is a common primary metabolic pathway which regenerates L-methionine **29** from waste by-products of SAM **27a**. Within this pathway, MTRP **83** is converted to 5-methylthioribulose-1-phosphate **84** (MTRPuIP) *via* ring opening of the hemiacetal group of the ribose ring to form an acyclic keto-sugar under the action of an MTRP isomerase as shown in Scheme 25.

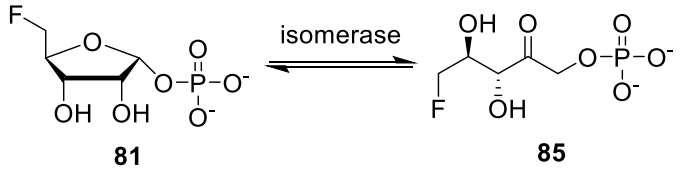


83 **84**

Scheme 25. MTRP isomerase in the methionine salvage pathway.

By homology, it was thought that a similar isomerase could accept FDRP **81** as a substrate and convert it to a 5-fluoro-5-deoxyribulose-1-phosphate **85** (FDRuIP). Moreover, ribulose phosphate compounds are well-known substrates for aldolase enzymes, creating aldehyde intermediates through a retro-aldol reaction. Therefore, fluoroacetaldehyde **79** could be generated from FDRuIP **85** by the action of a similar aldolase. FDRP **81** was incubated in cell-free extract experiments of *Streptomyces cattleya* and a new fluorinated compound was observed by ^{19}F-NMR. This proved to be FDRuIP **85** by GC-MS analysis after derivatisation of the reaction mixture with *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA).[94] PCR amplification of genomic DNA from *S. cattleya*, using primers generated from conserved regions of two different *Streptomyces* MTRP isomerases, led to the discovery of a similar gene

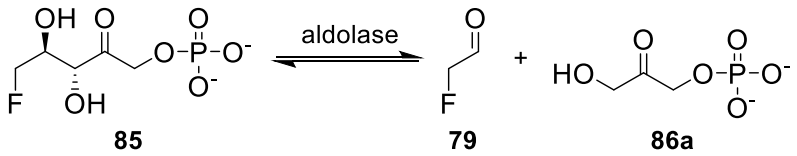
in the genome of *S. cattleya*. This gene was amplified, cloned into and over-expressed in *E. coli*. The protein obtained was able to convert FDRP **81** to FDRuIP **85**.



81 **85**

Scheme 26. Isomerase in the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a** in *S. cattleya*.

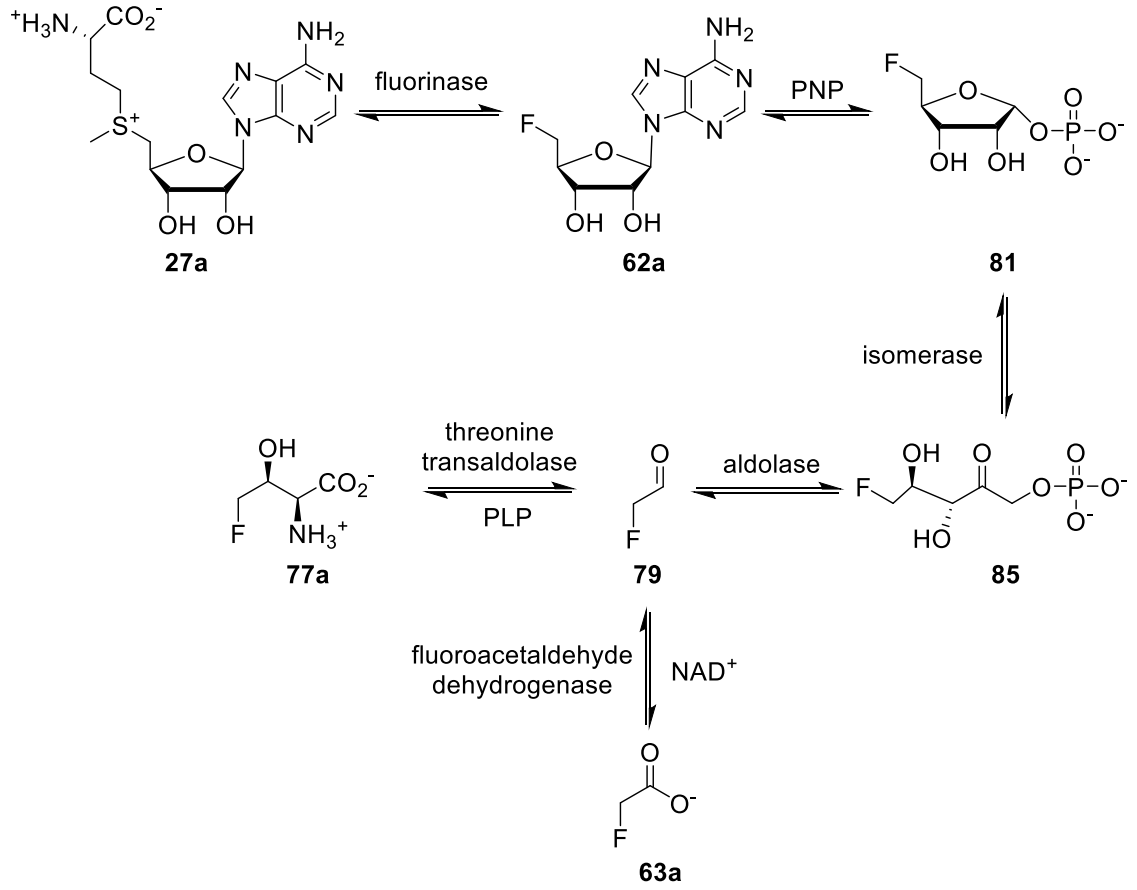
After isolation of FDRuIP **85** as an intermediate, it remained to carry out further investigations into how FDRuIP **85** was converted to fluoroacetaldehyde **79**. Ribulose phosphate analogues of FDRuIP are known to undergo a retro-aldol reaction which is catalysed by DHAP-dependent aldolases. Two aldolases (fructose 1,6-bisphosphate aldolase and putative fuculose 1-phosphate aldolase) were found in *S. cattleya*. The first enzyme produced an incorrect diastereoisomer in assays and the second could not be purified from cell-free extract experiments in *S. cattleya*.[95]



85 **79** **86a**

Scheme 27. The aldolase reaction.

However, a putative fuculose aldolase from *S. coelicolor* was over-expressed, along with all of the other enzymes (fluorinase, PNP, isomerase and fluoroacetaldehyde transaldolase) involved in the biosynthesis of fluorinated natural products. In the presence of various co-factors and reactants, this experiment produced fluorothreonine **77a** and indicated that an aldolase is responsible for the generation of fluoroacetaldehyde **79**.[96] Though the aldolase from *S. cattleya* is yet to be isolated and purified. The complete metabolic pathway of fluoroacetate **63a** and fluorothreonine **77a** is described in Scheme 28.



Scheme 28. Metabolic pathway of fluoroacetate **63a** and fluorothreonine **77a**.

The complete genome of *S. cattleya* was elucidated in 2011. This genome contains a linear chromosome of 6.28 Mb (5822 genes) and a linear mega-plasmid of 1.81 Mb (1747 genes).[97,98] The genes involved in the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a** are not all located in close proximity. The transaldolase and the putative aldolase are located on the megaplasmid, whereas the "Spencer cluster" containing the *flA* gene is located on the chromosome. The full genome map of *S. cattleya* is depicted in Figure 23, with the fluorinase gene identified as *flA* and the PNP gene identified as *flB*. In order to confirm the six enzymes involved in the metabolic pathway of fluoroacetate **63a** and fluorothreonine **77a**, specific gene knock-out mutant strains of *S. cattleya* of all the revelant genes were prepared and production of fluorometabolites was analysed. During analysis of the knock-outs of each of the first four enzymes of the metabolic pathway, no production of fluorometabolites was observed. However, the knock-out mutant of the threonine transaldolase showed production

of fluoroacetate **63a** only**,** confirming the branch point of the fluorometabolite biosynthesis pathway.[99]

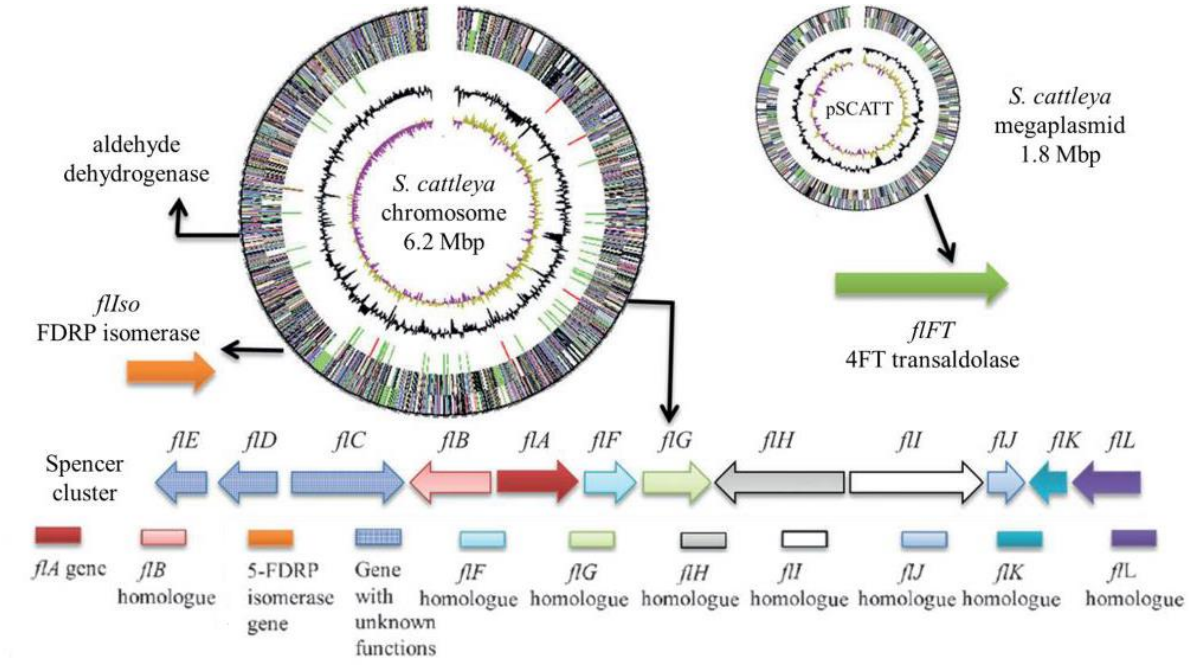


Figure 27. The full genome map of *S. cattleya*.

1.6. Conclusions and project aims

The secondary metabolism of *Streptomyces* offers a wide range of natural products, many of which have found potential application in various fields, particularly in the pharmaceutical industry.[15] To date, over 4000 halogenated natural products have been isolated from various sources in nature. However, among all these halogenated natural products, only five contain a fluorine atom.

The rarity of fluorinated natural products can be explained by the fact that biological fluorination can only proceed *via* a nucleophilic fluoride ion. Indeed, cationic and radical fluorine cannot be generated in nature.[21] Fluoride in aqueous medium is a poor nucleophile due to its hydrogen bonding with water, which creates a tight hydration shell around the anion. Therefore, the desolvation of the fluoride ion, which requires energy, is necessary for nucleophilic fluorination in nature.

Due to the unique properties of the fluorinated natural products, investigations into new fluorometabolite producing bacteria were conducted. Recently, the production of new and unidentified fluorinated natural products by a soil bacterium named *Streptomyces* sp. MA37 was observed. Therefore, the initial aim of this work was the identification of these new fluorometabolites by both total synthesis and enzyme assay, as discussed in **Chapter 2**.

Due to the complexity within secondary metabolism, several techniques have been developed in order to elucidate the metabolic pathways involved. Until now, the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a** in *Streptomyces cattleya* represented the only fully elucidated metabolic pathway of fluorometabolites. The elucidation of this divergent metabolic pathway allowed for the discovery and isolation of the only known native fluorinating enzyme, the fluorinase.

Significant efforts have previously been deployed towards the discovery of a novel native fluorinating enzyme. According to the structure of nucleocidin **78a** produced by *Streptomyces*

calvus, its biosynthesis is based on a different fluorinating enzyme than the previously discovered fluorinase. The aims of this second project was the elucidation of the metabolic pathway to nucleocidin **78a** by conducting feeding experiments as described in **Chapter 3**. The second objective was to perform cell-free extract experiments for the determination of the native substrate for this new fluorinating enzyme, as reported in **Chapter 4**.

1.7. References

1. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell*, 4th edition, New-York, Garland Science, 2002.
2. M. D. Brand, *J. Exp. Biol.*, 1997, **200**, 193-202.
3. P. A. Wright, *J. Exp. Biol.*, 1995,**198**, 273-281.
4. K. Nealson, P. Conrad, *Philos. Trans. R. Soc. Lond. B.*, 1999, **354**, 1923-1939.
5. M. Cox, D. L. Nelson, *Lehninger Principles of Biochemistry*, Palgrave Macmillan, 2004.
6. K. Tamano, *Front. Microbiol.*, 2014, **5**, 718.
7. A. L. Demain, A. Fang, *Adv. Biochem. Eng. Biotechnol.*, 2000, **69**, 1-39.
8. J. B. McAlpine, *J. Nat. Prod.*, 2009, **72**, 566-572.
9. D. C. Coleman, D. A. Crossley, *Fundamentals of Soil Ecology*, 2nd Ed., Academic Press, 2004.
10. P. A. Paul, F. E. Clark, *Soil Microbiology and Biochemistry*, 2nd Ed., Academic Press, 1996.
11. C. C. Thornburg, T. M. Zabriskie, K. L. McPhail, *J. Nat. Prod.*, 2010, **73**, 489-499.
12. T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, Norwich, The John Innes Foundation, 2000.
13. D. H. Williams, M. J. Stone, P. R. Hauck, S.K. Rahman, *J. Nat. Prod.*, 1989, **52**, 1189-1208.
14. R. A. Mapleston, M. J. Stone, D. H. Williams, *Gene*, 1992, **115**, 151-157.
15. A. Hasani, A. Kariminik, K. Issazadeh, *Int. J. Adv. Biol. Biomedical. Res.*, 2014, **2**, 63-75.
16. J. S. Kahan, M. F. 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, J. Birnbaum, *J. Antibiot.*, 1979, **32**, 1-12.

17. P. J. Cassidy, G. Albers-Schonberg, R. T. Goegelman, T. Miller, B. Arison, E. O. Stapley, J. Birnbaum, *J. Antibiot.*, 1981, **34**, 637-648.

18. S. Namiki, K. Kangouri, T. Nagate, H. Hara, K. Sugita, S. Omura, *J. Antibiot.*, 1982, **35**, 1234-1236.

19. G. W. Gribble, *J. Chem. Edu.*, 2004, **81**, 1441-1449.

20. D. H. Williams, *Nat. Prod. Rep.*, 1996, **13**, 469-477.

21. C. S. Neumann, D. G. Fujimori, C. T. Walsh, *Chem. Biol.*, 2008, **15**, 99-109.

22. D. G. Fujimori and C. T. Walsh, *Curr. Opin. Chem. Biol.*, 2007, **11**, 553-560.

23. D. Galonic, E. Barr, C. Walsh, J. Bollinger Jr, and C. Krebs, *Nat. Chem. Biol.*, 2007, **3**, 113-116.

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

25. R. D. Libby, T. M. Beachy, and A. K. Phipps, *J. Biol. Chem.*, 1996, **271**, 21820-21827.

26. J. Littlechild, E. Garcia Rodriguez, and M. Isupov, *J. Inorg. Biochem.*, 2009, **103**, 617-621.

27. A. Butler, J. N. Carter-Franklin, *Nat. Prod. Rep.*, 2004, **21**, 180-188.

28. S. Keller, T. Wage, K. Hohaus, M. Hölzer, E. Eichhorn, and K.-H. van Pée, *Angew. Chem. Int. Ed.*, 2000, **39**, 2300-2302.

29. A. S. Eustaquio, F. Pojer, J. P. Noel, *Nat. Chem. Biol.*, 2008, **4**, 69-74.

30. J. A. K. Howard, V. J. Hoy, D. O'Hagan, G. T. Smith, *Tetrahedron*, 1996, **52**, 12613-12622.

31. M-C. Chapeau, P. A. Frey, *J. Org. Chem.*, 1994, **59**, 6994-6998.

32. C. Walsh, *Tetrahedron*, 1982, **38**, 871-909.

33. M. H. Gelb, J. P. Svaren, R. H. Abeles, *Biochemistry*, 1985, **24**, 1813-1817.

34. B. E. Smart, *J. Fluorine Chem.*, 2001, **109**, 3-11.

35. B. D. Roth, *Prog. Med. Chem.*, 2002, **40**, 1-22.

36. A. Fung, Z. Jin, N. Dyatkina, G. Wang, L. Belgelman, J. Deval, *Antimicrob. Agents Chemother.*, 2014, **58**, 3636-3645.

37. A. J. Elliot, Organofluorine Chemistry: Principles and Commercial Applications, 1994, 145-157.

38. J-P. Bégué, D. Bonnet-Delpon, *Chimie bioorganique et médicinale du fluor*, CNRS Editions, 2005, Avant-propos.

39. T. Zhou, W. Zhou, Q. Wang, P. L. Dai, F. Liu, Y. L. Zhang, J. H. Sun, *Pestic. Biochem. Physiol.*, 2011, **100**, 35-40.

40. T. G. E. Davies, L. M. Field, P. N. R. Usherwood, M. S. Williamson, *Life*, 2007, **59**, 151-162.

41. E. Kissa, *Fluorinated Surfactants and Repellents*; Marcel Dekker: New-York, NY, USA, 2001.

42. M. Schultz, D. Barofsky, J. Field, *Environ. Eng. Sci.*, 2004, **20**, 487-501.

43. T. Furuya, C. A. Kuttruff, T. Ritter, *Curr. Opin. Drug Discovery Dev.*, 2008, **11**, 803-819.

44. T. Umemoto, S. Fukami, G. Tomizawa, K. Harasawa, K. Kawada, K. Tomita, *J. Am. Chem. Soc.*, 1990, **112**, 8563-8575.

45. T. D. Beeson, D. W. C. MacMillan, *J. Am. Chem. Soc.*, 2005, **127**, 8826-8828.

46. R. E. Banks, S. N. Mohialdine-Khaffaf, G. Sankar Lal, I. Sharif, R. G. Syvret, *J. Chem. Soc.*, *Chem. Commun.*, 1992, 595-596.

47. R. E. Banks, *J. Fluor. Chem.*, 1998, **87**, 1-17.

48. Y. Yoshida, Y. Kimura, *Chem. Lett.*, 1988, **17**, 1355-1358.

49. H. Sun, S. G. DiMagno, *J. Am. Chem. Soc.*, 2005, **127**, 2050-2051.

50. G. A. Olah, M. Nojima, I. Kerekes, *Synthesis*, 1973, **12**, 780-783.

51. W. J. Middleton, *J. Org. Chem.*, 1975, **40**, 574-578.

52. A. L'Heureux, F. Beaulieu, C. Bennett, D. R. Bill, S. Clayton, F. LaFlamme, M. Mirmehrabi, S. Tadayon, D. Tovell, M. Couturier, *J. Org. Chem.*, 2010, **75**, 3401-3411.

53. T. Umemoto, R. P. Singh, Y. Xu, N. Saito, *J. Am. Chem. Soc.*, 2010, **132**, 18199-18205.

54. J. G. Stark, H. G. Wallace, *Chemistry Data Book*, 2nd Ed., Hodder Education, 1989.

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

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

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

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

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

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

61. M. A. Vincent, I. H. Hellier, *Chem. Commun.*, 2006, 5902-5903.

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

63. J. S. C. Marais, *Onderstepoort J. Vet. Sci. Anim. Ind.*, 1943, **18**, 203-206.

64. J. S. C. Marais, *Onderstepoort J. Vet. Sci. Anim. Ind.*, 1944, **20**, 67-73.

65. F. L. M. Pattison, *Toxic aliphatic fluorine compounds*, 1959, Elsevier, Amsterdam.

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

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

68. R. A. Peters, *Proc. R. Soc. Lond. B Biol. Sci.*, 1952, **139**, 143-170.

69. R. A. Peters, R. W. Wakelin, P. Buffa, *Proc. Roy. Soc. B*, 1953, **140**, 497-506.

70. R. Keck, H. Haas, J. Rétey, *FEBS Letts.*, 1980, **114**, 287-290.

71. J. F. Morrison, R. A. Peters, *Biochem. J.*, 1954, **58**, 473-479.

72. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert, C. D. Stout, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 13699-13703.

73. R. Z. Eanes, D. N. Skilleter, E. Kun, *Biochem. Biophys. Res. Comm.*, **46**, 1618-1622.

74. B. Tecle, J. E. Casida, *Chem. Res. Toxicol.*, 1989, **2**, 429-435.

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

76. R. A. Peters, R. J. Hall, *Biochemical Pharmacology*, 1959, **2**, 25-36.

77. R. A. Peters, R. J. Hall, P. F. V. Ward, N. Sheppard, *Biochem J.*, 1960, **77**, 17-22.

78. M. C. Walsh, W. E. Klopfenstein, J. L. Harwood, *Phytochemistry*, 1990, **29**, 3797-3799.

79. K. A. Reid, *Ph.D. Thesis*, The Queen's University of Belfast, 1994.

80. S. O. Thomas, V. L. Singleton, J. A. Lowery, R. W. Sharpe, L. M. Pruess, J. N. Porter, J. H. Mowat, N. Bohonos, *Antibiotics Annual*, 1957, 716-721.

81. K. Isono, M. Uramoto, H. Osada, M. Ubukata, H. Kusakabe, T. Komaya, N. Miyata, S. K. Sethi, J. A. McCloskey, *Nucleic Acids Symp. Ser.*, 1984, **15**, 65-67.

82. H. Osada, K. Isono, *Antimicrob. Agents Chemother.*, 1985, **27**, 230-233.

83. M-L. Sung, L. Fowden, D.S. Millington, R.C. Sheppard, *Phytochemistry*, 1969, **8**, 1227-1233.

84. I. D. Jenkins, J.P.H. Verheyden, J.G. Moffatt, *J. Am. Chem. Soc.*, 1976, **98**, 3346-3357.

85. M. R. Amin, *Ph.D. Thesis*, University of Durham, 1996.

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

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

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

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

90. C. Schaffrath, *Ph.D. Thesis*, University of St Andrews, 2002.

91. C. D. Murphy, S. J. Moss, D. O'Hagan, *Appl. Environ. Microbiol.*, 2001, **67**, 4919-4921.

92. C. D. Murphy, D. O'Hagan, C. Schaffrath, *Angew. Chem. Int. Ed.*, 2001, **40**, 4479-4481.

93. S. L. Cobb, H. Deng, J. T. G. Hamilton, R. McGlinchey, D. O'Hagan, *Chem. Commun.*, 2004, 592-593.

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

95. S. Cross, *Ph.D. Thesis*, University of St Andrews, 2008.

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

97. V. Barbe, M. Bouzon, S. Mangenot, B. Badet, J. Poulain, B. Segurens, D. Vallenet, P. Marliere, J. Wesseinbach, *J. Bacteriol.*, 2001, **193**, 5055-5056.

98. H. Y. Ou, P. Li, C. Zhao, D. O'Hagan, Z. Deng, *Nucleotide Sequence submitted to the EMBL/GenBank/DDBJ databases*, 2011.

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

2. Identification of new fluorometabolite from *Streptomyces* sp. MA37

2.1. Analogues of the fluorinase enzyme

2.1.1. Four new fluorinase enzymes from different bacterial species

In 2014, four new fluorinase enzymes were isolated from the bacterial species; *Streptomyces* sp. MA37, *Nocardia brasiliensis*, *Actinoplanes* sp. N902-109 and *Streptomyces xinghaiensis* NRRL B-24674.[1,2] These fluorinases were identified through the mining of publically available databases of bacterial genomes and each shared high homology (>80%) with the fluorinase, isolated from *S. cattleya*.[1]

N. brasiliensis is a species of *Nocardia,* a genus of Gram-positive, rod-shaped bacteria, which mimics fungi by forming small branching filaments. Like most bacteria, *Nocardia* can be found across the world where the soil is rich with organic matter. Some species of *Nocardia* are harmless to humans; however most are responsible for nocardiosis, an infectious disease which can be acquired by either inhalation or, more rarely, through traumatic introduction. This disease affects either the lungs (*pulmonary nocardiosis*) or the whole body (*systemic nocardiosis*).[3,4] Consequently, it can be fatal for patients with weak immune systems, for example: children, the elderly and immunocompromised people. In 2012, the genome sequence of this hospital pathogen became publically available.[5,6] Comparative genome analysis showed a new fluorinase enzyme which possessed 81% similarity to the fluorinase from *S. cattleya*. Unfortunately, cultures of *N. brasiliensis* in the lab were not able to produce either fluoroacetate **63a** or 4-fluoro-L-threonine **77a**, suggesting that this new fluorinase enzyme possesses a latent activity within this organism.[1]

Actinoplanes sp. N902-109 is a bacterium possessing aerial mycelia and spherical, motile spores.[7] As with *N. brasiliensis*, the full genome sequence of *Actinoplanes* sp. N902-109 has been sequenced, and this became publically available in 2013.[8,9] A fluorinase gene was identified within the genome, and this was subsequently over-expressed in *E. coli*.[1] The enzyme possessed 80% similarity to the fluorinase enzyme from *S. cattleya*.[10] Unfortunately, the organism was not placed in the public domain so it has not been determined whether this soil bacterium is capable of producing fluoroacetate **63a** or 4-fluoro-L-threonine **77a** in culture; however the over-expressed enzyme was found to be a functional fluorinase.[1]

Streptomyces xinghaiensis NRRL B-24674 is a marine bacterium isolated in 2009 around Xinghai Bay, in Dalian, China.[11,12] Its full genome was sequenced in 2011 and this showed the presence of a fluorinase gene sharing a high sequence identity (84%) to that of *S. cattleya*.[2] Fermentation of *Streptomyces xinghaiensis* NRRL B-24674 in the presence of inorganic fluoride was carried out but did not show fluorometabolites production. However, when this marine bacterium was incubated in the presence of potassium fluoride and artificial sea salt, the production of fluoroacetate **63a** was observed.[2] *Streptomyces xinghaiensis* NRRL B-24674 is the only microorganism from the marine environment known to produce a fluorinated natural product.

Additionally, a fourth fluorinase was discovered from the bacterium *Streptomyces* sp. MA37, isolated from a soil sample collected in Legon, Ghana, Dr. Kwaku Kyeremeh from the University of Ghana and Dr. Deng from the University of Aberdeen. The full genome sequence of this soil bacterium was determined in 2011, revealing that its fluorinase possessed 87% similarity to the fluorinase from *S. cattleya*. Cultures of *S.* sp. MA37 grown by Dr. Deng from the University of Aberdeen showed the production of fluoroacetate **63a** and 4-fluoro-L-threonine **77a**, along with novel, as yet unidentified, fluorometabolites, as described in Figure 2.[1]

Figure 1. Photograph of Dr. Deng (second from left) in Ghana where *Streptomyces* sp. MA37 was found.

These findings suggested that the unknown fluorometabolites observed from *S.* sp. MA37 cultures, were produced *via* a metabolic pathway related to that of fluoroacetate **63a** and 4-fluoro-L-threonine **77a**. Additionally it was found that the concentration of fluoroacetate **63a** and 4-fluoro-L-threonine **77a** was greater than that observed for the other unidentified fluorinated natural products. The chemical shifts of 4-fluoro-L-threonine **77a** and the unidentified fluorometabolites are in a similar region of the ^{19}F-NMR spectrum, suggesting common fluoromethyl groups. Moreover, 4-fluoro-L-threonine **77a** and the other unidentified fluorinated natural products share the same multiplicity with different coupling constants. Therefore, it was suggested these unidentified fluorometabolites have the same metabolic origin as fluoroacetate **63a** and 4-fluoro-L-threonine **77a** but there is a divergence creating two competing metabolic pathways (Figure 2).

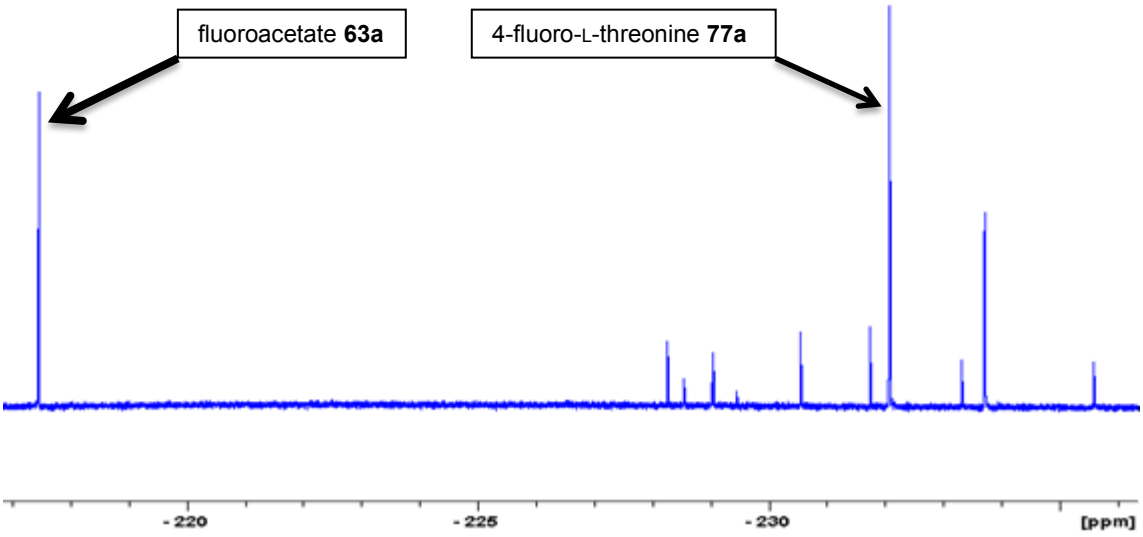


Figure 2. ^{19}F-NMR spectrum of an extract of *Streptomyces* sp. MA37 culture.

2.1.2. Genetic and kinetic studies

The amino acid sequences of the newly discovered fluorinase enzymes revealed that all of the active site residues were highly conserved. The three fluorinase enzymes from *S.* sp. MA37, *N. Brasiliensis*, *Actinoplanes* sp. N902109 were overexpressed in *E. coli* and were shown to be capable of catalysing the reaction between fluoride anion and SAM **27a** forming the product FDA **62a**. However, all were identified as slow enzymes, with turnover numbers (k_{cat}) of less than 0.3 min^{-1}, as described below in Figure 3.[1] The three newly identified enzymes possess close k_{cat} values, each of which were a little higher than that of the fluorinase isolated from *S. cattleya*. The catalytic efficiency (k_{cat}/Km) is a constant that indicates the relative efficiencies. It was found that new fluorinase enzymes are all very slow such as that of *Streptomyces cattleya*.

	SAM *K*m (μM)	Turnover number k_{cat} (min^{-1})	Catalytic efficiency k_{cat}/Km ($mM^{-1}min^{-1}$)
S. cattleya	29.2	0.083	2.84
***S.* sp. MA37**	82.4	0.262	3.18
N. Brasiliensis	27.8	0.122	4.40
***Actinoplanes* sp. N902109**	45.8	0.204	4.44

Figure 3. Kinetic data of the four fluorinase enzymes.

The organisation of the biosynthetic genes responsible for fluorometabolite production by the four species, *Streptomyces* cattleya, Streptomyces sp. MA37, *Nocardia brasiliensis* and *Actinoplanes* sp. N902-109, was investigated (see Figure 4, image reproduced from O'Hagan *et al*.[1]). The genome maps of all three bacteria, from which the new fluorinase enzymes were discovered, shared high homology to fluorometabolite biosynthetic genes found in *S. cattleya*. In Figure 4, gene maps are centred around the fluorinase encoding gene *flA*. The genes encoding the FDRP isomerase and the fluorothreonine transaldolase are both located remote from the Spencer cluster in *S. cattleya*. However, in the three other organisms these genes were found to be located close to the gene encoding for the fluorinase, suggesting a greater evolutionary regulation of fluorometabolite biosynthesis.

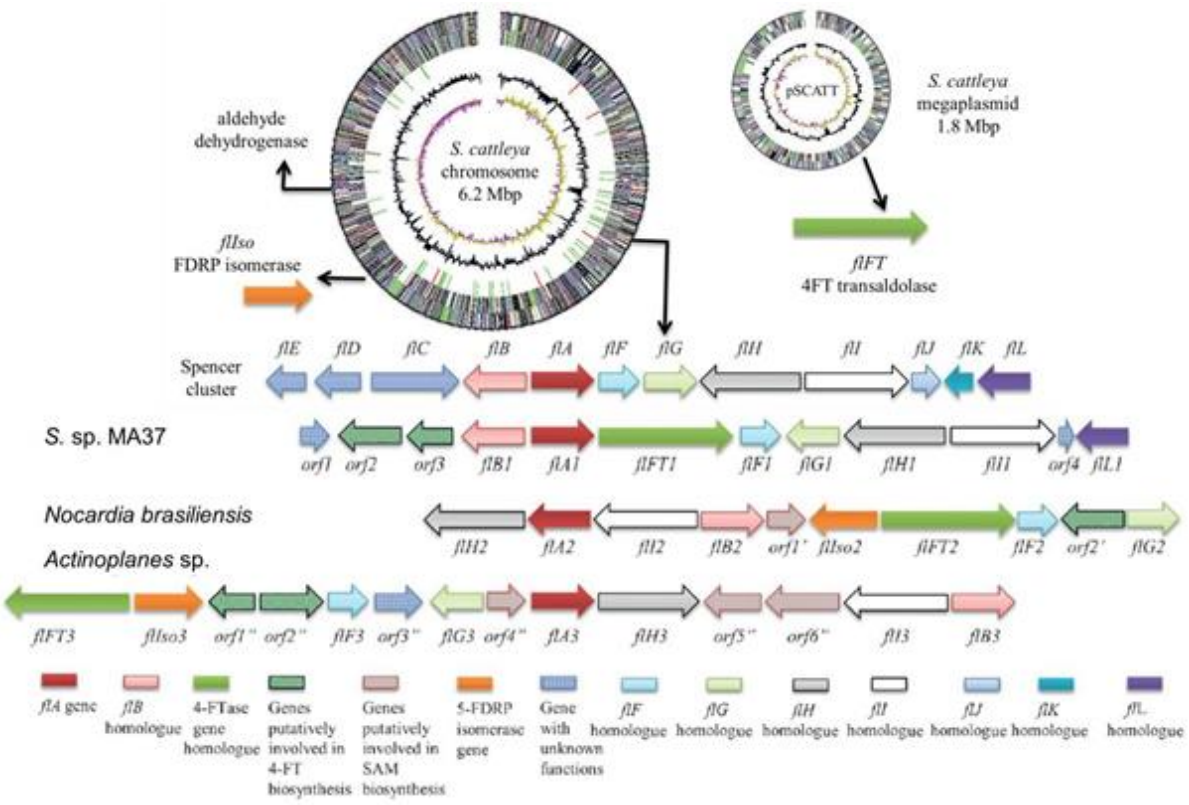


Figure 4. Fluorometabolite biosynthesis gene organisations in *S. cattleya*, *S.* sp. MA37, *N. brasiliensis*, *Actinoplanes* sp. N902109.

2.2. Analogy between *Streptomyces* sp. MA37 and *Salinospora tropica*

2.2.1. A chlorinated natural product: salinosporamide A 11

The full genome sequence of *S.* sp. MA37 was found to possess a high homology with that of *Salinispora tropica* and *Salinispora arenicola.* These marine bacteria, which are found in ocean sediments, are known to biosynthesise the anticancer agent, salinosporamide A **11**, as shown below in Figure 5.

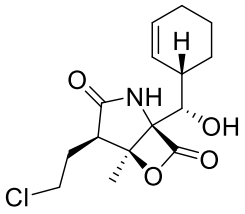


Figure 5. Structure of salinosporamide A **11**.

Salinosporamide A **11** was isolated in 2003 and has exhibited potent cancer cell cytotoxicity.[13,14] This chlorinated natural product belongs to a family of compounds collectively called salinosporamides, each of which contains a densely functionalised γ-lactam-β-lactone ring structure.[15] Its structure was confirmed by single-crystal X-ray diffraction analysis,[13] and it is very similar to *clasto*-lactacystin-β-lactone, also known as omuralide **87** (see Figure 6), which was the first inhibitor of the proteasome discovered. The proteasome is a multicatalytic proteinase complex which is responsible for most nonlysosomal protein degradation within a living cell. Omuralide **87** specifically inhibits the proteolytic activity of the 20S subunit of the proteasome, without interfering with the other protease activities.

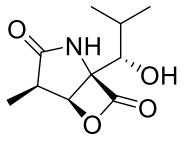


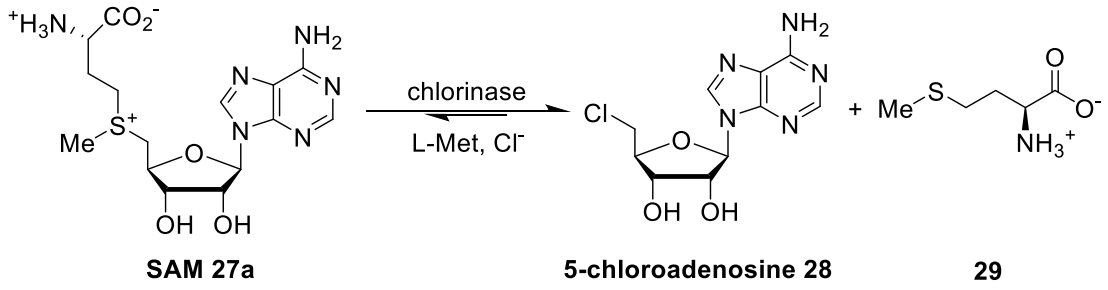
Figure 6. Structure of omuralide **87**.

Due to the structural similarity between omuralide and salinosporamide A **11**, the latter was investigated as a potential inhibitor of the proteasome. This study revealed that salinosporamide A **11** is 35 times more potent than omuralide **87**.[16] This chlorinated natural product **11** contains a cyclohexene substituent, which suggested perhaps that the inhibition mechanism of the proteasome by salinosporamide A **11** was different from the mechanism of action of omuralide **87**. Indeed, the mechanism of this inhibition was revealed to be based on inhibition of all three proteolytic functions by covalent binding to the active site threonine residues of the 20S proteasome.[13] Three years after its discovery, salinosporamide A **11**

entered phase I in human clinical trials for the treatment of multiple myeloma, which is a cancer of plasma cells.

2.2.2. Metabolic pathway of salinosporamide A 11

Salinosporamide A **11** is a chlorinated natural product and the enzyme responsible for the chlorination step during its biosynthesis was identified in 2008.[16] This enzyme shares a moderate homology of 30% to the fluorinase enzyme isolated from *Streptomyces cattleya*, and exhibits a very high substrate specific towards the specific halide ions, principally chloride ion, as fluoride ion cannot be utilised by this enzyme. This chlorinase enzyme is responsible for the conversion of SAM **27a** and chloride ion to 5'-chloro-5'-deoxyadenosine **28** (ClDA), in an equivalent manner to the function of the fluorinase enzyme, as shown below in Scheme 1.



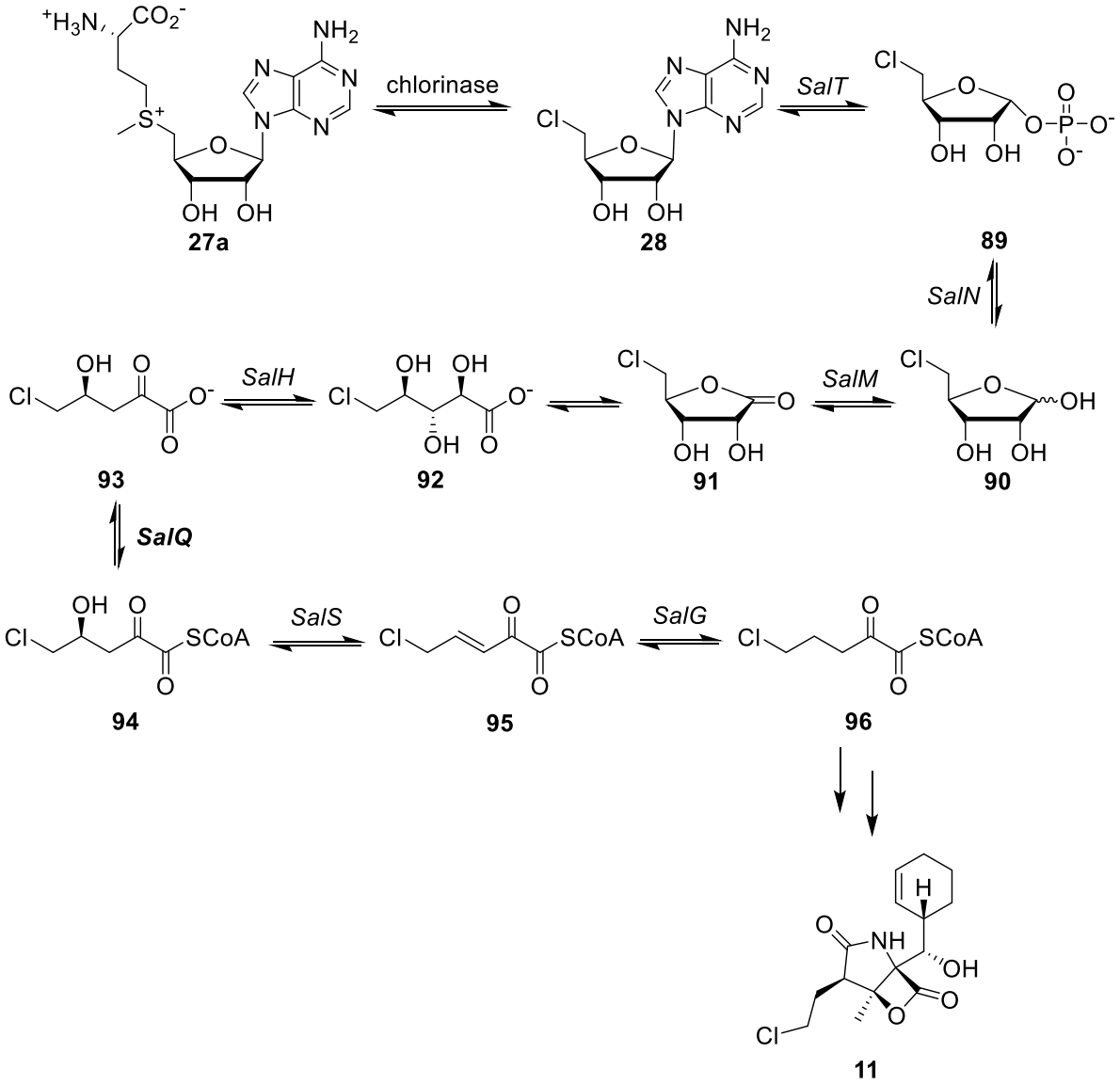
Scheme 1. Formation of a chlorine-carbon bond by the chlorinase enzyme.

Knockout experiments using *Salinispora tropica* were performed where the gene encoding the chlorinase was mutated and rendered inoperable.[16] Incubation of FDA **62a** or 5-fluoro-5-deoxyribose **88** (FDR) in chlorinase-knockout mutant of *S. tropica* revealed the production of fluorosalinosporamide A,[17] a fluorinated analogue of salinosporamide A **11**, subsequently the biological activity and cytoxicity of fluorosalinosporamide A were evaluated. Its inhibition efficiency was half that of chlorosalinosporamide A **11**, and its cytotoxicity was reduced by a factor of 20, as stated in Figure 7.[17] Consequently, a *S. tropica* mutant was designed in which the gene encoding the chlorinase enzyme was chromosomally substituted with the gene encoding the fluorinase enzyme from *S. cattleya*. The mutant organism showed the production of fluorosalinosporamide A.[18]

	Proteasome inhibition (nM)	Cytotoxicity (nM)
salinosporamide A 11	0.7	9.5
fluorosalinosporamide A	1.5	210

Figure 7. IC_{50} values for proteasome inhibition and cytotoxicity.

In 2009, the complete metabolic pathway of the chlorinated building block of salinosporamide A **11** was determined by a combination of gene deletions and chemical complementation experiments using putative intermediates.[19] The biosynthesis of the chlorinated building block involves 8 steps, beginning with the conversion of SAM **27a** to ClDA **28** by a chlorinase enzyme. ClDA **28** is then transformed into 5-deoxy-5-chloro-D-ribose-1-phosphate **89** (ClDRP) by *SalT*, a purine nucleoside phosphorylase (PNP) responsible for the displacement of the purine moiety by phosphate group. In contrast to the biosynthesis of fluorometabolites in *S. cattleya*, ClDRP **89** undergoes hydrolysis of its phosphate group affording 5-chloro-5-deoxyribose **90** (ClDR). Oxidation of the anomeric alcohol of ClDR **90** is followed by an elimination reaction generating a chlorinated α-ketoacid **92** which is then subjected to oxidative decarboxylation producing the corresponding CoA derivative **94**. The latter **94** undergoes an elimination reaction followed by hydrogenation affording the chlorinated intermediate **96** as described below in Scheme 2.

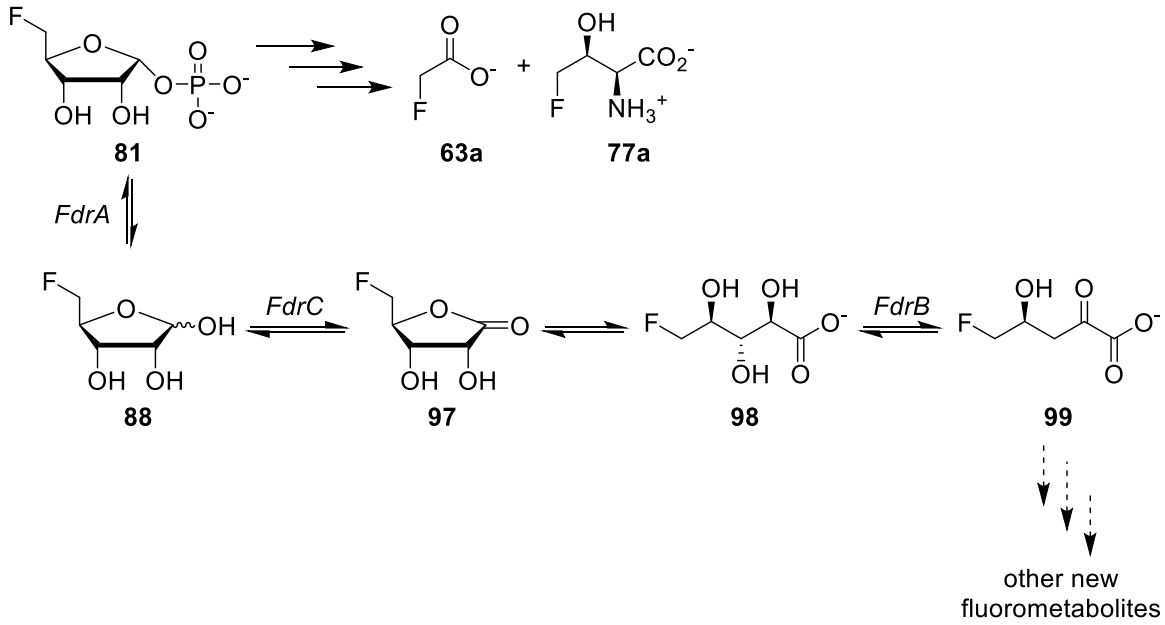


Scheme 2. Metabolic pathway of salinosporamide A **11**.

2.2.3. Aims of the project

Comparison of the genomes of *S.* sp. MA37 and *S. tropica* revealed that they share high homology, especially for genes involved in the biosynthesis of the chlorinated intermediate **96**; *SalN*, *SalM* and *SalQ*. Genome studies of the *Streptomyces* sp. MA37 strain showed the presence of a *FdrA* gene that is thought to encode a metal-dependent phosphoesterase, which shares 56% sequence identity with *SalN* of the salinosporamide gene cluster. Moreover, *FdrB* and *FdrC* genes also share high sequence identities with *SalH* (68% identity) and *SalM* (69% identity) and are predicted to encode a hydroxy-acid dehydratase and a chain dehydrogenase,

respectively. Therefore, based on information obtained from the investigation of salinosporamide A **11** biosynthesis, a metabolic pathway of the new fluorometabolites found in *S.* sp. MA37 (which require further structural and biosynthetic investigation) was proposed, as shown in Scheme 3. 5-Deoxy-5-fluoro-D-ribose-1-phosphate **81** was assumed to be a common intermediate to fluoroacetate **63a**, 4-fluoro-L-threonine **77a** and the new, as yet unidentified, fluorometabolites.



Scheme 3. Proposed metabolic pathway of new fluorometabolites in *S.* sp. MA37.

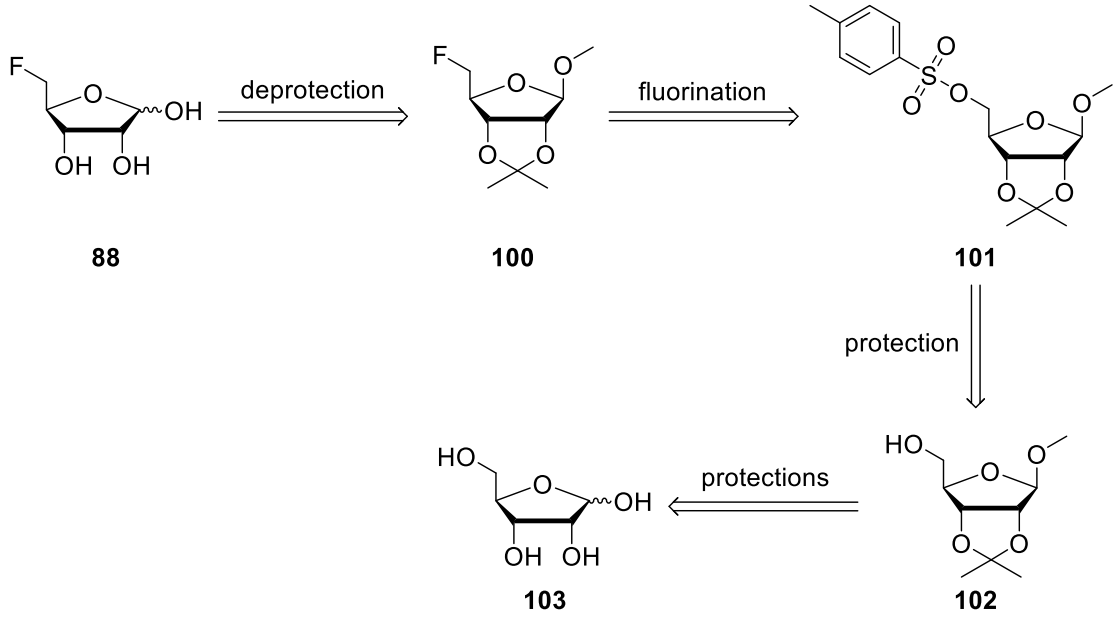
The enzyme FdrC was over-expressed by Dr. Long Ma from the University of St Andrews by designing a plasmid containing both an antibiotic resistance gene and the desired gene encoding for *FdrC*. Over-expression of the FdrC enzyme was accomplished, employing the same conditions used for the over-expression of fluorinase.[20]

The initial aim of this project was to carry out enzyme assays with FdrC. In order to achieve this, the synthesis of FDR **88** was required. Absolute stereochemistry of the product of the enzyme FdrC catalysed reaction was verified through the chemical synthesis of (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid **98**. Fluorolactone **97** was also prepared by synthesis, in order to investigate whether the enzyme FdrC catalyses the conversion of this lactone **97** to fluorocarboxylic acid **98**.

2.3. Synthesis of putative substrates of the FdrC enzyme

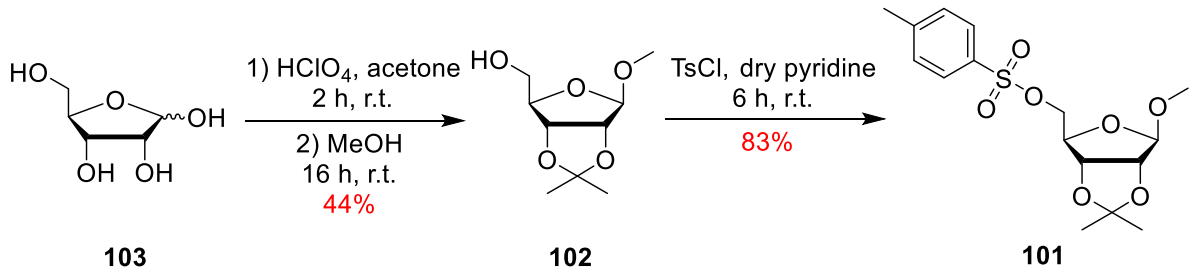
2.3.1. Synthesis of 5-fluoro-5-deoxyribose 88

An initial aim was to synthesise FDR **88**. This was accomplished by following a procedure developed by Li *et al*., (Scheme 4).[21] After preparation of FDR **88**, it was envisaged that the anomeric alcohol could be oxidized giving fluororibolactone **97**, which would then be subject to base-mediated hydrolysis to afford fluorocarboxylic acid **98**.



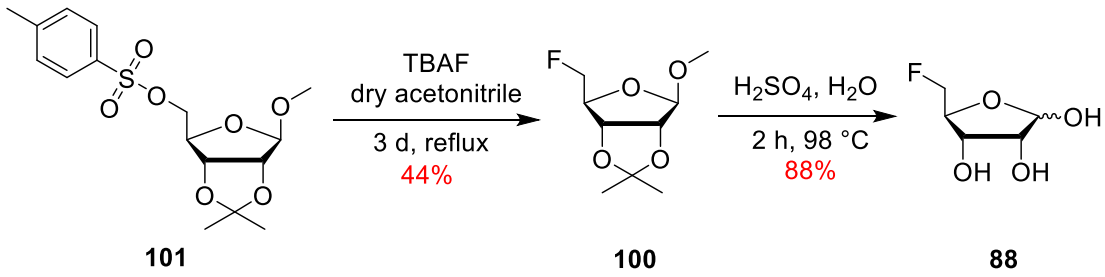
Scheme 4. Retrosynthetic strategy to FDR **88** by Li *et al*.[21]

In order to synthesis FDR **88**, an acetal protection of the 2'- and 3'-hydroxyl groups of D-ribose **103** was accomplished using acetone and DMP in the presence of catalytic perchloric acid. This was followed by methanolysis of the anomeric hydroxyl-group *via* addition of MeOH to the acidic mixture.[22] Acetonide **102** was isolated in 44% yield. Tosylation of the primary hydroxyl group of the ribose **102** was then achieved with a good conversion using tosyl chloride in dry pyridine. This reaction mixture was then purified by column chromatography affording **101** in 83% yield, as illustrated in Scheme 5.



Scheme 5. Protection of D-ribose **103** and tosylation of **102**.

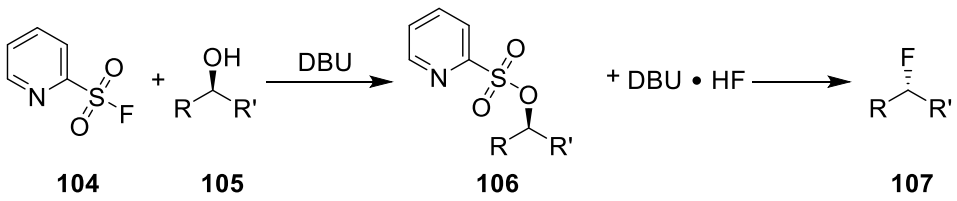
Fluorination of tosylate **101** was achieved using an excess of TBAF **53** (5 eq.) in dry acetonitrile. The conversion was improved by increasing the reaction temperature from room temperature to reflux (22% to 44%). In an effort to further optimise this conversion, the reaction was carried out using a range of TBAF **53** concentrations; however both the yield and the conversion could not be further improved. This fluorination reaction was monitored by ^{19}F-NMR and after three days the conversion did not progress beyond 66%. Deprotection to reveal the hydroxyl groups of **100** was then achieved utilising aqueous sulphuric acid. The reaction mixture was quenched by addition of $BaCO_3$, filtered and the product freeze-dried to give a mixture of α and β anomers of FDR **88** in 0.4:1 ratio (Scheme 6). This material was subjected to enzymatic assay studies with FdrC enzyme and the results will be described later in section **2.4**.



Scheme 6. Synthesis of FDR **88**.

In 2015, a mild and selective deoxyfluorination reagent with increased stability, with respect to conventional reagents, was reported. PyFluor **104** was prepared from 2-mercaptopyridine. 2-Mercaptopyridine is oxidised with aqueous sodium hypochlorite to form 2-pyridinesulfonyl chloride, and was then treated with potassium bifluoride to afford PyFluor **104**.[23] This reagent

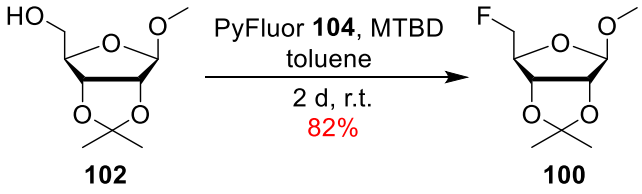
claims to be a viable and inexpensive alternative to the reagents such as DAST **57** due to its selectivity against elimination and ease of reaction work-ups. Fluorination reactions with PyFluor **104** typically require a longer reaction time, along with the use of strong bases such as DBU and MTBD. A mode of action for PyFluor **104** has been proposed and is shown below in Scheme 7.



104 **105** **106** **107**

Scheme 7. Proposed mechanism of fluorination reaction with PyFluor **104**.

PyFluor **104** was used for the optimisation of the synthesis of FDR **88**. Indeed, this reagent allows the fluorination of acetonide **102**, in the presence of MTBD as a base, in a very good yield (82%) as depicted in Scheme 8. The overall yields of both strategies are approximately 75% but activation of the 5-hydroxyl of **102** into tosylate **101** was not necessary with the direct use of PyFluor **104**.



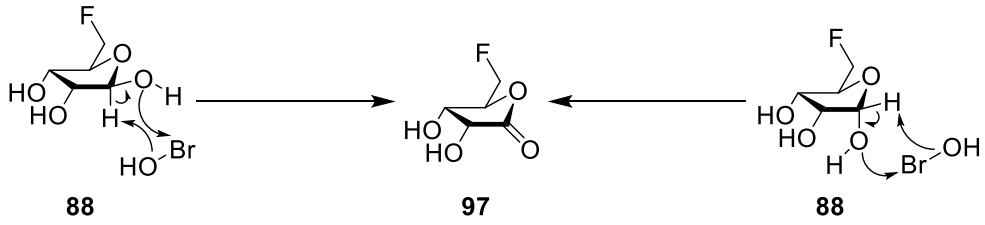
102 **100**

Scheme 8. Fluorination reaction with PyFluor **104**.

2.3.2. Synthesis of (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid 98 from 5-fluoro-5-deoxyribose 88

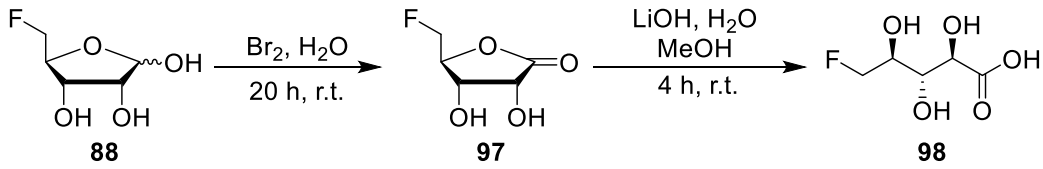
With FDR **88** in hand, oxidation at the anomeric position was completed by employing bromine in water. At neutral pH, previous kinetic studies indicated that the mechanism of such oxidations involve the formation of hypobromous acid.[24-26] These mild conditions allow the selective oxidation of the anomeric alcohol of FDR **88** as illustrated in Scheme 9. ^{19}F-NMR of

the crude product revealed the formation of the fluororibolactone **97** along with other unidentified side-products. Purification of the mixture by C-18 cartridge or ion exchange column proved to be difficult.



Scheme 9. Mechanism of oxidation of α and β FDR **88** with Br_2.

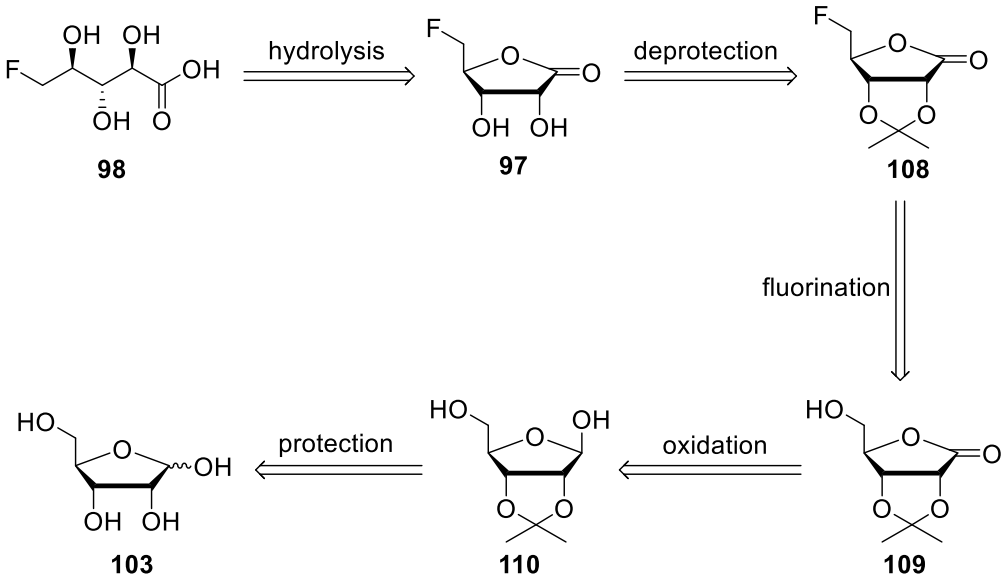
It was therefore decided to carry out a one pot reaction with direct hydrolysis of fluorolactone **97** in aqueous LiOH in the hope that purification of the newly formed carboxylic acid **98** could be more readily achieved (Scheme 10).[27] This approach proved somewhat successful; however the fluorocarboxylic acid **98** was contaminated with side products, which could not be readily purified using a C-18 cartridge or ion exchange column.



Scheme 10. Synthesis approach to fluorometabolite **98**.

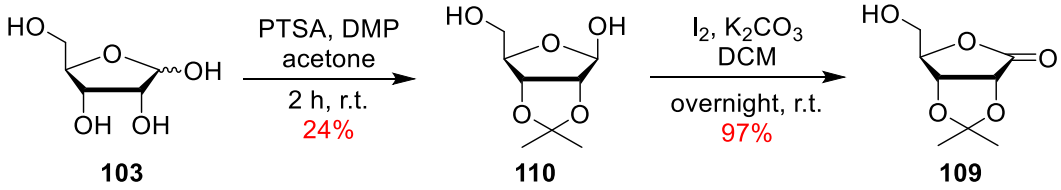
2.3.3. Synthesis of (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid 98

Due to the challenge in purifying (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid **98**, an alternative approach was utilised, in which oxidation of the anomeric alcohol of the ribose ring was carried out before the fluorination reaction as illustrated in Scheme 11. The last two steps of this new strategy, which are acetal deprotection and ring-opening hydrolysis, are known to allow for a less challenging purification when compared to purifications involved in the previous strategy.



Scheme 11. Second retrosynthetic strategy to fluorometabolite **98**.

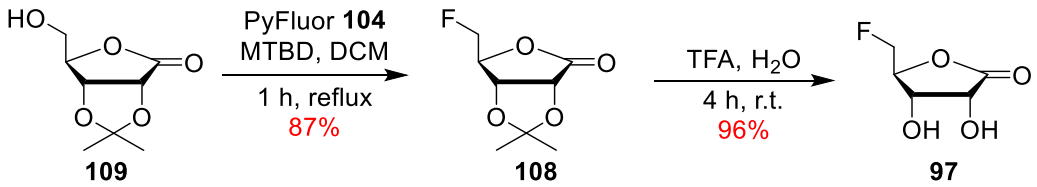
The 2'- and 3'-hydroxyl groups of D-ribose **103** were subjected to acetonide protection, using DMP in the presence of a catalytic amount of PTSA in acetone.[28] Only the β anomer **110** was obtained in low yield after purification by column chromatography. Oxidation of the anomeric alcohol was then achieved with iodine in DCM in the presence of potassium carbonate, affording **109** in good yield, as shown below in Scheme 12. This reaction, along with the oxidation using bromine in water, was chemoselective as only the anomeric alcohol was oxidised.



Scheme 12. Protection of D-ribose **103** and chemoselective oxidation of **110**.

Fluorination of the acetonide protected lactone **109** with Deoxofluor® **58** in DCM, at room temperature, afforded the protected fluorolactone **108** in good yield.[29] In contrast to the first strategy; this route did not require formation of the tosylate as a leaving group. Indeed, the overall yield of the tosylation and the fluorination in the previous strategy was 37%. To

conclude, this strategy proved to be quicker, shorter and the yield of the fluorination step improved significantly. The yield of this fluorination reaction was improved from 52% to 87% by using the reagent PyFluor **104** instead of Deoxofluor® **58**, as described previously in Scheme 7.



Scheme 13. Fluorination of **109** and deprotection of acetonide **108**.

Moreover the crystal structure of the fluorinated compound **108** was determined by X-ray crystallography (Figure 8). The structure revealed that the fluorine atom adopts a *gauche* conformation relative to the oxygen in the ring and an *endo* conformation with respect to the ribose ring system. According to the previous studies by Buissonneaud *et al.*[30], this *endo* conformation can be explained by the fact that the fluorine atom is partially electronegatively charged and the carbon C-1 is partially electropositively charged, therefore there is an electrostatic interaction between these two atoms. The acetonide **108** was then deprotected using TFA in water to afford pure fluorolactone **97**. After work up there was still some residual TFA observable by ^{19}F-NMR and as such several co-evaporations with Et_2O were required to remove all of the remaining TFA (Scheme 13).

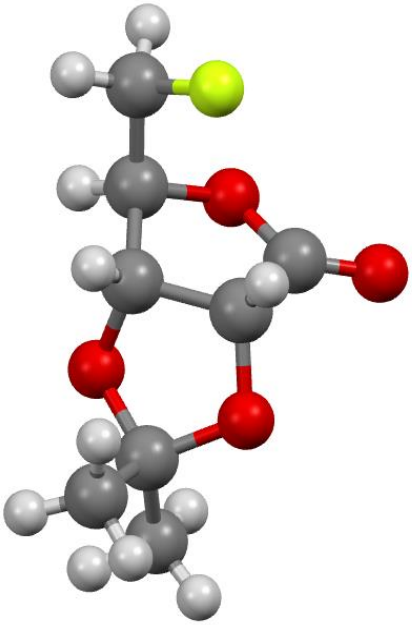
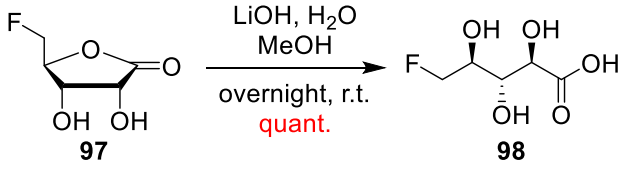


Figure 8. X-ray crystal structure of fluorolactone **108**.

Finally, the hydrolysis of fluororibolactone **97** was explored using aqueous LiOH (Scheme 14).[27] After purification, the ^{19}F-NMR revealed the presence of fluoride anion in addition of the desired product **98**.



Scheme 14. Final step of the synthesis of the novel fluorometabolite **98**.

In order to confirm that the synthetic fluorocarboxylic acid **98** exactly correlated to the fluorometabolite observed from *S.* sp. MA37 cultures, a spiking experiment was carried out, as shown below in Figure 9. ^{19}F-NMR of a culture of *S.* sp. MA37 after 19 days revealed an increased presence of new fluorometabolites. Among them there was a peak at -233.30 ppm. This ^{19}F-NMR sample was then spiked with the synthetically prepared fluorocarboxylic acid

98, and ^{19}F-NMR revealed that the peak at -233.30 ppm had increased. This experiment confirmed that fluorocarboxylic acid **98** is a fluorometabolite of *S.* sp. MA37.

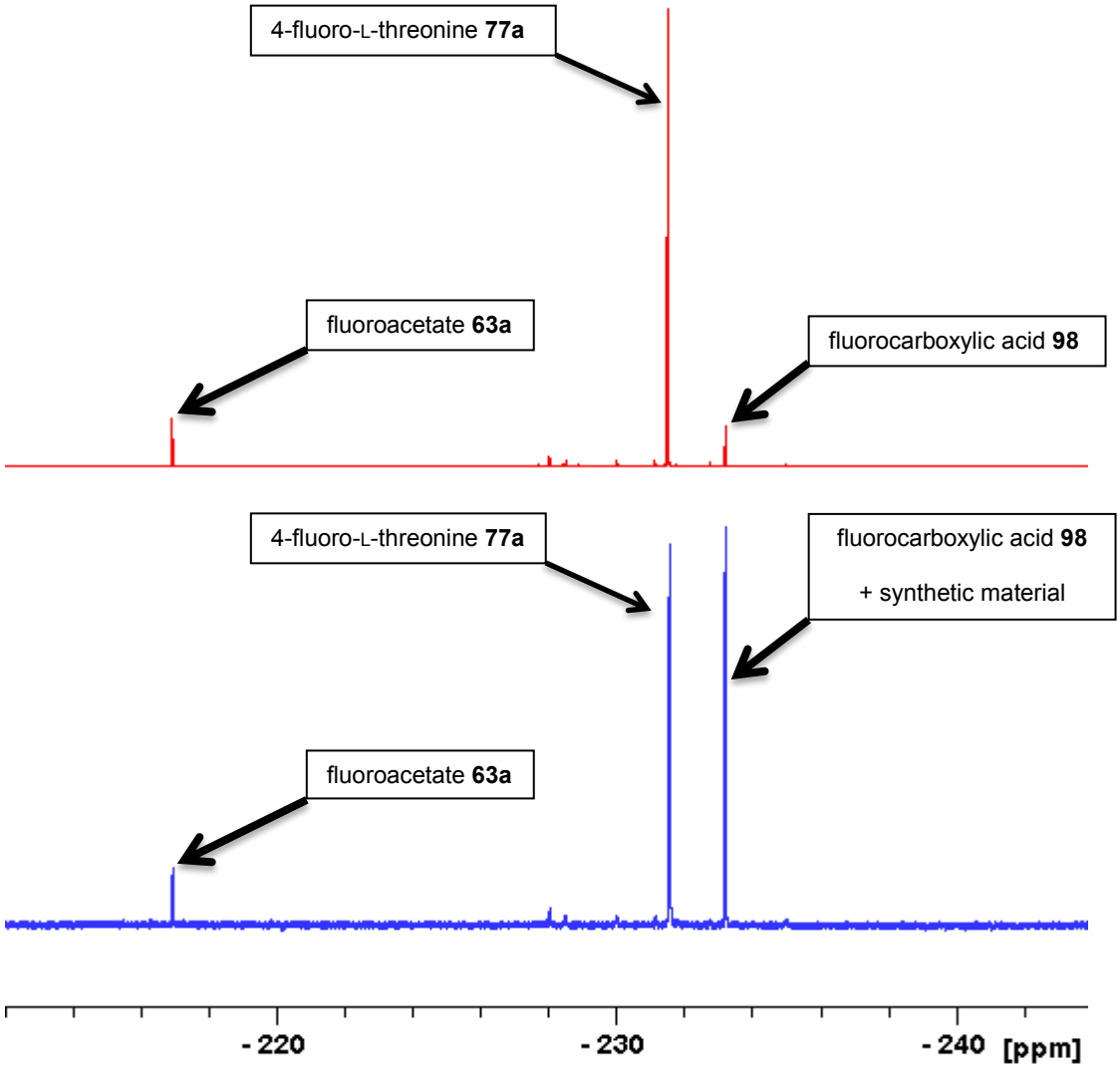


Figure 9. ^{19}F{^{1}H}-NMR spiking experiment of *Streptomyces* sp. MA37 culture after 19 days in D_2O (blue) and after adding synthetic fluorometabolite **98** in D_2O (red).

2.4. Enzymatic assays of FdrC enzyme

2.4.1. Identification of FDR 88 as a key intermediate

ClDRP **89** is the last intermediate in salinosporamide A **11** biosynthesis which possesses a corresponding fluorinated analogue (FDRP **81**) within the metabolic pathway of fluoroacetate **63a** and fluorothreonine **77a** in *Streptomyces cattleya*. Therefore, it was hypothesised that FDRP **81** is the branching point for the biosynthesis of fluoroacetate **63a**, fluorothreonine **77a** and the new fluorometabolites in *Streptomyces* sp. MA37.

To investigate this hypothesis, the role of FDR **88** in the biosynthesis of these unknown fluorometabolites was examined. FDR **88**, which was prepared by synthesis, was incubated in a cell-free extract of *Streptomyces* sp. MA37 and monitored by ^{19}F-NMR. The spectrum obtained was very different from that of whole cell incubations in the presence of inorganic fluoride. The resulting ^{19}F-NMR spectrum showed a dominance of fluoroacetate **63a** and fluorothreonine **77a** over the seven fluorometabolites produced by this organism. Indeed, in this new experiment signals of some of the 'minor fluorometabolites' had increased and subsequently signals corresponding to fluoroacetate **63a** and fluorothreonine **77a** had diminished, as shown below in Figure 10.[20] This indicates that FDR **88** may be an intermediate of some of the new fluorometabolites of *Streptomyces* sp. MA37. However, it did not upregulate the biosynthesis of fluoroacetate **63a** and fluorothreonine **77a**.

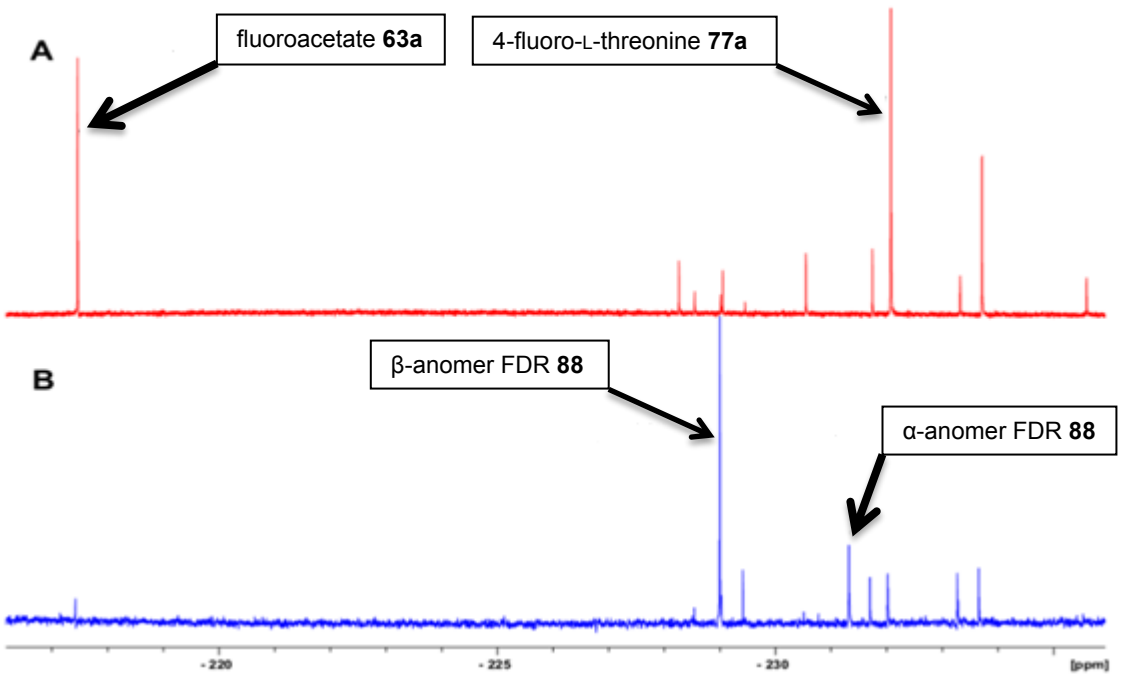


Figure 10. ^{19}F-NMR spectroscopic analysis of fluorometabolites:
(**A**) in the supernatant of culture from *Streptomyces* sp. MA37.
(**B**) in the cell free extract of MA37 strain incubated with FDR **88**.

2.4.2. Enzyme assays

With FDR **88** in hand, it became an objective to determine whether the FdrC enzyme is able to utilise FDR **88** as a substrate. Subsequently, over-expression of a codon-optimized synthetic gene for FdrC in *E. coli*, including a His_6 tag and a TEV protease cleavage site was performed in our laboratory by Dr. Long Ma. After purification of the coded protein, the molecular weight of this isolated FdrC enzyme was estimated to be approximately 31 kDa.[31]

Enzyme assays with the FdrC enzyme using various conditions were carried out and monitored by ^{19}F-NMR spectroscopy. The enzyme reaction of FdrC in the presence of FDR **88** without any co-factors did not show any conversion. Incubation of FdrC enzyme with FDR **88** and $NADP^+$ still resulted in no observable turnover. However, the enzymatic reaction in the presence of FDR **88** and NAD^+ revealed a complete conversion of FDR **88** to a new organofluorine compound at -233.15 ppm with the same multiplicity (dt) by ^{19}F-NMR. This result revealed that the FdrC enzyme is a NAD^+-dependent enzyme. According to the

proposed metabolic pathway of salinosporamide A **11** in *Salinospora tropica* (see Scheme 2), fluorocarboxylic acid was thought to be formed *via* a multi-step FdrC enzyme catalysed reaction. This reaction can reasonably proceed in two steps: firstly oxidation of the anomeric alcohol to lactone **97**, followed by a ring opening hydrolysis. Incubation of fluorolactone **97** with the FdrC enzyme in the presence of NAD^+ revealed the appearance of FHPA **98** as indicated by ^{19}F-NMR but there was no obvious presence of fluorolactone **97**, suggesting an instantaneous hydrolysis. In addition to this, enzymatic reaction products along with bacterial culture products were derivatised with *N*-methyl-*N*-trimethylsilyl trifluroacetamide (MSTFA) and analysed by GC-MS in the James Hutton Institute by Dr. Thomas Shepherd. The same protocol was followed for the derivatisation of the synthetic fluorocarboxylic acid **98** that showed an identical fragmentation and retention time, proving that they are both the same fluorometabolite. All these results are summarised in Figure 11 and Figure 12.[31]

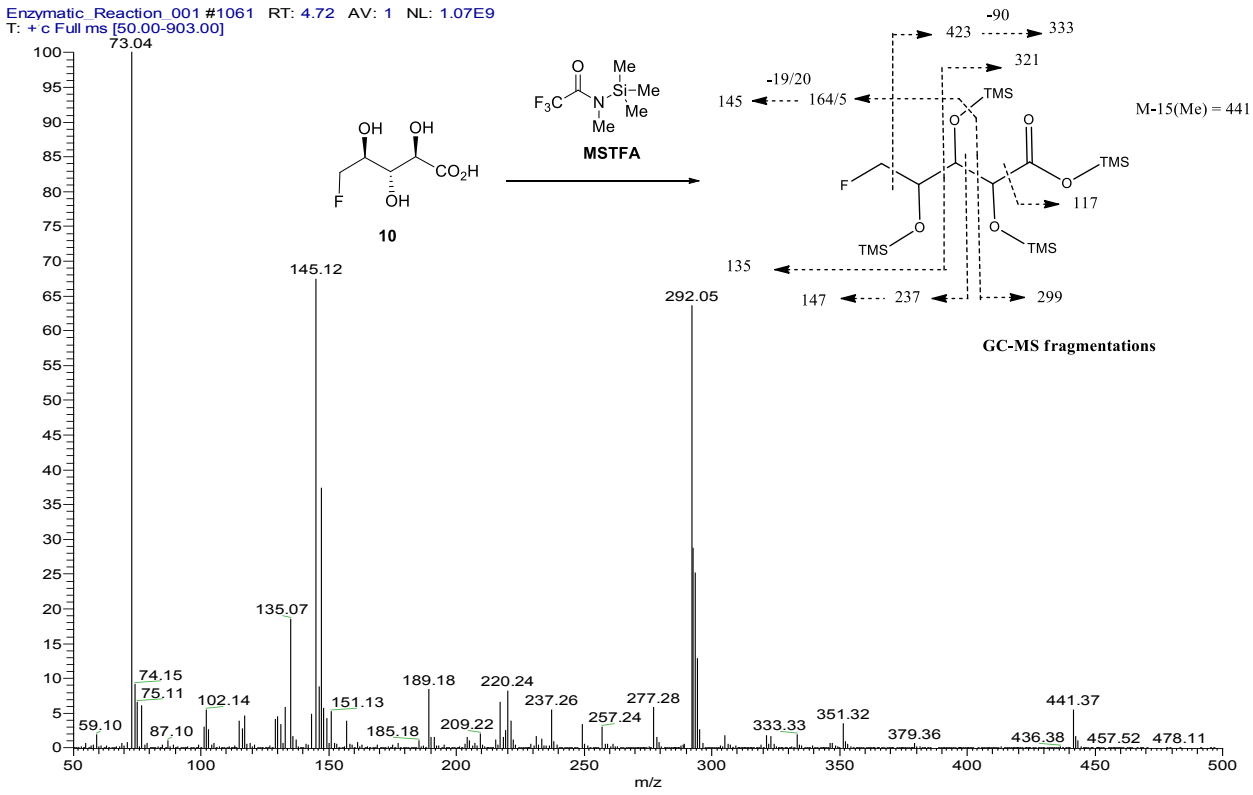


Figure 11. GC-MS fragmentation analysis of MSTFA treated extract showing the mass spectrum and fragmentation pattern of FHPA $(TMS)_4$.

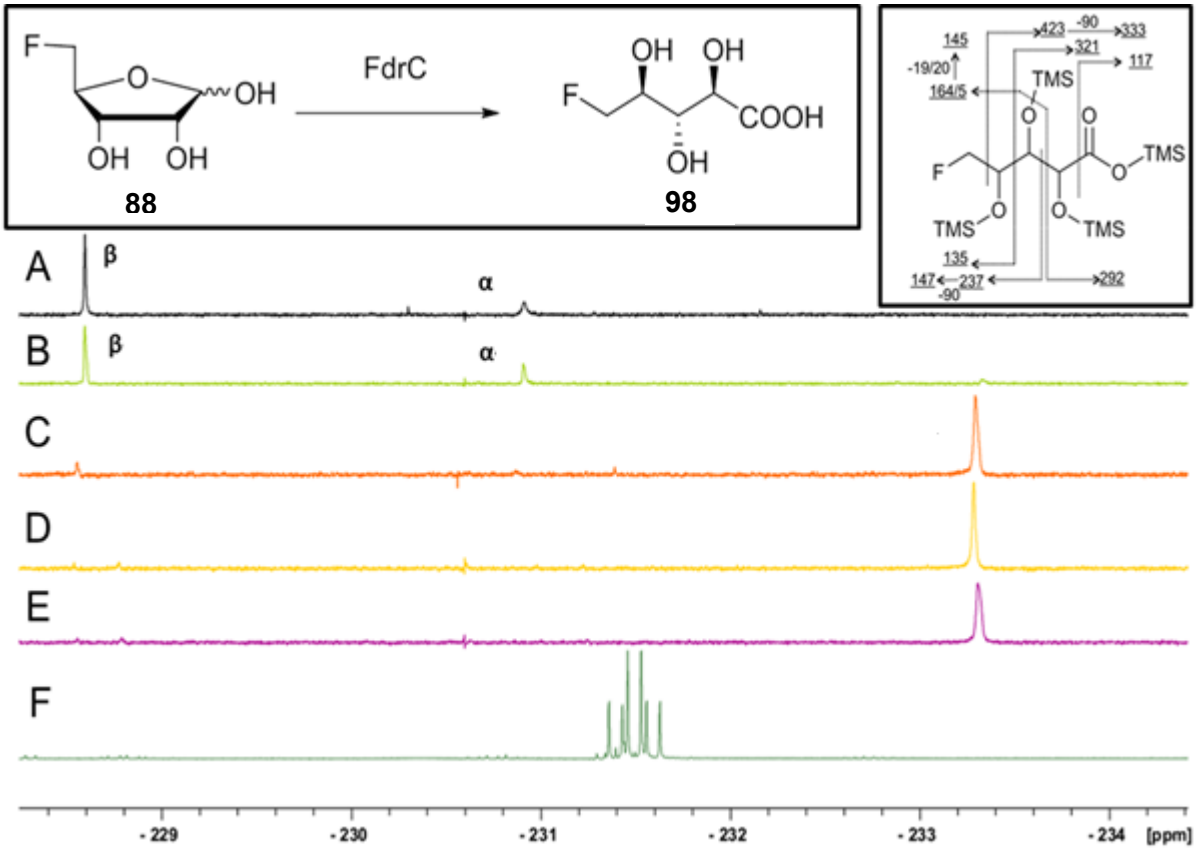


Figure 12. ^{19}F-NMR spectroscopic analysis of:
(**A**) ^{19}F{^{1}H}-NMR of FdrC after overnight incubation with FDR **88** at 37 °C.
(**B**) ^{19}F{^{1}H}-NMR of the FdrC enzymatic reaction after overnight incubation with $NADP^+$ and FDR **88** at 37 °C.
(**C**) ^{19}F{^{1}H}-NMR of the FdrC enzymatic reaction after overnight incubation with NAD^+ and FDR **88** at 37 °C.
(**D**) ^{19}F{^{1}H}-NMR of FdrC enzymatic reaction after overnight incubation with NAD^+ and fluorolactone **97** at 37 °C.
(**E**) ^{19}F{^{1}H}-NMR of aqueous solution of **97** after overnight incubation at 37 °C.
(**F**) ^{19}F-NMR of synthetic fluorolactone **97**.
Left insert: the enzymatic reaction; right insert: the GC-MS fragmentation pattern of FHPA **98**.

Kinetic studies with FdrC enzyme carried out in our laboratory by Dr. Long Ma, demonstrated that FdrC has a higher affinity for D-ribose **103** over FDR **88** as indicated by the Michaelis contant (K_m) values. However, both substrates are oxidised with a very similar efficiency (k_{cat}/K_m) as shown below in Figure 13.[20]

	***K*m (μM)**	**Turnover number k_{cat} (min^{-1})**	**Catalytic efficiency k_{cat}/Km ($\mu M^{-1}min^{-1}$)**
D-ribose	0.84	0.32	0.38
FDR	2.73	0.58	0.21

Figure 13. Kinetic data of FdrC enzyme.[31]

2.5. Conclusion

In conclusion, identification of FHPA **98** as a new fluorometabolite was achieved by synthesis. Its biosynthesis was revealed to eminate from branching of the established metabolic pathway to fluoroacetate **63a** and 4-fluoro-L-threonine **77a**. Genetic studies revealed the presence of a gene cluster containing three genes (*FdrA*, *FdrB* and *FdrC*) which share high homology with genes involved in the biosynthesis of salinosporamide A **11** in *Salinospora tropica*. Over-expression of the *FdrC* gene was carried out by Dr. Long Ma, and enzymatic assays of the FdrC enzyme demonstrated that FDR **88** is a substrate for the enzyme. FdrC enzyme is responsible for the oxidation of the anomeric alcohol to give fluorolactone **97**, followed by hydrolysis generating FHPA **98**. Synthetic synthesis of FHPA **98** allowed the confirmation of the absolute stereochemistry of the product of FdrC enzyme reaction. Moreover, GC-MS analysis was achieved after derivatisation of synthetic FHPA **98** and derivatisation of the product of the enzymatic assay. This demonstrated that they are indeed the same compound by retention time and fragmentation pattern. All of these experiments demonstrate that FHPA **98** is a new fluorinated natural product of *Streptomyces* sp. MA37 and the first novel fluorinated natural product to be identified in over two decades.

2.6. References

1. H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, D. O'Hagan, *ChemBioChem*, 2014, **15**, 364-368.
2. S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan, H. Deng, *Org. Biomol. Chem.*, 2014, **12**, 4828-4831.
3. B. J. Marsland, M. Kurrer, R. Reissmann, N. L. Harris, M. Kopf, *Eur. J. Immunol.*, 2008, **38**, 479-488.
4. J. G. Mercer, P. I. Mitchell, K. M. Moar, A. Bissett, S. Geissler, K. Bruce, L. H. Chappell, *Parasitology*, 2000, **120**, 641-647.
5. L. Vera-Cabrera, R. Ortiz-Lopez, R. Elizondo-Gonzalez, A. A. Perez-Maya, J. Ocampo-Candiani, *J. Bacteriol.*, 2012, **194**, 2761-2762.
6. L. Vera-Cabrera, R. Ortiz-Lopez, R. Elizondo-Gonzalez, J. Ocampo-Candiani, *PLoS One*, 2013, **8**, e65425.
7. C. E. Bland, N. J. Couch, *The Prokaryotes: a Handbook on Habitats, Isolation, and Identification of Bacteria*, 1981, Springer, New York, 2004-2010.
8. June 11th 2013; Actinoplanes genome http://www.ncbi.nlm.nih.gov/nuccore/494682416.
9. H. Huang, S. X. Ren, S. Yang, H. F. Hu, *Chin. J. Nat. Med.*, 2015, **13**, 90-98.
10. H. Y. Ou, P. Li, C. Zhao, D. O'Hagan, Z. Deng, *Nucleotide Sequence submitted to the EMBL/GenBank/DDBJ databases*, 2011.
11. X. Q. Zhao, W. J. Li, W. C. Jiao, Y. Li, W. J. Yuan, Y. Q. Zhang, H. P. Klenk, J. W. Suh, F. W. Bai, *Int. J. Syst. Evol. Microbiol.*, 2009, **59**, 2870-2874.
12. W. Jiao, F. Zhang, X. Zhao, J. Hu, J-W. Suh, M-J. Virolle, 2013, *PLoS ONE*, **8**, e75994.
13. R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew. Chem. Int. Ed.*, 2003, **42**, 355-357.

14. D. Chauhan, L. Catley, G. Li, K. Podar, T. Hideshima, M. Velankar, C. Mitsiades, N. Mitsiades, H. Yasui, A. Letai, H. Ovaa, C. Berkers, B. Nicholson, T-H. Chao, S. T. C. Neuteboom, P. Richardson, M. A. Palladino, K. C. Anderson, *Cancer Cell*, 2005, **8**, 407-419.
15. L. L. Beer, B. S. Moore, *Org. Lett.*, 2007, **9**, 845-848.
16. A. S. Eustaquio, F. Pojer, J. P. Noel, B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69-74.
17. A. S. Eustaquio, B. S. Moore, *Angew. Chem. Int. Ed.*, 2008, **47**, 3936-3938.
18. A. S. Eustaquio, D. O'Hagan, B. S. Moore, *J. Nat. Prod.*, 2010, **73**, 378-382.
19. A. S. Eustaquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Flovora, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *PNAS*, 2009, **106**, 12295-12300.
20. C. Dong, F. L. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan, J. H. Naismith, *Nature*, 2004, **427**, 561-565.
21. X-G. Li, S. Dall'Angelo, L. F. Schweiger, M. Zanda, D. O'Hagan, *Chem. Commun.*, 2012, **48**, 5247-5249.
22. M. P. Bosch, F. Campos, I. Niubó, G. Rosell, J. L. Díaz, J. Brea, M. I. Loza, A. Guerrero, *J. Med. Chem.*, 2004, **47**, 4041-4053.
23. M. K. Nielsen, C. R. Ugaz, W. Li, A. G. Doyle, *J. Am. Chem. Soc.*, 2015, **137**, 9571-9574.
24. H. H. Bunzel, A. P. Mathews, *J. Am. Chem. Soc.*, 1909, **31**, 464-479.
25. N. C. Deno, N. H. Potter, *J. Am. Chem. Soc.*, 1967, **89**, 3555-3556.
26. G. N. Branimir, D. L. Žugić, M. M. Gvozdenović, T. Lj. Trišović, *Carbohydr. Res.*, 2006, **341**, 1779-1787.
27. B. Dayal, G. Salen, B. Toome, G. S. Tint, S. Shefer, J. Padia, *Steroids*, 1990, **55**, 233-237.
28. J. H. Cho, D. L. Bernard, R. W. Sidwell, E. R. Kern, C. K. Chu, *J. Med. Chem.*, 2006, **49**, 1140-1148.

29. P. Nasomjai, D. O'Hagan, A. M. Z. Slawin, *Beilstein J. Org. Chem.*, 2009, **5**, 37.

30. D. Y. Buissonneaud, T. Van Mourik, D. O'Hagan, *Tetrahedron*, 2010, **66**, 2196-2202.

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

3. Isotopic labelling studies into nucleocidin 78a

3.1. Introduction

3.1.1. Previous studies

Since the initial discovery and isolation of nucleocidin **78a** from *Streptomyces calvus* in 1957,[1] its subsequent production has proved challenging. First of all, nucleocidin **78a** production is difficult to observe, this is primarily due to the low titre during the cultivation of *S. calvus*, which is approximately 1 mg in 1 L. Moreover, the biosynthesis of nucleocidin **78a** by *S. calvus* is very unpredictable, and publically deposited strains appear to have lost the ability to produce nucleocidin **78a** altogether.[2-4] Moreover, *S. calvus* is known to have a "bald" phenotype because of its lack of spores, and thus is defficient in the synthesis of secondary metabolites.

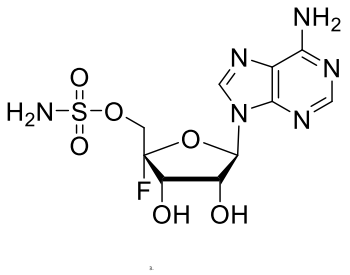


Figure 1. Structure of nucleocidin **78a**.

Over the last 50 years, extensive investigations have been carried out to elucidate the biosynthesis of nucleocidin **78a**.[2-4] However, these studies have proved challenging due to the low level or lack of production. As a result, subsequent investigations have focused on revealing the reasons behind this inconsistent production. Firstly, it has been widely accepted, though not without some opposition, that the formation of spores and aerial mycelium in *Streptomyces* is crucial for the production of secondary metabolites.[5] The absence of spores from *Streptomyces calvus* in solid media has been postulisated to be one reason for its seemingly unpredictable nucleocidin **78a** production.

Furthermore, several studies have reported that certain mutations within the *rpoB* gene, which is known to encode the RNA polymerase *β*-subunit protein, favoured the production of secondary metabolites in *Streptomyces*.[6] In one study, a wild-type strain of *S. calvus* underwent mutation of the *rpoB* gene by ultraviolet light irradiation,[2] this mutant strain was then observed to exhibit resistance to rifampicin, which is known to inhibit transcription by binding to the *β*-subunit of RNA polymerase.[7] The production of antibiotics, including nucleocidin **78a**, was revealed to be enhanced in this mutant strain.[2]

In 2013, genome sequencing of wild-type *S. calvus* identified a mutation in the *bldA* gene.[3] The *bld* genes are crucial for initiating aerial hyphae development and sporulation in solid media. The *bldA* gene is responsible for encoding a tRNA which is important for the translation of mRNA involved in nucleocidin **78a** biosynthesis. Correction of this mutation in the *bldA* gene, by complementation of a wild-type strain of *S. calvus* with a functional copy, allowed for the restoration of the production of nucleocidin **78a**.[4]

3.1.2. Production of nucleocidin 78a

For this study, an in-house strain of *S. calvus* T-3018 was provided by Pfizer. Genome sequencing of this wild-type strain of *S. calvus* was carried out in China in 2014, and comparison of its genome sequence with that of the public strain of *S. calvus* revealed that *S. calvus* T-3018 does not possess a mutation within the *bldA* gene.

As such, the first aim of the project was to establish nucleocidin **78a**-producing cultures, from glycerol stocks of *S. calvus* T-3018. *S. calvus* T-3018 was grown on solid ISP4 medium for 20 days (solid ISP2 medium was also trialed, however this did not provide sufficient nucleocidin **78** production). A specific fermentation medium, developed by Pfizer, was then used for the growth of *S. calvus* for 18 days (see section **5.3.2**).

After centrifugation, the supernatant was lyophilised and production of nucleocidin **78a** was assayed by mass spectrometry and ^{19}F-NMR. Mass spectrometry indicated that nucleocidin

78a was produced, however in ^{19}F-NMR experiments the chemical shifts of nucleocidin **78a** and fluoride ion are close, and as the fluoride peak is very broad the nucleocidin **78a** peak was obscured and could not be easily identified.

In 1957, Thomas *et al*. reported that nucleocidin **78a** could be dissolved in *n*-butanol.[1] In 2009, successful extraction of antibiotics from *S. calvus'* cultures was achieved using *n*-butanol by *Fukuda et al*.[2] The smallest volume of *n*-butanol required for complete extraction of antibiotics was found to be one fifth of the volume of the culture broth.[2]

Therefore, cultures of *S. calvus* were centrifuged and the supernatant extracted with *n*-butanol. The organic layer was concentrated and the ^{19}F-NMR spectrum recorded. This method allowed for the separation, and the resolution, of nucleocidin **78a** and fluoride ion peaks as shown in Figure 2. The production of nucleocidin **78a** during these experiments was very low, and due to this the number of scans for ^{19}F-NMR was increased to a minimum of 3600, over a 2 h acquisition time.

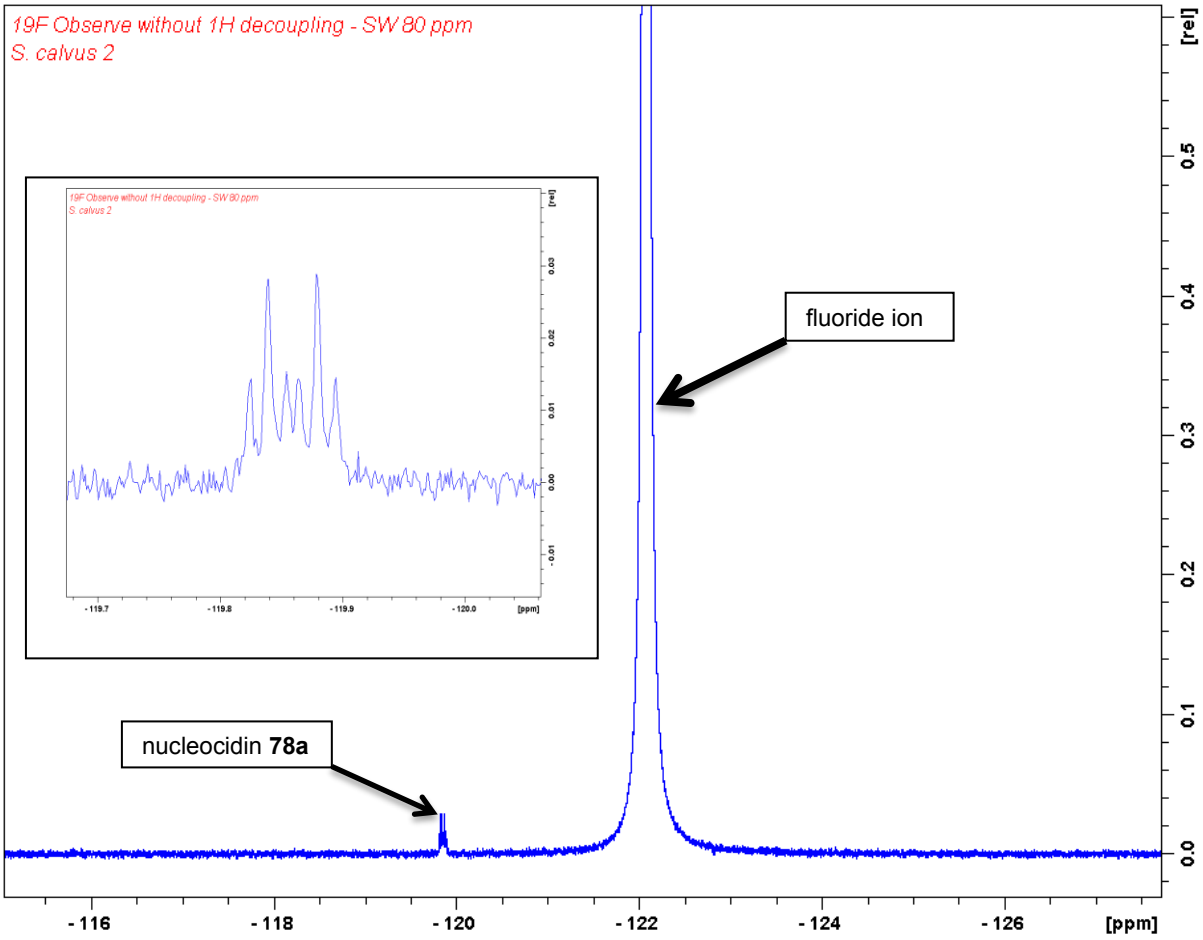


Figure 2. ^{19}F-NMR spectrum of a butanol extract in D_2O of *S. calvus* culture after 18 days.

The experimental multiplicity (doublet of triplets), and coupling constants of nucleocidin **78a** proved to be comparable to the reported coupling constants found in the literature, as presented in Figure 3.[8]

	Experimental	**Literature[8]**
Multiplicity	dt	dt
Chemical shift	-119.86 ppm	-121.64 ppm
***J* 3'-F**	19.0 Hz	17.2 Hz
***J* 5'-F**	7.3 Hz	7.5 Hz

Figure 3. Comparison of ^{19}F-NMR spectrum between *Streptomyces calvus* culture in D_2O and synthetic nucleocidin **78a** in D_2O.

LC-MS was then carried out on the butanol extract of the supernatant of *S. calvus* cultures to further confirm the production of nucleocidin **78a** as shown in Figure 4. Moreover, purified nucleocidin **78a** was provided by Pfizer and a ^{19}F-NMR spiking experiment was performed, which confirmed undoubtedly the biosynthesis of nucleocidin **78a**. These results confirmed that nucleocidin-producing cultures had been successfully established.

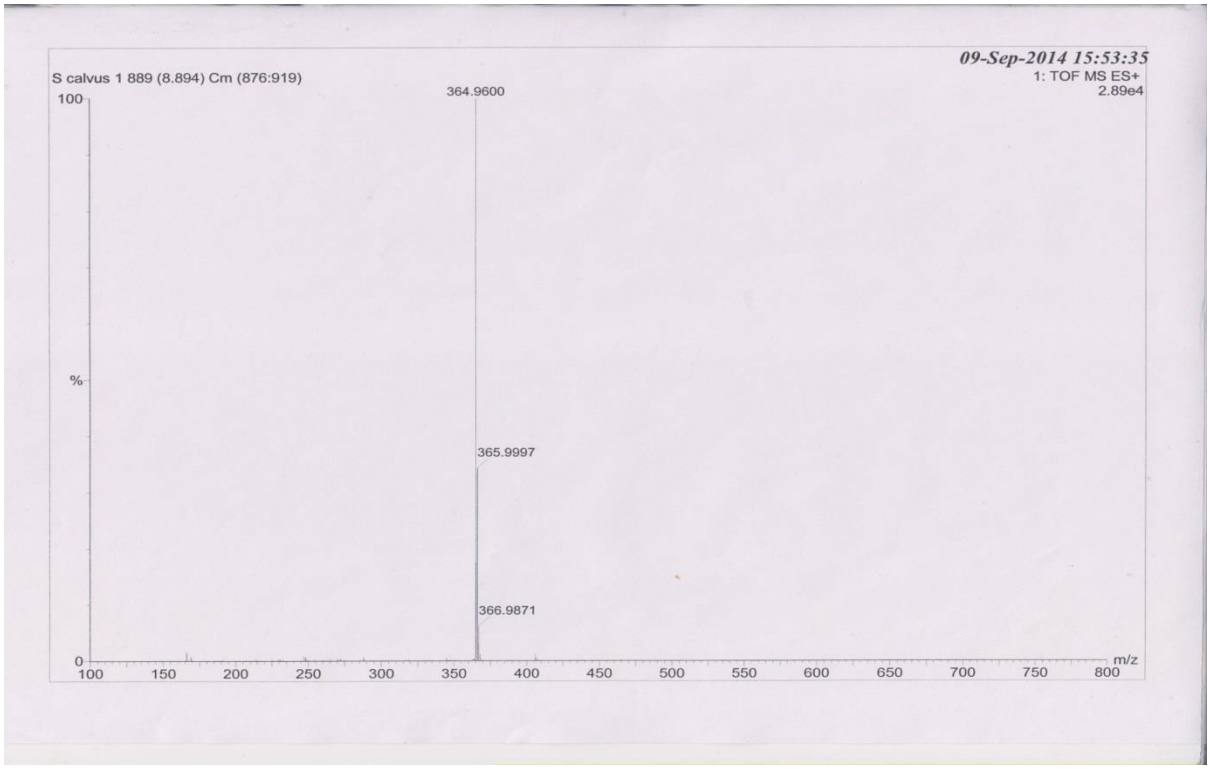


Figure 4. LC-MS analysis of *n*-butanol extract showing the production of nucleocidin **78a** in *Streptomyces calvus* culture.

Not only was the amount of nucleocidin **78a** produced from *S. calvus* T-3018 cultures very low, its production was also found to be very fickle. Indeed, most petri plates containing *Streptomyces calvus* did not show many spores. Cultures from petri plates containing spores did show enhanced nucleocidin **78a** production. However, cultures from petri plates without spores could also support the production of nucleocidin **78a**, although at much lower levels. It was also observed that the titres of nucleocidin **78a** dissapeared with glycerol stock over time. Therefore, glycerol stocks of *S. calvus* T-3018 were regularly supplied by Pfizer.

Comparison of nucleocidin **78a** and adenosine **111a** reveal that they are structurally similar molecules. Both contain a ribose ring system as well as an adenine base. In addition to this, nucleocidin **78a** contains a fluorine atom attached at the C-4' of the ribose ring and a sulfamoyl group bound to the primary alcohol. Subsequently, it was hypothesised that adenosine could be either a putative substrate for a new fluorinating enzyme or an early intermediate in the biosynthesis of nucleocidin **78a**. Therefore, the role and biosynthesis of adenosine **111a**, in relation to nucleocidin **78a** biosynthesis, were investigated.

3.2. Biosynthesis of adenosine 111a

3.2.1. The role of adenosine 111a

Nucleosides are glycosylamines that contain nitrogenous bases such as: adenine, guanine, uracil, thymine and cytosine. This nucleobase is linked to a 5-carbon sugar such as ribose or deoxyribose *via* a *β*-N-glycosidic bound. All naturally occurring ribonucleosides are represented in Figure 5.

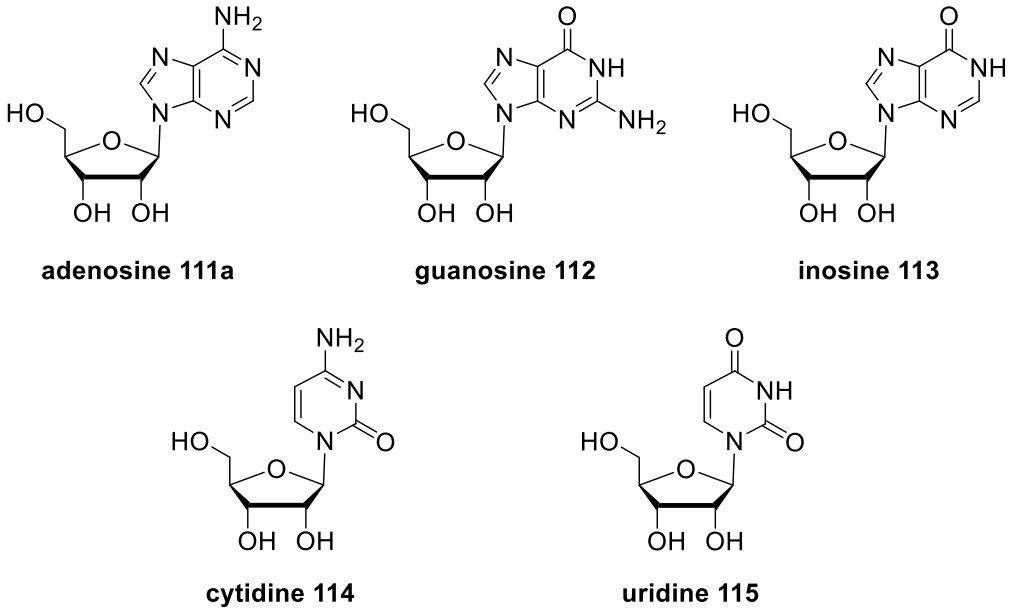


Figure 5. Structures of ribonucleosides.

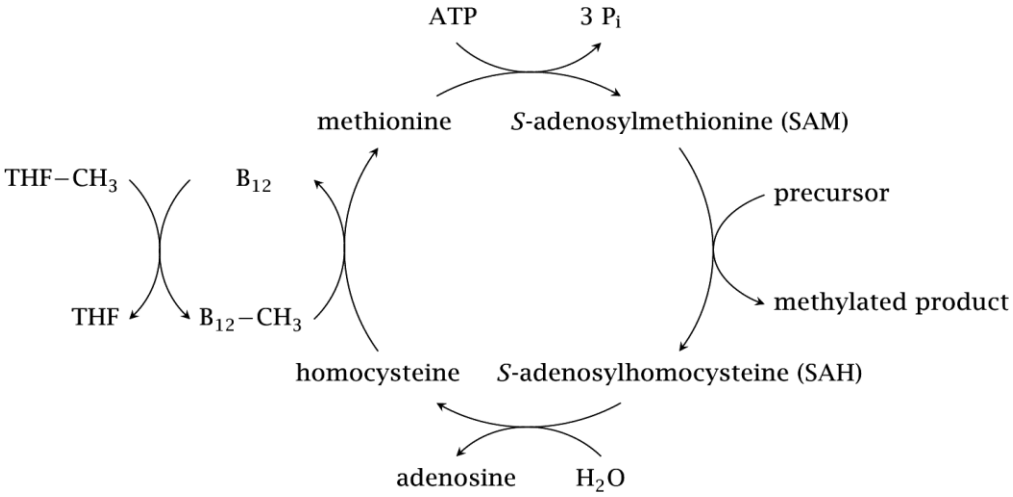
Certain kinases, possess the ability to phosphorylate nucleosides at the 5'-primary alcohol. After such phosphorylation, they are known as nucleotides, which are the building blocks of DNA and RNA. These nucleotides can be acquired either from the diet or biosynthesised by *de novo* pathways and are essential to almost all biological process within a cell.

Nucleosides can be phosphorylated up to three times to form crucial nucleotides such as adenosine triphosphate (ATP), which provides the chemical energy needed for metabolism within cells; as the hydrolysis of ATP, forming adenosine diphosphate (ADP) and inorganic phosphate, allows for the release of energy that can be utilised in a variety of metabolic reactions.

3.2.2. From glycerol 116a to adenosine 111a

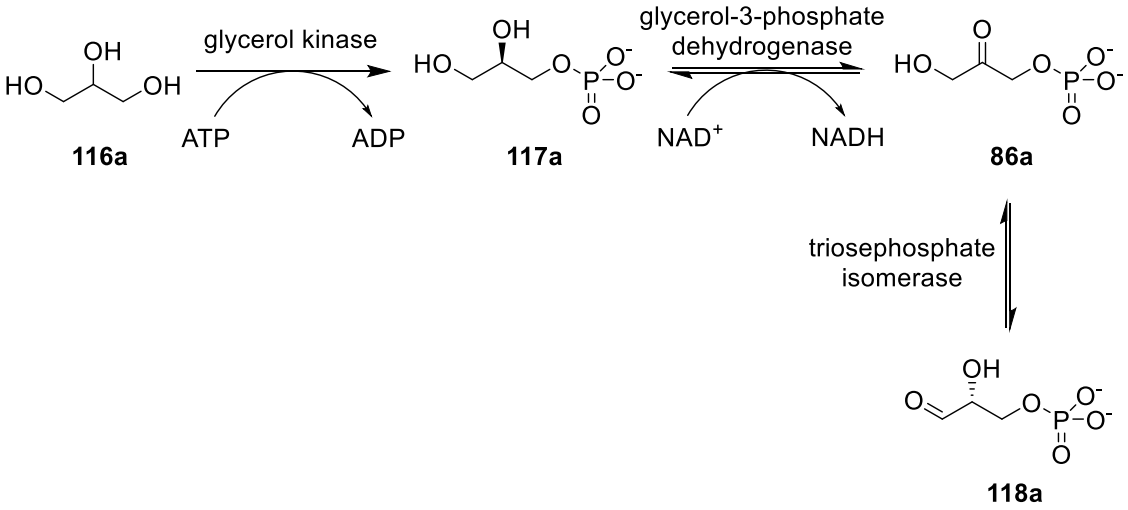
To establish that adenosine **111a** is involved in the biosynthesis of nucleocidin **78a**, the biosynthesis of adenosine **111a** in *S. calvus* T-3018 was investigated. An overview of the numerous metabolic pathways producing adenosine **111a** is described below. Since the strategy of this project was to conduct feeding experiments with labelled glycerols, the investigation was focused on the biosynthesis of adenosine **111a** from glycerol **116a** within the cell.

One of these involves the *S*-adenosylmethionine (SAM **27a**) cycle as described below in Scheme 1. In the SAM cycle, methionine **29** is first activated to generate SAM **27a**, which then donates a methyl-group to various precursors which produce *S*-adenosylhomocysteine (SAH). SAH then undergoes cleavage to afford adenosine **111a** and homocysteine. Finally, the regeneration of methionine from homocysteine requires vitamin B_{12} in the form of methylcobalamin.[9]



Scheme 1. SAM cycle producing adenosine **111a**.

Additionally, adenosine **111a** can also be biosynthesised by a different metabolic pathway derived from glycerol **116a**. Glycerol **116a** is a prochiral molecule and a precursor of lipid triglycerides. This precursor can enter two different, parallel, metabolic pathways: glycolysis and the pentose phosphate pathway. Before entering into one of these two metabolic pathways, glycerol **116a** must first be activated and transformed to a common intermediate known as glyceraldehyde-3-phosphate **118a**, as described in Scheme 2.

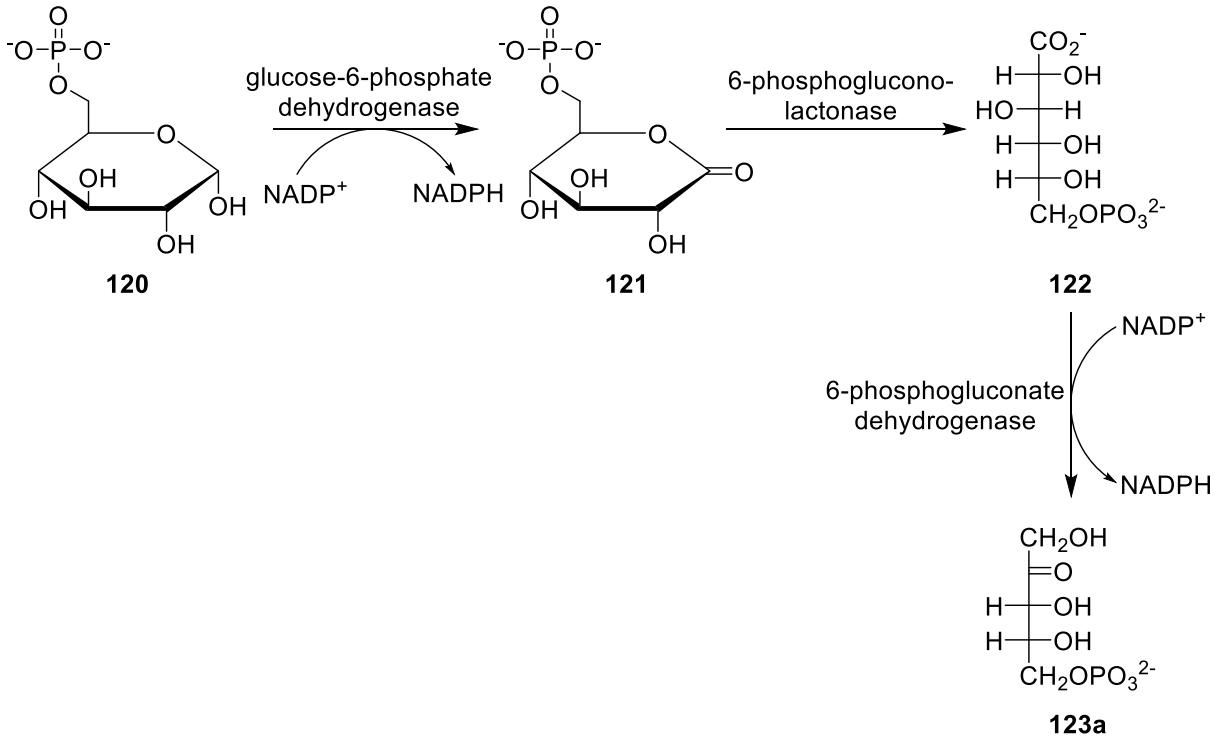


Scheme 2. Biosynthesis of glyceraldehyde-3-phosphate **118a** from glycerol **116a**.

Firstly, phosphorylation of glycerol **116a** by glycerol kinase to afford *sn*-glycerol-3-phosphate **117a** is stereoselective, since glycerol **116a** is always phosphorylated on the same pro-chiral arm (pro-

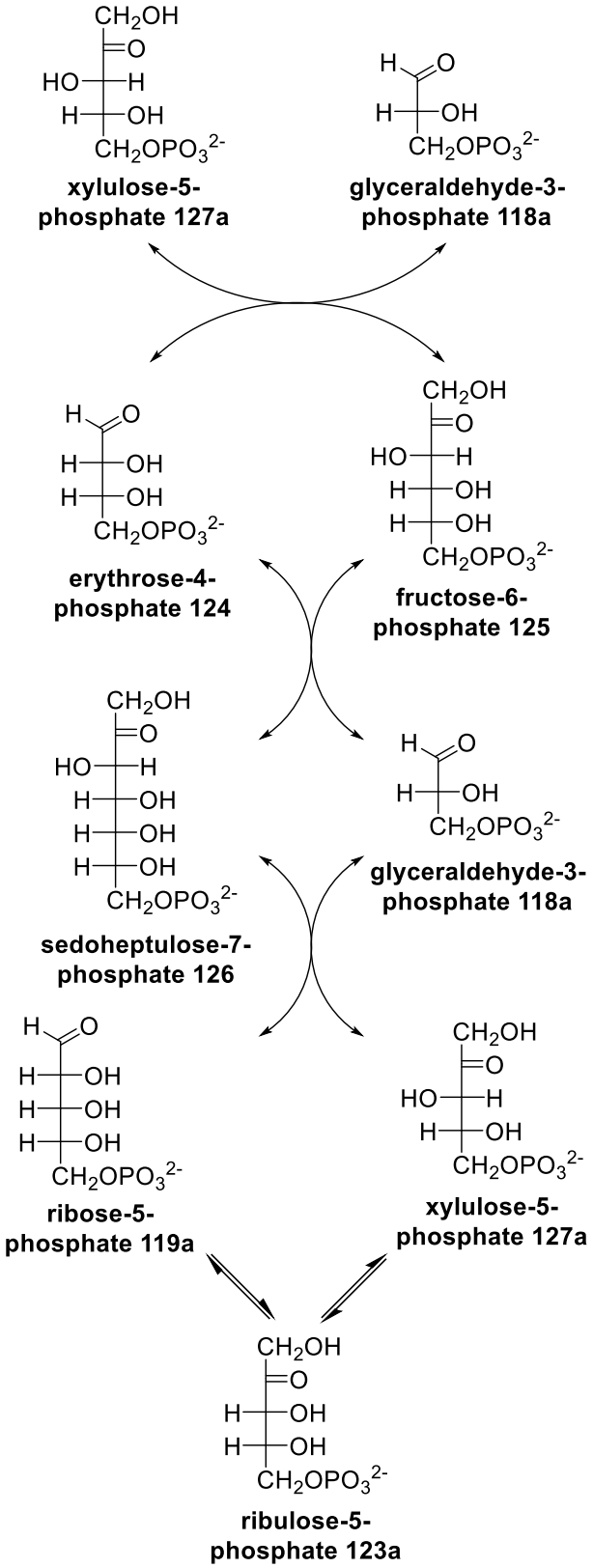
R). Oxidation of the secondary alcohol of *sn*-glycerol-3-phosphate **117a** is then achieved by a dehydrogenase enzyme, resulting in the formation of dihydroxyacetone phosphate **86a**. Finally, this undergoes isomerisation by triosephosphate isomerase affording glyceraldehyde-3-phosphate **118a**.

Glyceraldehyde-3-phosphate **118a** can enter the pentose phosphate pathway to generate ribose-5-phosphate **119a** as shown in Scheme 4, an advanced intermediate in the biosynthesis of adenosine **111a**. The pentose phosphate pathway contains two different phases: an oxidative and a non-oxidative phase.[10] The oxidative phase comprises the transformation of glucose-6-phosphate **120** to ribulose-5-phosphate **123a** in three enzyme-catalysed reactions as represented in Scheme 3. This phase begins with the oxidation of the hydroxyl group at C-1 of glucose-6-phosphate **120**, generating 6-phosphogluconolactone **121** and co-enzyme NADPH. 6-Phosphogluconolactonase is responsible for the hydrolysis of 6-phosphogluconolactone **121** into 6-phosphogluconate **122**. Finally, 6-phosphogluconate **122** undergoes oxidative decarboxylation affording ribulose-5-phosphate **123a** and NADPH.



Scheme 3. Oxidative phase of pentose phosphate pathway.

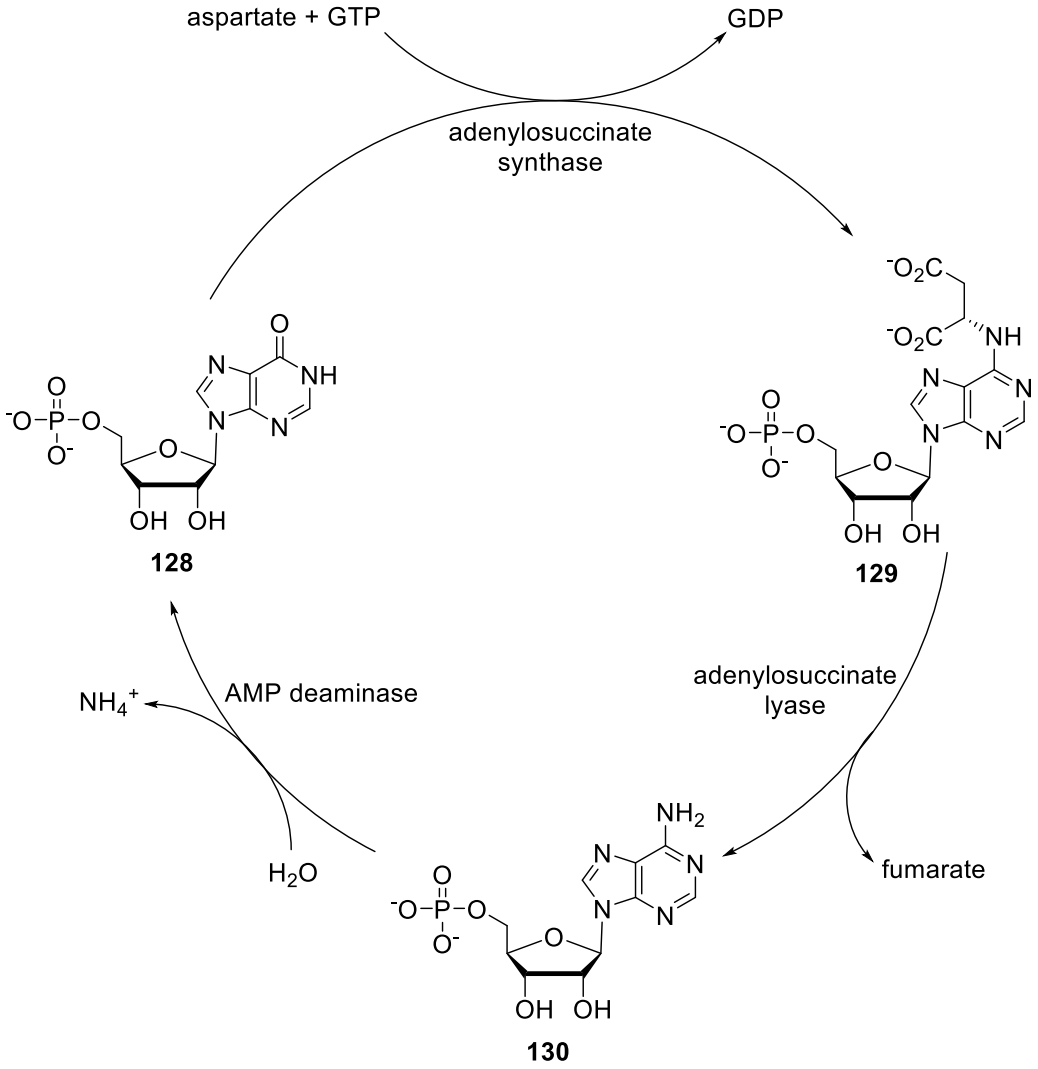
The non-oxidative phase is responsible for the formation of 5-carbon sugars, as shown in Scheme 4. During this phase, two very important sugars are biosynthesised: erythrose-4-phosphate **124** is a precursor of the aromatic amino acids and ribose-5-phosphate **119a** is a precursor of nucleotides and nucleic acids. Glyceraldehyde-3-phosphate **118a** can either react with fructose-6-phosphate **125** or sedoheptulose-7-phosphate **126** in reactions mediated by the same transketolase enzyme. Firstly, transfer of a 2-carbon fragment from fructose-6-phosphate **125** to glyceraldehyde-3-phosphate **118a** affords xylulose-5-phosphate **127a** and erythrose-4-phosphate **124**. Alternatively, transfer of a 2-carbon fragment from sedoheptulose-7-phosphate **126** to glyceraldehyde-3-phosphate **118a** gives ribose-5-phosphate **119a** and xylulose-5-phosphate **127a**. Epimerisation of xylulose-5-phosphate **127a** to ribulose-5-phosphate **123a** is catalysed by ribulose 5-phosphate 3-epimerase enzyme.[11] Finally, ribulose 5-phosphate isomerase catalyses the biotransformation of ribulose-5-phosphate **123a** to ribose-5-phosphate **119a**, which is an important intermediate in the biosynthesis of adenosine **111a**.[12]



Scheme 4. Non-oxidative phase of pentose phosphate pathway.

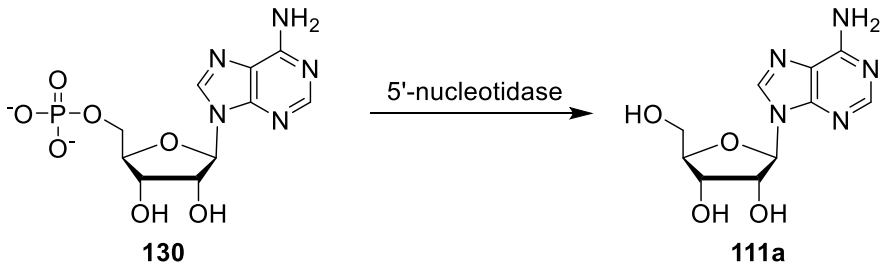
After its production, ribose-5-phosphate **119a** undergoes eleven enzymatic transformations to afford the important intermediate inosine monophosphate **128** (IMP). Conversion of IMP **128**

to adenylosuccinate **129** is then catalysed by adenylosuccinate synthase.[13] Adenylosuccinate **129** undergoes *β*-elimination which causes its cleavage, generating AMP **130** and fumarate. Finally, IMP **128** can be regenerated from AMP **130** through the release ammonia mediated by AMP deaminase enzyme as described in Scheme 5.



Scheme 5. Biosynthesis of AMP **130** from IMP **128**.

AMP **130** can either be phosphorylated to form adenosine diphosphate (ADP), and then again to form adenosine triphosphate (ATP) or be hydrolysed to give adenosine **111a**. Indeed, 5'-nucleotidase is an enzyme responsible for catalysing the phosphorolytic cleavage of 5'-nucleotides such as AMP **130**, as shown in Scheme 6.



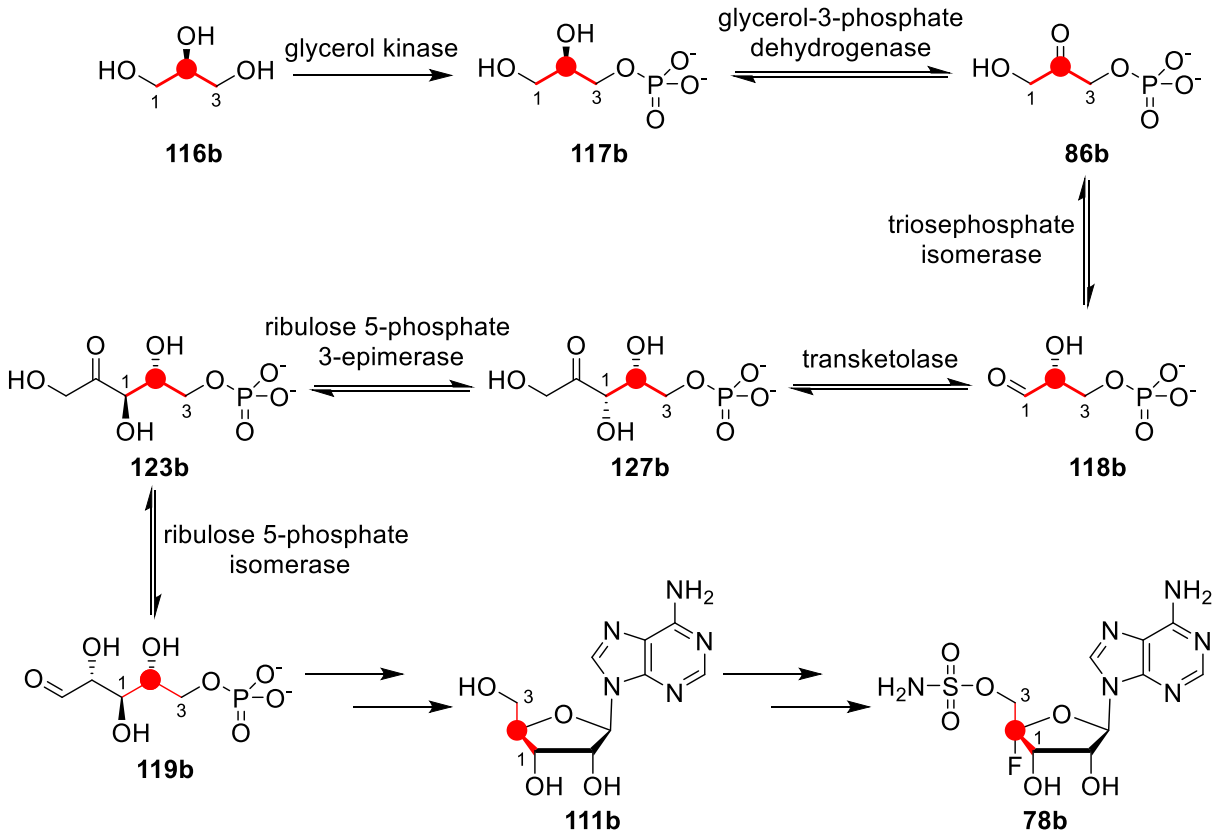
Scheme 6. Conversion of AMP **130** to adenosine **111a** by 5'-nucleotidase.

3.3. Feeding experiments of [2-^{13}C]-glycerol 116b and [1,1,2,3,3-$^{2}H_5$]-glycerol 116c

3.3.1. Isotopic incorporation of [2-^{13}C]-glycerol 116b into nucleocidin 78a

According to the pentose phosphate pathway (described previously), the carbon C-2 of glycerol **116a** should be incorporated as the carbon C-4' of nucleocidin **78b**, the carbon directly attached to fluorine. If this hypothesis proves to be true, a feeding experiment with [2-^{13}C]-glycerol **116b** would reveal this incorporation by ^{19}F-NMR, as coupling should be observed between the fluorine atom and the ^{13}C-carbon.

In order to determine the optimal time for the addition of isotopically labelled glycerol into *Streptomyces calvus* fermentations, a profile of nucleocidin production was carried out by performing a work-up of the cultures every two days. It was observed that nucleocidin **78a** production begins after 9-10 days of incubation. The labelled glycerols were then added to different *S. calvus* fermentations (100 mL) after 4 days, followed by subsequent similar additions every two days until a total of 6 additions of either glycerol had been achieved. The final concentration of added glycerol was 10 mM in each case. After 18 days of fermentation, the cells were removed after centrifugation and the supernatant was extracted into *n*-butanol (20 mL). The organic layer was concentrated under reduced pressure, and the extract was analysed by ^{19}F-NMR in order to detect nucleocidin **78a**.[14] The outcome of this feeding experiment can be rationnalised, as shown in Scheme 7.



Scheme 7. [2-^{13}C]-Glycerol **116b** incorporation into nucleocidin **78b**.

After a successful work-up, it was clear that the fluorine signal of nucleocidin **78a** was accompanied by satellites due to $^{1}J_{CF}$ coupling, consistent with a population of nucleocidin **78b** molecules enriched with ^{13}C at the carbon C-4' of the ribose ring. This experiment showed a high incorporation (25%), with a ^{13}C-^{19}F coupling constant of 232 Hz, as shown in Figure 4. The satellite signals are off-centre relative to the natural abundance ^{19}F-NMR signal due to a heavy atom α-shift.[15] This incorporation is entirely consistent with the prediction shown previously in Scheme 7. This experiment proved that the ribose ring of nucleocidin **78a** is indeed biosynthesised *via* the pentose phosphate pathway.

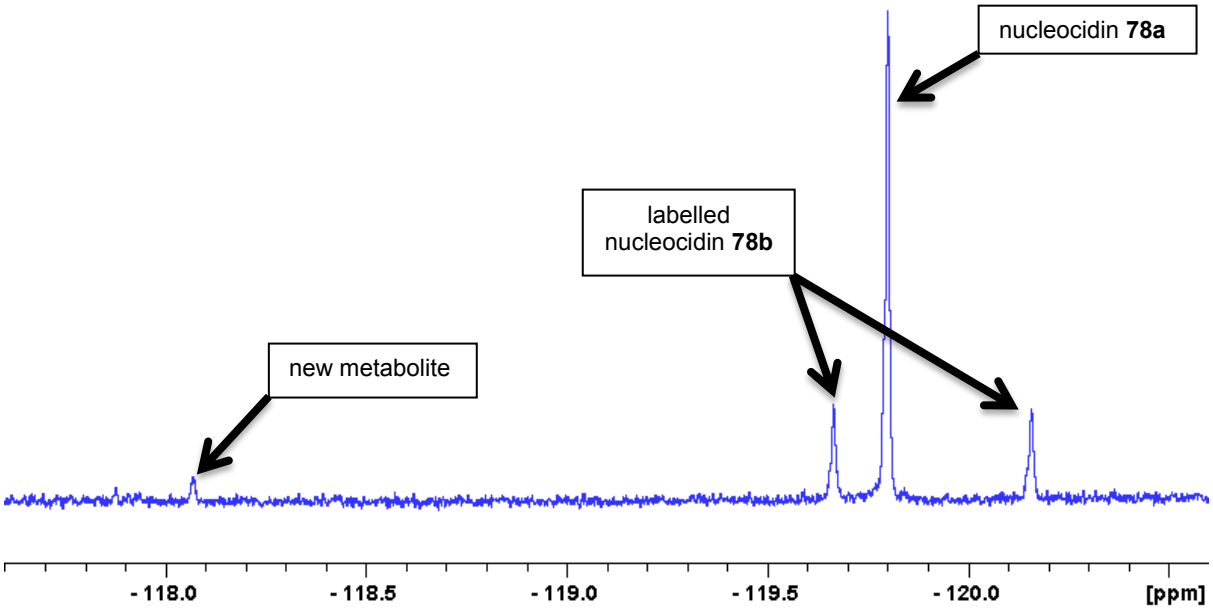


Figure 6. $^{19}F\{^{1}H\}$-NMR spectrum of feeding experiment with [2-^{13}C]-glycerol **116b** in D_2O.

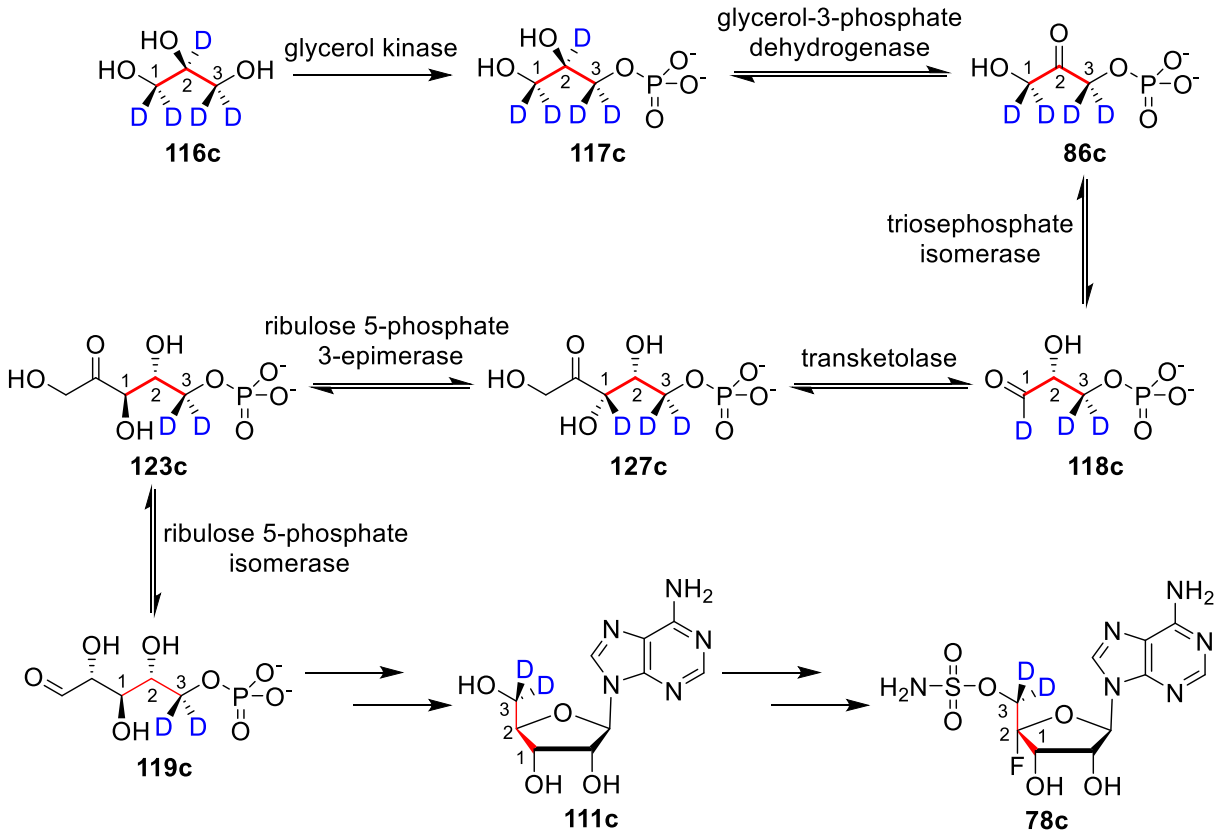
Moreover, an additional organofluorine compound was observed at -118.1 ppm, which revealed the presence of a new fluorinated metabolite in very low quantity as shown in Figure 6.[14] Such a fluorometabolite was also observed in the recent study where nucleocidin production was elicited by reversing the *bldA* mutation.[4] Coupled-fluorine NMR proved that nucleocidin **78a** and this new metabolite shared the same multiplicity and very similar coupling constants, as shown in Figure 7. Therefore, it appears to be metabolically related to nucleocidin **78a** either as a biosynthetic precursor, or as a metabolite.[16]

	nucleocidin 78a	**new fluorometabolite**
Multiplicity	dt	dt
Chemical shift	-119.8 ppm	-118.1 ppm
J 3'-F	19.0 Hz	18.9 Hz
J 5'-F	7.3 Hz	6.6 Hz

Figure 7. Comparison of ^{19}F-NMR spectrum between nucleocidin **78a** and new fluorometabolite in D_2O.

3.3.2. Isotopic incorporation of [1,1,2,3,3-$^{2}H_5$]-glycerol 116c into nucleocidin 78a

After the successful pulse feeding experiments with [2-^{13}C]-glycerol **116b**, the commercially available [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** was then incubated with *Streptomyces calvus* cultures, using a similar protocol. The outcome of this feeding experiment can be rationalised, as shown in Scheme 8. Labelled glycerol **116c** undergoes phosphorylation, followed by oxidation of the secondary alcohol leading to the loss of a deuterium atom in **86c**. Isomerisation of tetradeuterated dihydroxyacetone phosphate **86c** leads to glyceraldehyde-3-phosphate **118c** and loss of another deuterium atom. The transketolation followed by epimerisation of **127c** catalysed by ribulose 5-phosphate 3-epimerase enzyme would result in the loss of a further deuterium atom, leaving only two deuterium atoms in ribose-5-phosphate **119c**. Therefore, [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** should generate [5',5'-$^{2}H_2$]-adenosine **111c**. This experiment would confirm that the carbon C-5' of nucleocidin **78a** is derived from the carbon C-3 of glycerol **116a**. Moreover, the chemical shift of the labelled nucleocidin **78c** would show whether the two deuterium atoms are still both attached to the carbon C-5' of nucleocidin **78a**.



Scheme 8. [1,1,2,3,3-$^{2}H_5$]-Glycerol **116c** incorporation into nucleocidin **78c**.

Addition of [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** to a final concentration of 10 mM into *Streptomyces calvus* fermentation was carried out, following the procedure described previously. The same work-up as previously mentioned was also performed. This experiment led to a 5% incorporation of deuterium into nucleocidin **78a**, as shown in Fugure 8.[14] The presence of the new fluorometabolite observed in [2-^{13}C]-glycerol **116b** experiment was not apparent in this case.

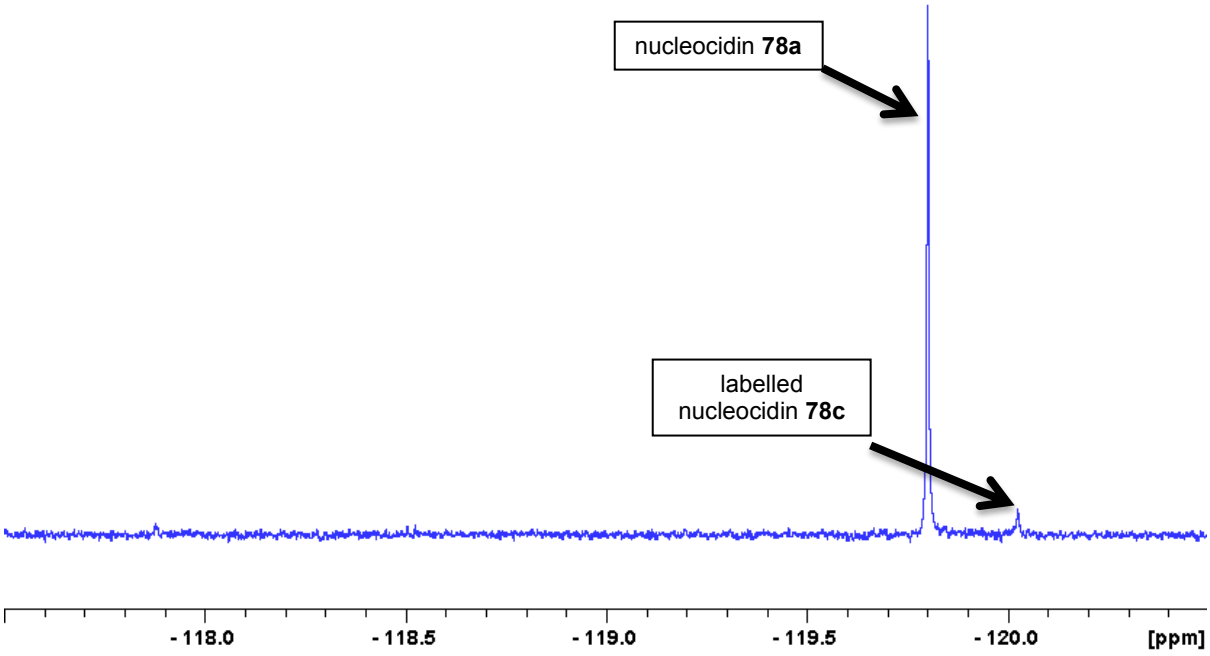


Figure 8. $^{19}F\{^{1}H\}$-NMR spectrum of feeding experiment with [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** in D_2O.

Feeding experiments with [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** led to a 5% incorporation of deuterium into nucleocidin **78a**, as indicated by an upfield isotope induced shift in the resultant $^{19}F\{^{1}H\}$-NMR spectrum, as shown in Figure 8. The magnitude of this upfield shift at 0.22 ppm is consistent with a single vicinal deuterium incorporated at the *β*-position to the fluorine atom. According to previous studies by Lambert and Greifenstein, the magnitude of this shift will change depending on the vicinal angle between the carbon-deuterium and the carbon-fluorine bonds.[17] Indeed, it was established that if the D-C-C-F angle is included between 0° and 60°, the isotope shift is about 0.15 ppm. This isotope shift increases to 0.35 ppm when the dihedral angle is between 120° and 180°, as shown in Figure 9.[18] This experiment confirmed that the origin of the carbon C-5' of nucleocidin **78a** is the carbon C-3 of glycerol 116a. The loss of one deuterium atom at the carbon C-5' of nucleocidin **78d** indicates that the mechanism of the biological fluorination might not be as straightforward as initially assumed.

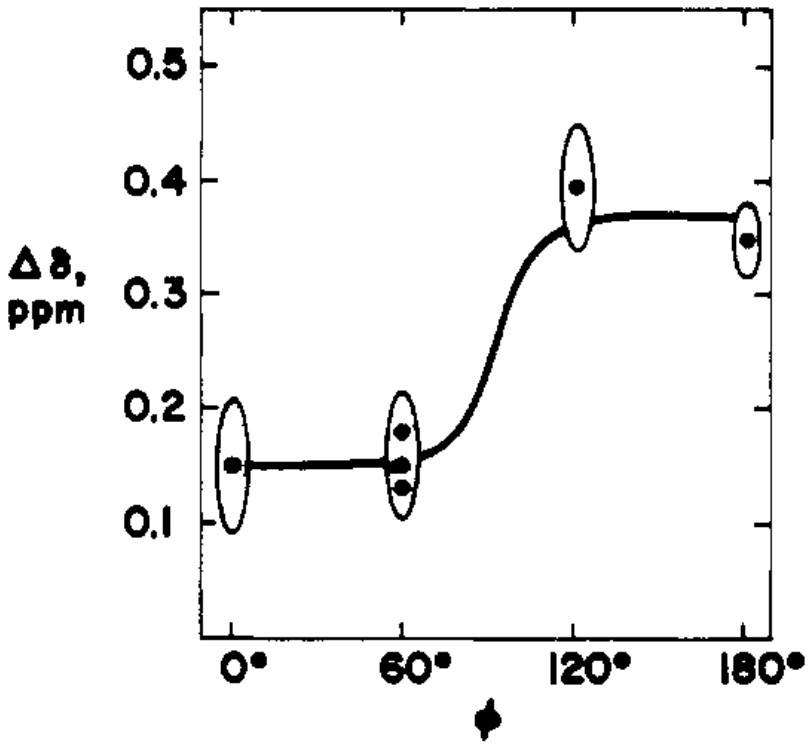


Figure 9. The D-C-C-F chemical shift isotope effect (Δδ) as a function of dihedral angle (ϕ) (figure taken from Lambert *et al.*[18]).

Therefore, the upfield shift of 0.22 ppm in the feeding experiment with glycerol-1,1,2,3,3-d_5 **116c** confirmed the presence of only one deuterium atom at the carbon C-5' of the ribose ring in nucleocidin **78d** with a D-C-C-F angle around 90°, as shown in Figure 10. This result could be explained by oxidation-reduction processes leading to a loss of one deuterium atom. This hypothesis will be discussed in more details in the **Chapter 4** in section **4.1.3**.

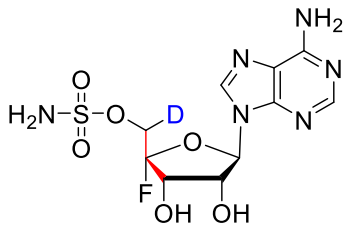
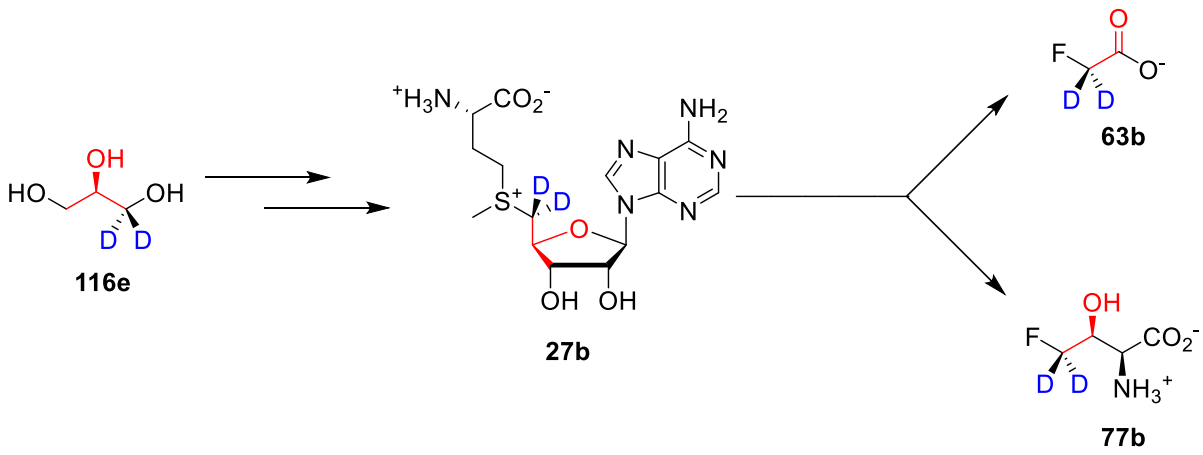


Figure 10. Structure of labelled nucleocidin **78d**.

3.4. Feeding of (2*S*)-[1-$^2H_2$]-glycerol 116d and (2*R*)-[1-$^2H_2$]-glycerol to *Streptomyces calvus* 116e

3.4.1. Feeding experiments in *Streptomyces cattleya*

Previously, a series of isotopic labelling studies with deuterated labelled glycerols in *Streptomyces cattleya* was repeated. This involved the synthesis and feeding experiments of (2*S*)-[1-$^2H_2$]-glycerol **116d** and (2*R*)-[1-$^2H_2$]-glycerol **116e**.[19] These isotopically enriched glycerols **116d** and **116e** were administered in distinct resting cell suspensions to cultures of *S. cattleya*. These experiments established that both hydrogens on the carbon C-5' of the ribose ring of SAM **27a** are retained in its conversion to fluoroacetate **63a** and 4-fluoro-L-threonine **77a**.[20] Further studies, along with a more comprehensive understanding of the individual steps within the biosynthesis of fluoroacetate **63a** and 4-fluoro-L-threonine **77a**, indicate that the carbons C-5' and C-4' of SAM **27c** were converted to fluoroacetate **63b** and, carbons C-3' and C-4' of fluorothreonine **77b** as indicated in Scheme 9.[15]

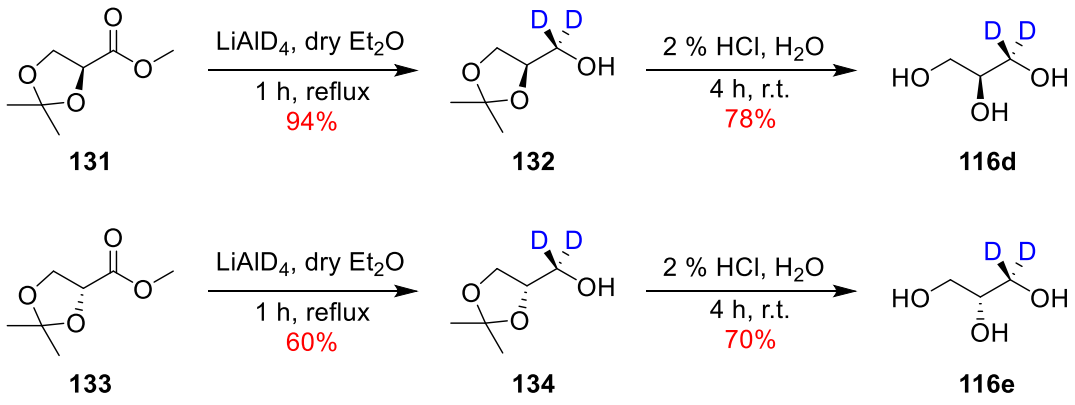


Scheme 9. Carbons C-5' and C-4' of SAM **27c** are retained from (2R)-[1-$^2H_2$]-glycerol **116e** in fluoroacetate **63b** and 4-fluoro-L-threonine **77b** biosynthesis.[20]

3.4.2. Synthesis of (2*S*)-[1-$^2H_2$]-glycerol 116d and (2*R*)-[1-$^2H_2$]-glycerol 116e

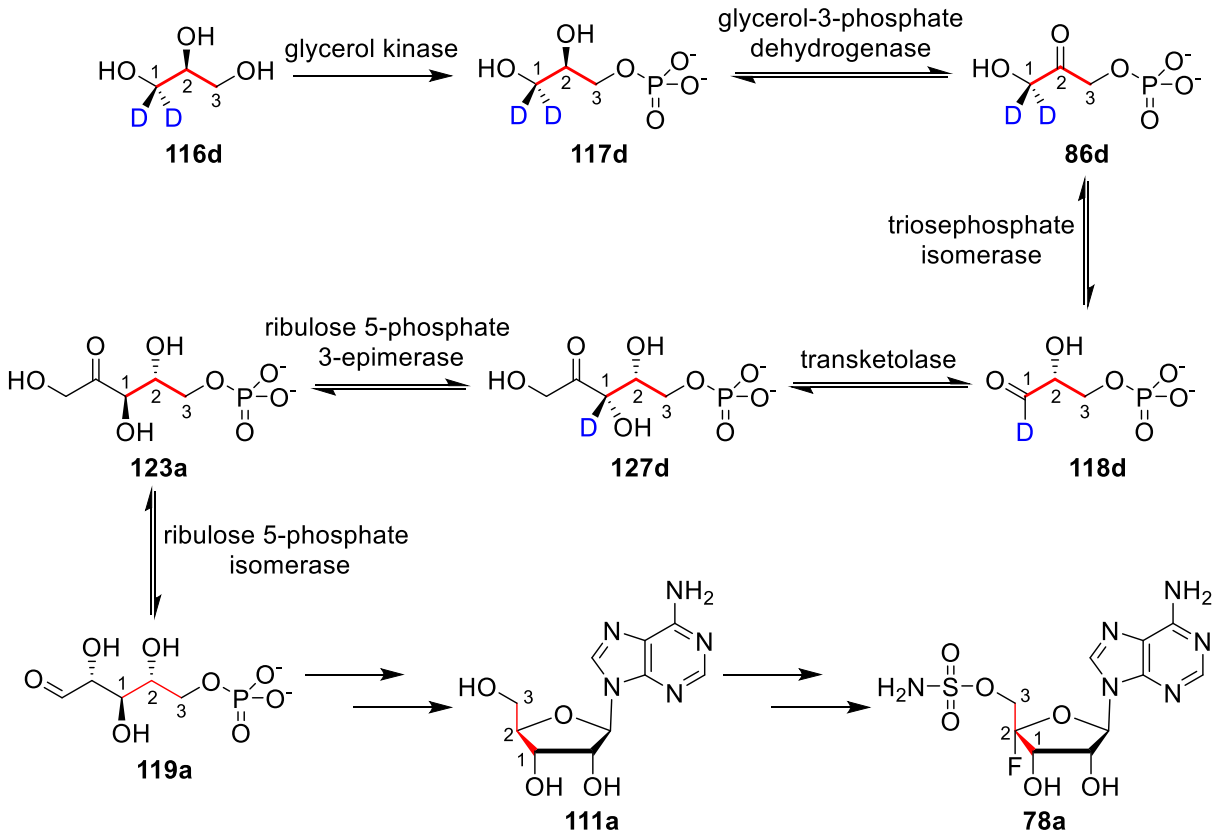
Feeding experiments with (2*S*)-[1-$^2H_2$]-glycerol **116d** and (2*R*)-[1-$^2H_2$]-glycerol **116e** in *Streptomyces calvus* would allow for further confirmation of the previous results shown above. Additionally, feeding experiment with (2*S*)-[1-$^2H_2$]-glycerol **116d** may provide information concerning the origin of carbon C-3' of nucleocidin **78a**.

In order to repeat the feeding experiments performed by Nieschalk *et al*., the synthesis of (2*S*)-[1-$^2H_2$]-glycerol **116d** and (2*R*)-[1-$^2H_2$]-glycerol **116e** was carried out using the procedure developed by *Hill et al.*[19,20] Commercially available (*S*)-methyl 2,2-dimethyl-1,3-dioxolane-4-carboxylate **131** and (*R*)-methyl 2,2-dimethyl-1,3-dioxolane-4-carboxylate **133** were both treated with $LiAlD_4$ in dry ether, under reflux, for 1 h. Aqueous quenching of these reactions forms aluminium hydroxides which are responsible for the formation of emulsions, making extractions of the product problematic. In order to improve the work-up, a method established by Fieser was used for purification of acetonides **132** and **134**.[21] This involved addition of H_2O to the reaction mixture at 0 °C, followed by aqueous sodium hydroxide (15%), and then a further addition of H_2O. Finally, HCl mediated hydrolysis of both alcohols **132** and **134** gave the two [1-$^2H_2$]-glycerol enantiomers **116d** and **116e** in good yield and purity as shown in Scheme 10.



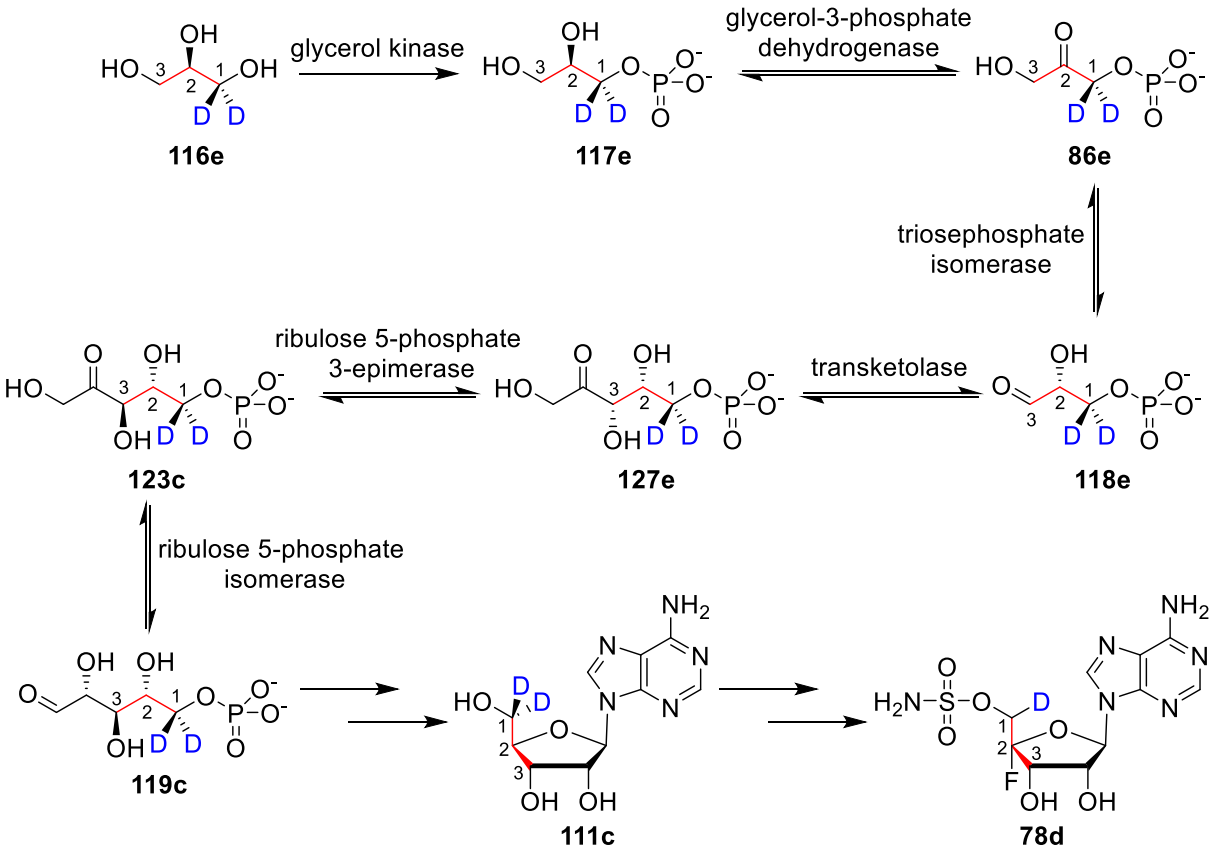
Scheme 10. Synthesis of (2*S*)-[1-$^2H_2$]-glycerol **116d** and (2*R*)-[1-$^2H_2$]-glycerol **116e**.

Once both deuterated glycerols **116d** and **116e** were in hand, feeding experiments were carried out. Based on the biosynthetic pathway of adenosine **111a**, the outcomes of each feeding experiment were anticipated. Feeding experiments using (2*S*)-[1-$^{2}H_2$]-glycerol **116d** were predicted to generate dihydroxyacetone phosphate **86d**. Isomerisation would then afford labelled glyceraldehyde-3-phosphate **118d** with a loss of one deuterium atom as shown in Scheme 11. The conversion of glyceraldehyde-3-phosphate **118d** to xylulose-5-phosphate **127d** is catalysed by transketolase.This would result in [3-^{2}H]-xylulose-5-phosphate **127d** which would then undergo epimerisation with loss of the last deuterium affording ribulose-5-phosphate **123a**, with exchange of the surviving deuterium with the bulk solvent.[16,22] Ribulose-5-phosphate **123a** is then metabolised to generate nucleocidin **78a** without any isotope. Therefore, no incorporation of deuterium atoms should be observed after the (2*S*)-[1-$^{2}H_2$]-glycerol **116d** feeding experiment.



Scheme 11. Rationalisation of the feeding experiment with (2*S*)-[1-$^{2}H_2$]-glycerol **116d** into nucleocidin **78a**.

On the other hand (2*R*)-[1-$^2H_2$]-glycerol **116e** would be converted to [1-$^2H_2$]-glyceraldehyde-3-phosphate **118e**, which would then enter the pentose phosphate pathway and be transformed to [1-$^2H_2$]-ribose-5-phosphate **119c** with retention of the two deuterium atoms. Like the feeding experiment with [1,1,2,3,3-$^2H_5$]-glycerol **116c**, deuterated adenosine **111c** would be formed from [1-$^2H_2$]-ribose-5-phosphate **119c**, as shown in Scheme 12. Therefore, the same outcome observed during this feeding experiment should be observed for the feeding experiment with (2*R*)-[1-$^2H_2$]-glycerol **116e**.



Scheme 12. Rationalisation of the feeding experiment with (2*R*)-[1-$^2H_2$]-glycerol **116e** into nucleocidin **78d**.

3.4.3. Pulse feeding experiment with dideuterated glycerols 116d and 116e

As expected, feeding experiments with (2*S*)-[1-$^{2}H_{2}$]-glycerol **116d** did not show isotopic incorporation into nucleocidin **78a**. However, the presence of the new metabolite observed after the feeding experiment with [2-^{13}C]-glycerol **116b** was apparent after this feeding experiment.

Pulse feeding experiments with (2*R*)-[1-$^{2}H_{2}$]-glycerol **116e** led to an 8% incorporation of deuterium into nucleocidin **78a**, which was determined by an upfield shift in the resultant $^{19}F\{^{1}H\}$-NMR spectrum, as shown in Figure 11. The magnitude of this upfield shift at 0.22 ppm is the same as that observed after the feeding experiment with [1,1,2,3,3-$^{2}H_{5}$]-glycerol **116c**. This result confirmed the formation of a population of monodeuterated nucleocidin **78d** molecules. Moreover, each feeding experiment with a labelled glycerol **116b-e** was performed twice at least with the approximately the same yield of incorporation.

Moreover, the presence of the second fluorometabolite was also observed. This fluorometabolite has the same chemical shift as that of the other organofluorine compound detected after the feeding experiments with (2*S*)-[1-$^{2}H_{2}$]-glycerol **116d** and [2-^{13}C]-glycerol **116b**, as shown in Figure 11. However, in feeding experiments with (2*R*)-[1-$^{2}H_{2}$]-glycerol **116e** the level of the unknown metabolite was greater than that of nucleocidin **78a**. It also showed a similar level of deuterium incorporation as nucleocidin **78a**, with a 0.22 ppm shift corresponding to the presence also of only one deuterium atom.[14]

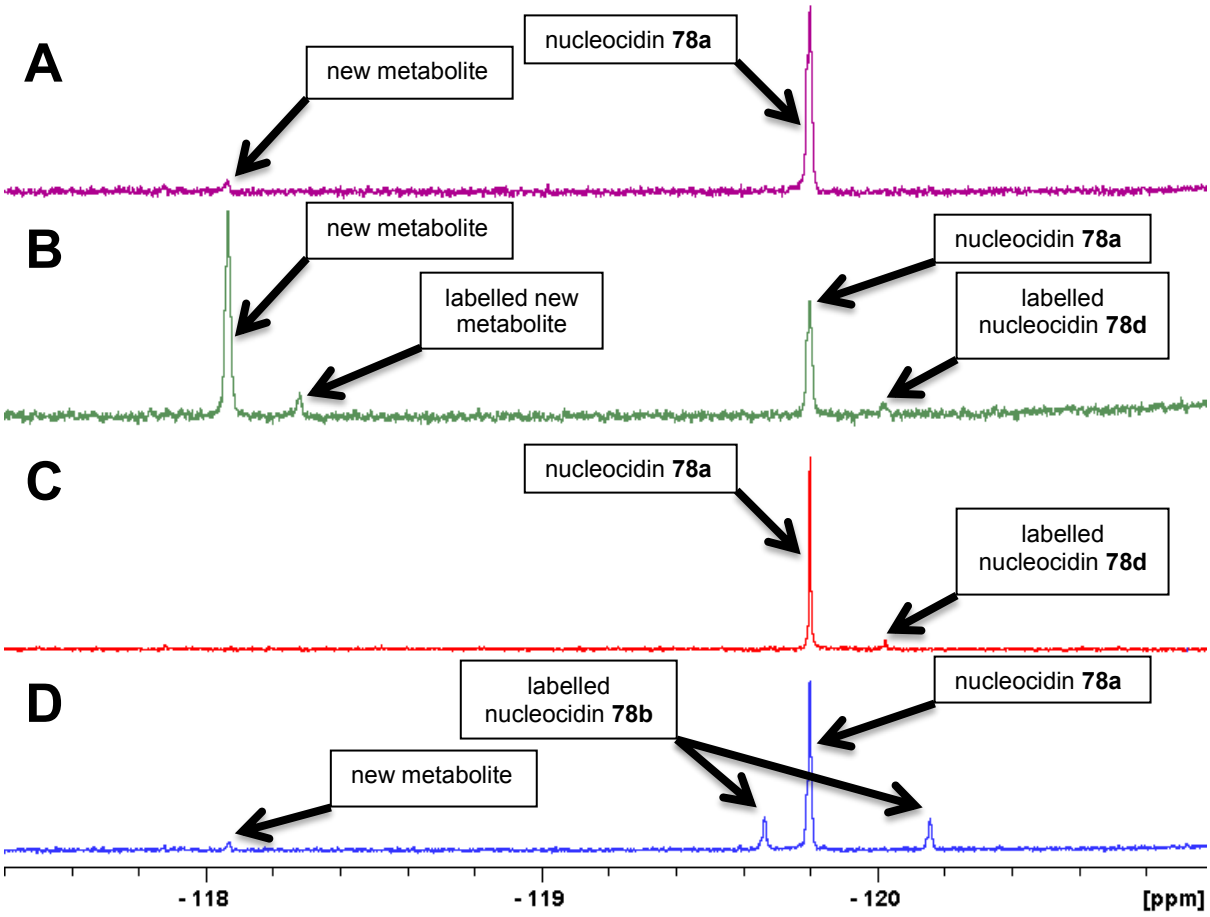


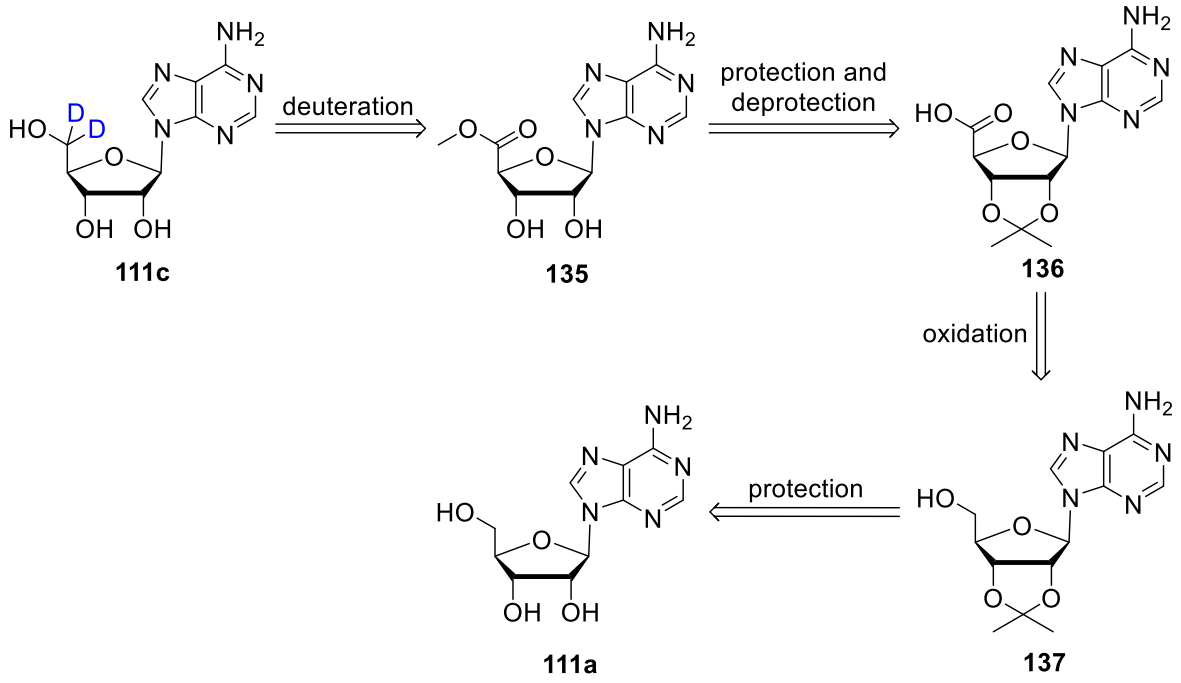
Figure 11. ^{19}F-NMR spectroscopic analysis of extracts of *S. calvus* after the addition of glycerol isotopomers:
(**A**) feeding experiment with (2*S*)-[1-$^{2}H_2$]-glycerol **116d**.
(**B**) feeding experiment with (2*R*)-[1-$^{2}H_2$]-glycerol **116e**.
(**C**) feeding experiment with [1,1,2,3,3-$^{2}H_5$]-glycerol **116c**.
(**D**) feeding experiment with [2-^{13}C]-glycerol **116b**.

3.5. Isotopically labelled adenosines 111c and 111d into nucleocidin 78a

After initial studies investigating the incorporation of labelled glycerol into nucleocidin **78a**, further feeding experiments using labelled adenosine were explored in order to test the hypothesis that adenosine **111a** is an advanced intermediate in the biosynthesis of nucleocidin **78a**. Therefore, the syntheses of [5',5'-$^2H_2$]-adenosine **111c** and [3'-^{2}H]-adenosine **111d** were addressed for this purpose.

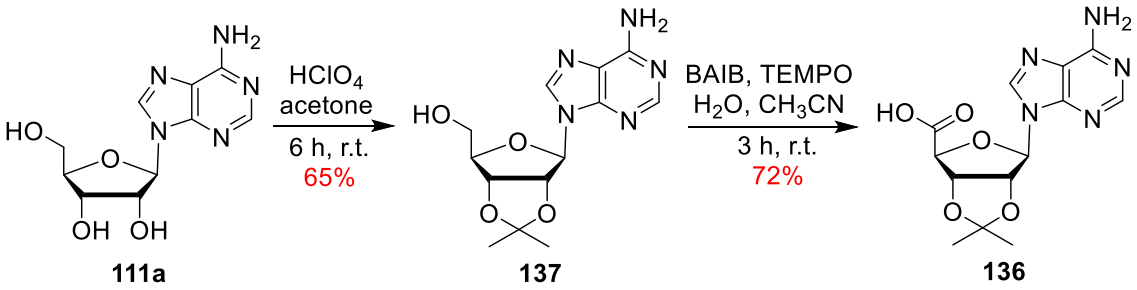
3.5.1. Synthesis of [5',5'-$^2H_2$]-adenosine 111c

In 1982, Pang *et al*.[23] reported the multistep synthesis of [5',5'-$^2H_2$]-adenosine **111c** from ethyl adenosine-5'-carboxylate. The synthesis of the latter was not reported, therefore, an alternative synthetic strategy was developed in this project for the preparation of [5',5'-$^2H_2$]-adenosine **111c**. After protection of adenosine **111a**, oxidation of acetonide **137** was envisaged to afford carboxylic acid **136**. Formation of ester **135** including acetal removal could be performed in a one-step process. Finally, [5',5'-$^2H_2$]-adenosine **111c** would be obtained after ester **135** reduction, as shown in Scheme 13.



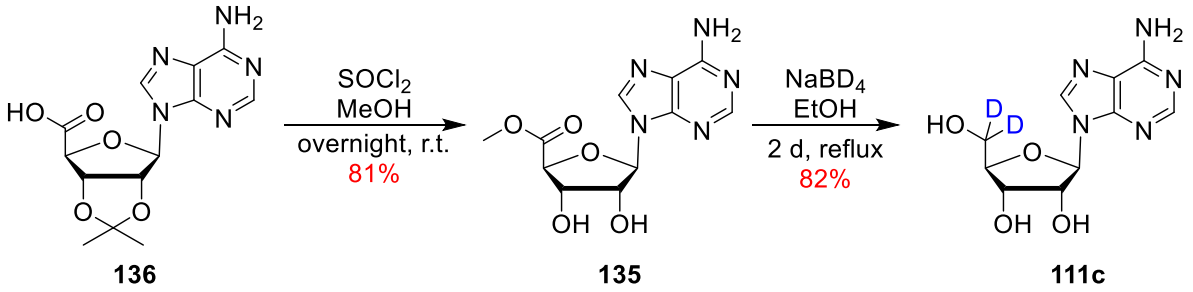
Scheme 13. Retrosynthetic strategy to [5',5'-$^{2}H_2$]-adenosine **111c**.

Acetonide **137** was obtained by reacting adenosine **111a** with acetone in the presence of perchloric acid. Purification of the reaction mixture by column chromatography afforded the product **137** in 65% yield.[24] Oxidation of acetonide **137** to adenosine-5'-carboxylic acid **136** following the procedure developed by Epp *et al*. was achieved in good yield.[25] Treating the acetal protected nucleoside **137** with a catalytic amount of 2,2,6,6-tetramethyl-1-piperidinyloxyl (TEMPO) and a stoichiometric amount of [bis(acetoxy)-iodo]benzene (BAIB) in a 1:1 ACN- H_2O afforded the carboxylic acid **136** in 72% yield in high purity, as shown in Scheme 14.



Scheme 14. Protection of adenosine **111a** and oxidation of **137**.

Acetonide **136** was treated with thionyl chloride in methanol to form an acyl chloride intermediate, which reacted with methanol affording the methyl ester **135**. Indeed, hydrochloric acid is generated in the reaction which also causes acetal hydrolysis. This deprotection is important for improving the solubility of ester **135** for the following deuteration reaction. The methyl ester **135** was purified by column chromatography in good yield (81%). Finally, ester **135** was dissolved in ethanol and sodium borodeuteride ($NaBD_4$) was added to the reaction mixture, as shown in Scheme 15.[23] Reduction of the methyl ester **133** generated [5',5'-$^2H_2$]-adenosine **111c** which was purified by column chromatography in high purity for use in feeding experiments. Despite the fact that proton NMR showed high purity, mass spectrometry analysis revealed the presence of a trace of unlabelled adenosine **111a** (2%).



Scheme 15. Synthesis of [5',5'-$^2H_2$]-adenosine **111c**.

The ^{1}H-NMR spectrum of [5',5'-$^2H_2$]-adenosine **111c** was compared to that of adenosine **111a**, as shown in Figure 12. This comparison showed an overall simplification of the coupling patterns. Indeed, the first significant observation is the disappearance of both peaks corresponding to H-5' in adenosine **111a**. Moreover, the multiplicity of OH-5' peak changed to a singlet, instead of a doublet of doublets in adenosine **111a** which is consistent with the presence of two deuterium atoms at the C-5' position. Also there was a tremendous change of the coupling pattern for the H-4' signal from a multiplet to a doublet. With all these observations, it was concluded that the synthesis of this adenosine isotopomer **111c** had been successful.

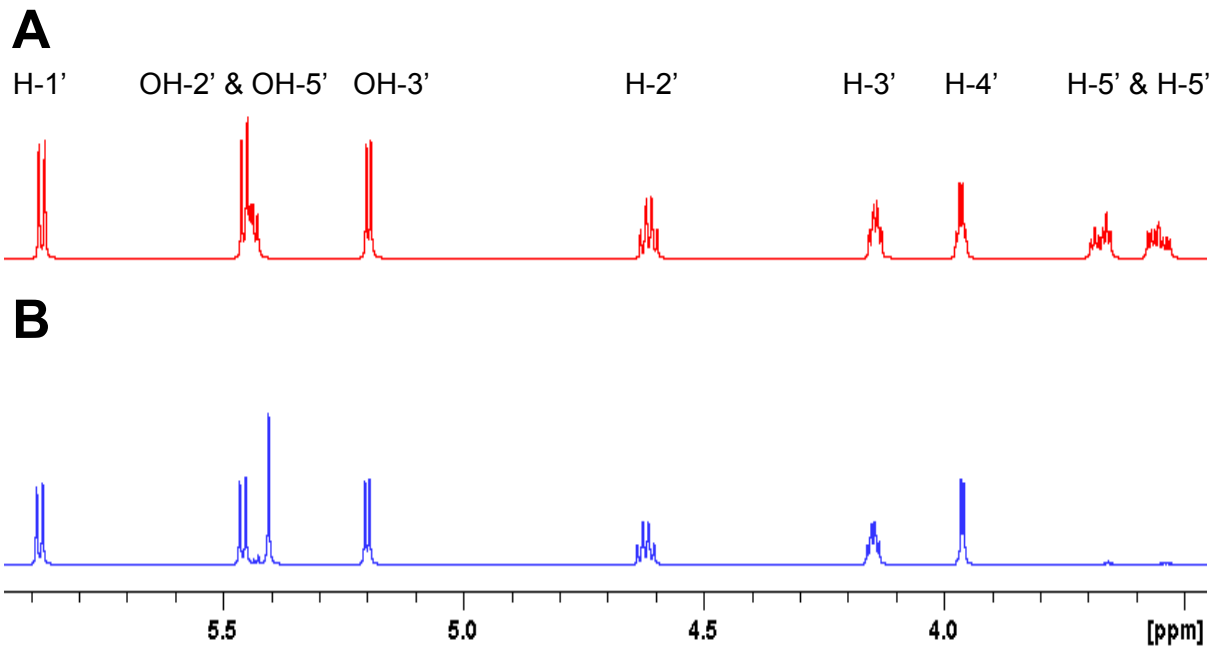
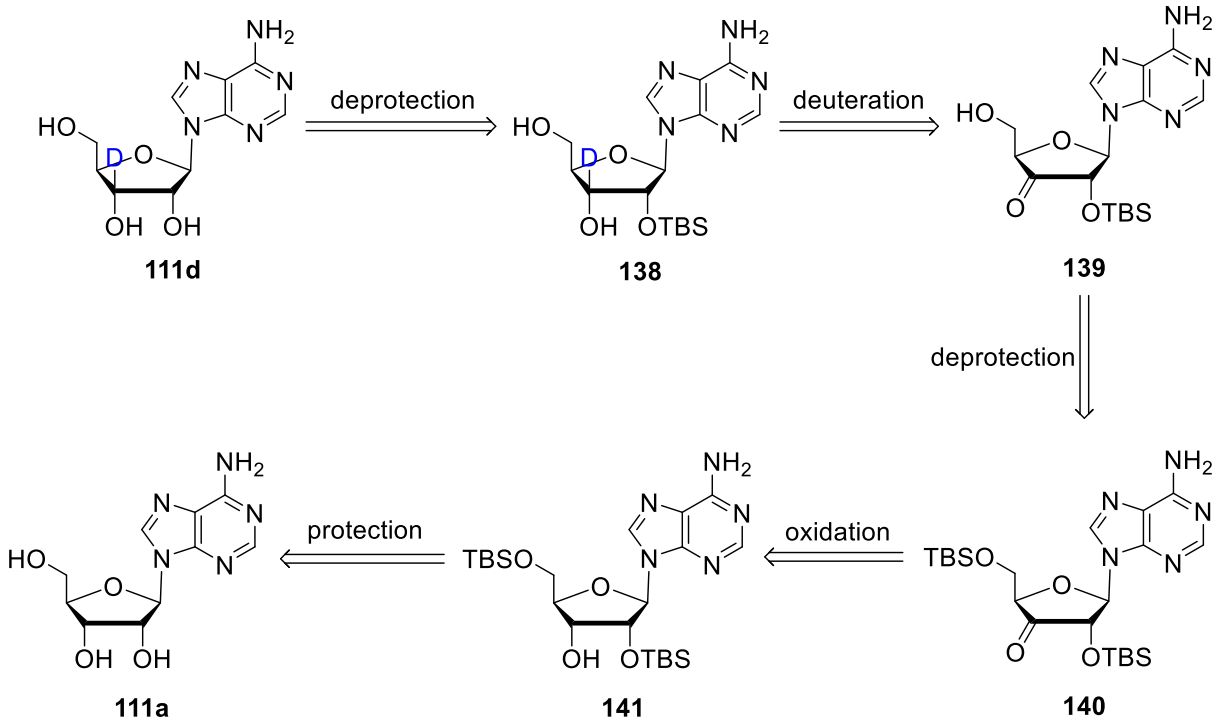


Figure 12. Comparison of ^{1}H-NMR spectrum of:
(**A**) adenosine **111a** in DMSO-*d*^{6}.
(**B**) [5',5'-$^{2}H_2$]-adenosine **111c** in DMSO-*d*^{6}.

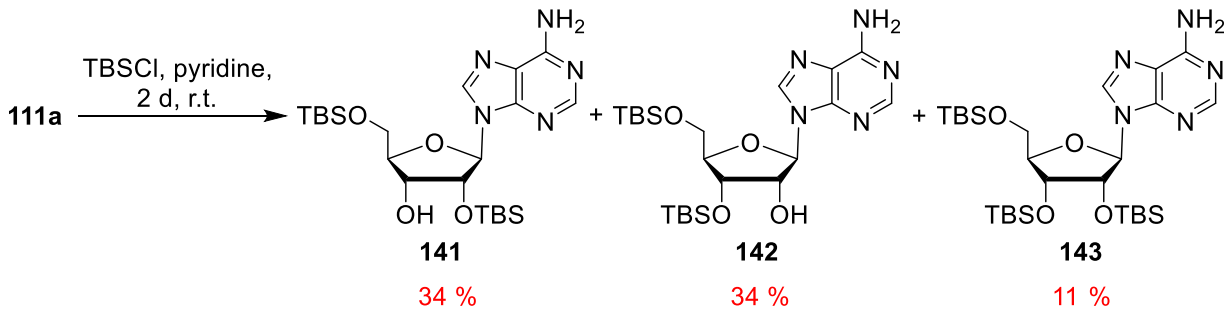
3.5.2. Synthesis of [3'-^{2}H]-adenosine 111d

The synthesis of monodeuterated [3'-^{2}H]-adenosine **111d** has been described in the literature by Robins *et al* in 1997.[26] The strategy involved the protection of the 2'- and 5'-hydroxyl groups of adenosine **111a**, and then oxidation of the free alcohol in **141** to give keto adenosine **140**. Selective deprotection of the primary alcohol of the keto adenosine **140**, followed by ketone reduction with $NaBD_4$ would offer monodeuterated adenosine **138**. Final deprotection of the 2'-hydroxyl would then give [3'-^{2}H]-adenosine **111d**.



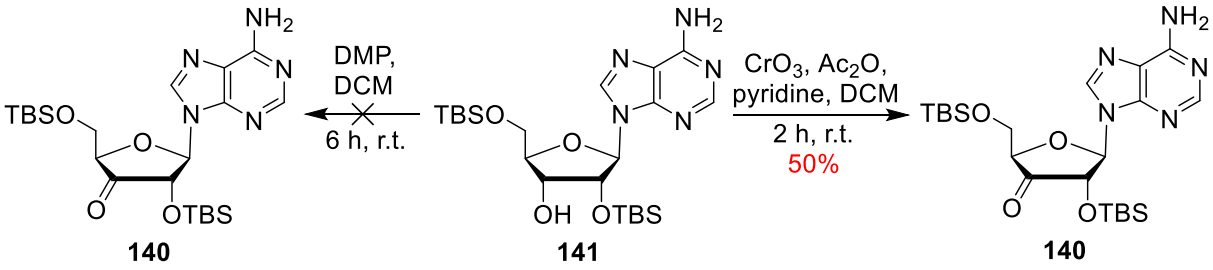
Scheme 16. Retrosynthetic strategy to [3'-^{2}H]-adenosine **111d**.

O-TBS protection of 2'- and 5'-hydroxyl positions of adenosine **111a** was achieved by treating adenosine **111a** with 3 equivalents of *tert*-butyldimethylsilyl chloride in pyridine. The product 2',5'-bis-*O*-(*tert*-butyldimethylsilyl)adenosine **141** was obtained as a mixture along with two side-products 3',5'-bis-*O*-(*tert*-butyldimethylsilyl)adenosine **142** and 2',3',5'-tris-O-(*tert*-butyldimethylsilyl)adenosine **143** in a 3:3:1 ratio. Each isomer was easily separated *via* column chromatography affording the desired product **141** in a relatively low yield (34%). The reaction outcome is summarised in Scheme 17.



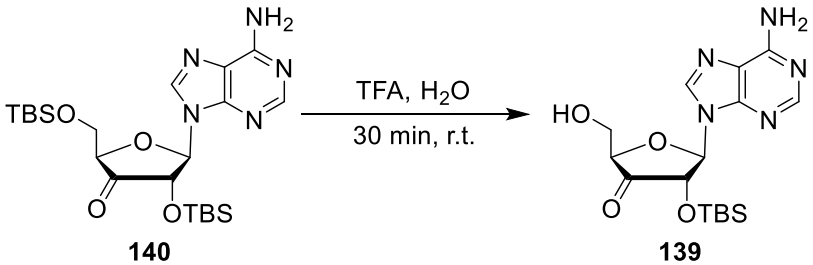
Scheme 17. Synthesis of the TBS protected adenosine **141**.

In the literature, the oxidation of the free 2'-hydroxyl in adenosine **141** was carried out using two different reagents: Dess-Martin periodinane (DMP) and chromium trioxide.[26] In our hands, oxidation of **141** employing DMP failed and only the starting material **141** was recovered. However, oxidation of **141** was successfully achieved using chromium trioxide, acetic anhydride and pyridine in DCM. The reaction mixture was filtered through glass microfibre and purified by flash column chromatography affording the desired product **140** in 50% yield, as shown in Scheme 18.



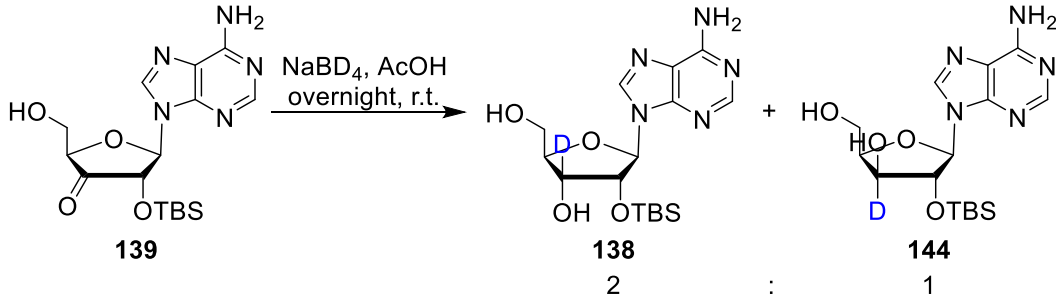
Scheme 18. Oxidation of the TBS protected adenosine **141**.

In order to increase its solubility, selective deprotection of the 5'-hydroxyl group of the protected adenosine **140** was performed using trifluoroacetic acid (TFA) in H_2O affording 2'-*O*-(*tert*-butyldimethylsilyl)adenosine **139**. Under these acidic conditions, primary alcohols are deprotected more easily than secondary. The reaction time was kept short to avoid unwanted deprotection of the 2'-hydroxyl group. Due to its instability, the keto adenosine **139** was not purified and was used immediately in the next reaction.



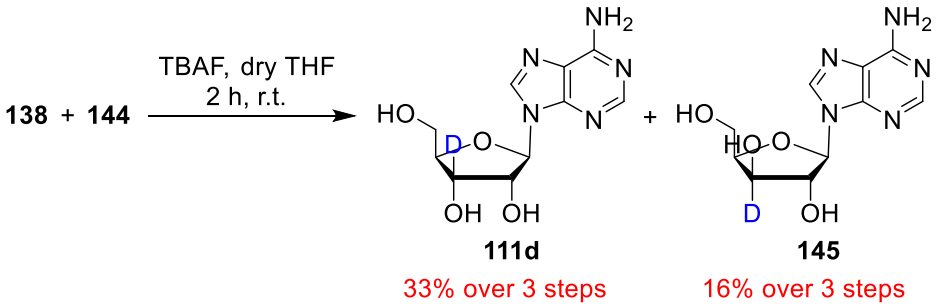
Scheme 19. Deprotection of the TBS protected adenosine **140**.

Unpurified protected adenosine **139** was treated with $NaBD_4$ and acetic acid, generating sodium triacetoxyborodeuteride ($NaBH(OAc)_3$/AcOH) *in situ*. The formation of sodium triacetoxyborodeuteride, an activated deuteride reducing agent, allowed for the reduction of the carbonyl group with incorporation of a deuterium atom at the carbon C-3'. A mixture of two deuterated diastereoisomers **138** and **144** was obtained in a 2:1 ratio, with the *beta* product **138** as the major product. This product mixture was used in the following reaction without purification.



Scheme 20. Reduction of the TBS protected adenosine **139** with $NaBD_4$.

Final deprotection of the **138** and **144** mixture was performed using *tert*-butylammonium fluoride **53** (TBAF) in dry THF, as shown below in Scheme 25. [3'-^{2}H]-Adenosine **111d** was purified by column chromatography (EtOAc/^{i}PrOH/H_2O, 7:2:1) allowing the separation of the two diastereoisomers.



Scheme 21. Final step in the synthesis of [3'-^{2}H]-adenosine **111d**.

The ^{1}H-NMR spectrum of [3'-^{2}H]-adenosine **111d** was compared to that of adenosine **111a** as shown below in Figure 13. The coupling patterns of monodeuterated **111d** were simplified due

to the presence of the isotope. The major observation was the significant reduction of the peak corresponding to H-3' in adenosine **111a**. Moreover, the multiplicity of H-4' peak changed from a multiplet to a triplet which confirmed the presence of one deuterium atom. The coupling patterns of monodeuterated **111d** were simplified due to the presence of the isotope. Finally, there is a clear simplification of the coupling pattern of H-2'. Based on the above mentioned observations, it was concluded that the reduction of keto adenosine **139** with $NaBD_4$ had proven successful. Despite the high purity of [3'-^{2}H]-adenosine **111d** shown by ^{1}H-NMR, mass spectrometry analysis showed the presence of unlabelled adenosine **111a** as a minor component (4%).

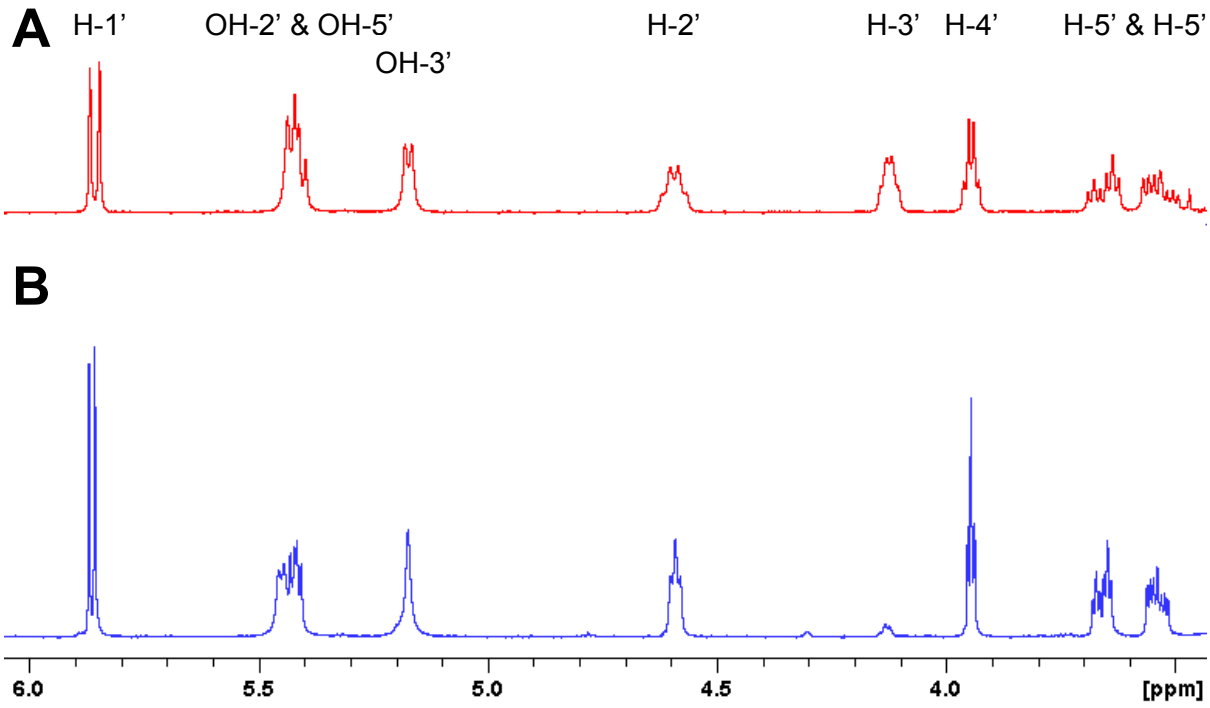
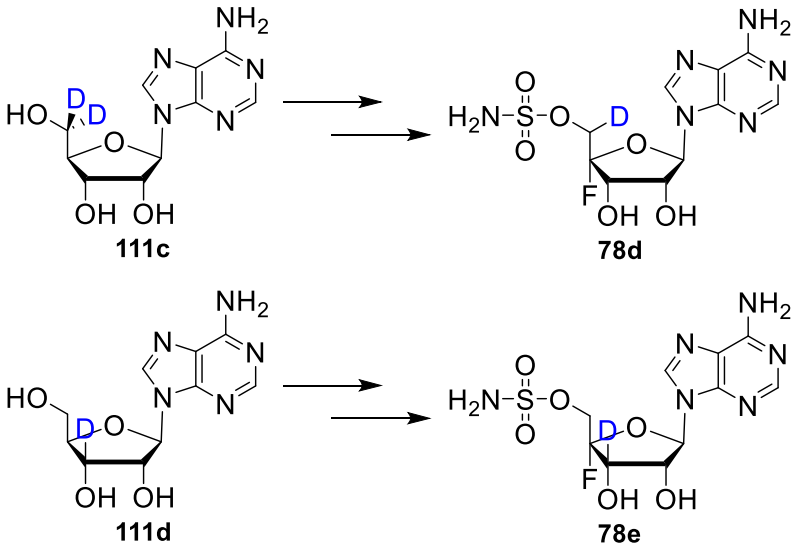


Figure 13. Comparison of ^{1}H-NMR spectrum of:
(**A**) adenosine **111a** in DMSO-d^6.
(**B**) [3'-^{2}H]-adenosine **111d** in DMSO-d^6.

3.5.3. Feeding experiments of labelled adenosines 111c and 111d

Since previous feeding experiments could not confirm that the carbon C-3' of nucleocidin **78a** is derived from the carbon C-1 of glycerol **116a**, a feeding experiment with [3'-^{2}H]-adenosine **111d** was conducted. Moreover, a feeding experiment with [5',5'-$^{2}H_2$]-adenosine **111c** was performed in order to discover whether the loss of deuterium atom at the carbon C-5' of nucleocidin **78a** occurs before or after the formation of adenosine **111a**. The anticipated isotopic incorporation sites for nucleocidin **78a** are illustrated in Scheme 22.



Scheme 22. Anticipated labelling patterns into nucleocidin **78a** from [5',5'-$^{2}H_2$]-adenosine **111c** and [3'-^{2}H]-adenosine **111d**.

A feeding experiment with [5',5'-$^{2}H_2$]-adenosine **111c** was achieved following the same procedure developed previously for that of the labelled glycerols. However, the final concentration of [5',5'-$^{2}H_2$]-adenosine **111c** (4 mM) was lower than the concentration of the glycerols (10 mM) because of the higher toxicity of adenosine **111a**.[27] Unfortunately, no incorporation of isotope was observed into the resultant nucleocidin **78a** in this experiment as judged by ^{19}F{^{1}H}-NMR, as shown below in Figure 14. Moreover, the production of nucleocidin **78a** was substantially diminished in the feeding experiment compared to the control experiment.

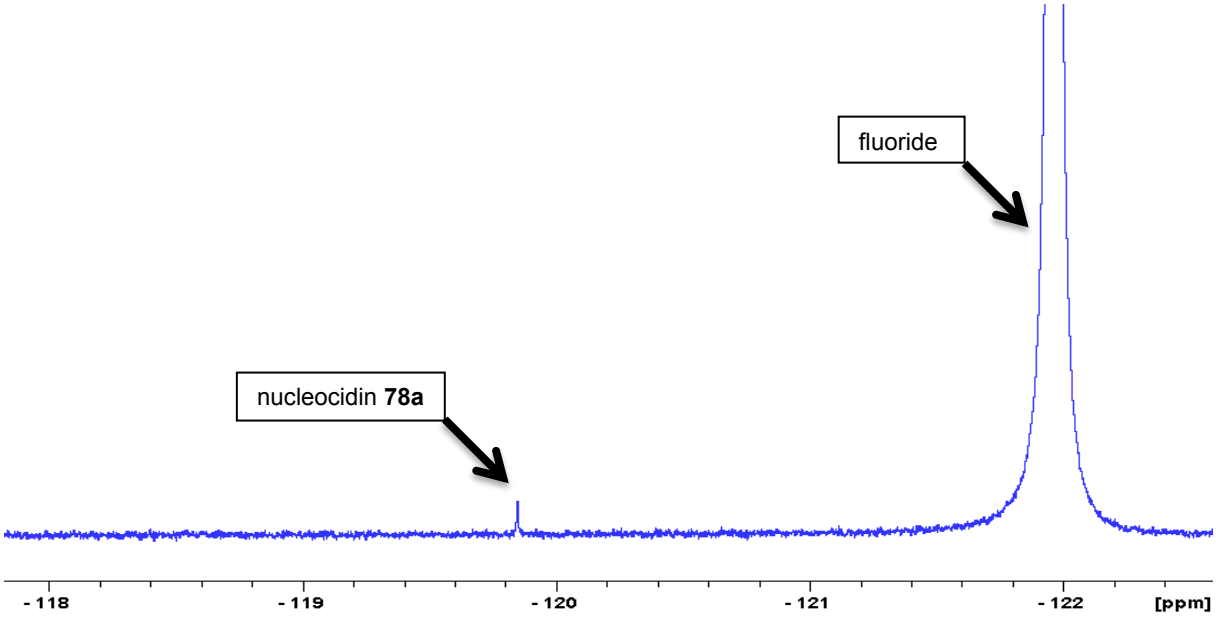


Figure 14. $^{19}F\{^{1}H\}$-NMR spectrum of feeding experiment with [5',5'-$^{2}H_2$]-adenosine **111c** in D_2O.

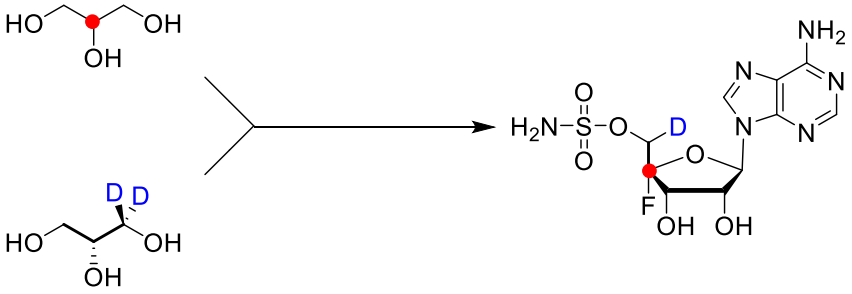
Pulse feeding experiment with [3'-^{2}H]-adenosine **111d** was carried out, but unfortunately, the production of nucleocidin **78a** had ceased during this experiment, as well as during the control experiment conducted in parallel. Therefore, these results could not successfully concluded.

3.6. Conclusion

An in-house strain of *Streptomyces calvus* T-3018 which was known to produce nucleocidin **78a** was acquired from Pfizer. Despite nucleocidin production being very fickle and with very low titres, nucleocidin **78a** was observed by ^{19}F-NMR after *n*-butanol extraction. Due to the structural similarity of adenosine **111a** and nucleocidin **78a** it was then hypothesised that adenosine **111a** may be an intermediate in the biosynthesis of nucleocidin **78a**. According to the pentose phosphate pathway, carbons C-4' and C-5' of the ribose ring of adenosine **111a** are derived from glycerol **116a**, as shown in Figure 15.

Therefore, in order to investigate the biosynthesis of nucleocidin **78a**, feeding experiments with different isotopomers of glycerol were carried out. Fermentation cultures of *Streptomyces calvus* were supplemented with [2-^{13}C]-glycerol **116b** and this experiment showed a 25% incorporation of ^{13}C into carbon C-4' of nucleocidin **78b**. This incorporation was consistent with pentose phosphate pathway involvement in ribose ring synthesis.

(2*S*)-[1-$^{2}H_2$]-Glycerol **116d** and (2*R*)-[1-$^{2}H_2$]-glycerol **116e** were synthesised following the procedure developed by Hill *et al*.[19] Pulse feeding experiment with (2*S*)-[1-$^{2}H_2$]-glycerol **116d** did not show any isotope incorporation into nucleocidin **78a**. This is consistent with the loss of the two pro-S deuterium atoms during the pentose phosphate pathway. Pulse feeding experiments with [1,1,2,3,3-$^{2}H_5$]-glycerol **116c** and (2*R*)-[1-$^{2}H_2$]-glycerol **116e** showed incorporation into the carbon C-5' of nucleocidin **78d**, and with an upfield shift of 0.22 ppm. According to studies by Lambert and Greifenstein, the value of this chemical shift is consistent with the presence of only one deuterium atom attached at the carbon C-5' of nucleocidin **78d** as depicted in Scheme 23.[17,18] The loss of one deuterium atom could be explained by an oxidation process ocurring after ribose ring assembly but prior to, or concominant with, fluorine introduction.



Scheme 23. Incorporation of labelled glycerols **116b** and **116c** into nucleocidin **78a**.[14]

The syntheses of [5',5'-$^{2}H_2$]-adenosine **111c** and [3'-^{2}H]-adenosine **111d** were achieved with high purity. Unfortunately, pulse feeding experiments with [5',5'-$^{2}H_2$]-adenosine **111c** resulted in a low production of nucleocidin **78a** and no observable isotope incorporation. Moreover, feeding experiments with [3'-^{2}H]-adenosine **111d** could not be interpreted due to the lack of nucleocidin **78a** production.

3.7. References

1. S. O. Thomas, V. L. Singleton, J. A. Lowery, R. W. Sharpe, L. M. Pruess, J. N. Porter, J. H. Mowat, N. Bohonos, *Antibiotics Annual*, 1957, 716-721.
2. K. Fukuda, T. Tamura, Y. Segawa, Y. Mutaguchi, K. Inagaki, *Actinomycetologica,* 2009, **23**, 51-55.
3. L. Kalan, A. Gessner, M. N. Thaker, N. Waglechner, X. Zhu, A. Szawiola, A. Bechthold, G. D. Wright, D. L. Zechel, *Chemistry & Biology*, 2013, **20**, 1214-1224.
4. X. M. Zhu, S. Hackl, M. N. Thaker, L. Kalan, C. Weber, D. S. Urgast, E. M. Krupp, A. Brewer, S. Vanner, A. Szawiola, G. Yim, J. Feldmann, A. Bechtold, G. D. Wright, D. L. Zechel, *ChemBioChem*, 2015, **16**, 2498-2506.
5. P. Yague, M. T. Lopez-Garcia, B. Rioseras, J. Sanchez, A. Manteca, *Curr. Trends Microbiol.*, 2012, **8**, 65-73.
6. H. Hu, Q. Zhang, K. Ochi, *J. Bacteriol.*, 2002, **184**, 3984-3991.
7. E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S. A. Darst, *Cell*, 2001, **104**, 901-912.
8. A. R. Maguire, W. D., Meng, S. M. Roberts, A. J. Willets, *J. Chem. Soc. Perkin Trans. I,* 1993, 1795–1808.
9. J. D. Finkelstein, J. J. Martin, *Int. J. Biochem. Cell Biol.*, 2000, **32**, 385-389.
10. N. J. Kruger, A. Von Schaewen, *Current Opinion in Plant Biology*, 2003, **6**, 236-246.
11. J. Caruthers, J. Bosch, F. Buckner, W. Van Voorhis, P. Myler, E. Worthey, C. Mehlin, E. Boni, G. DeTitta, J. Luft, A. Lauricella, O. Kalyuzhniy, L. Anderson, F. Zucker, M. Soltis, G. J. Wim, *PROTEINS: Structure, Function, and Bioinformatics*, 2006, **62**, 338-342.

12. R. Zhang, C. E. Andersson, A. Savchenko, T. Skarina, E. Evdokimova, S. Beasley, C. H. Arrowsmith, A. M. Edwards, A. Joachimiak, S. L. Mowbray, *Structure*, 2003, **11**, 31-42.
13. R. B. Honzatko, H. J. Fromm, *Arch. Biochem. Biophys.*, 1999, **370**, 1-8.
14. A. Bartholomé, J. E. Janso, U. Reilly, D. O'Hagan, *Org. Biomol. Chem.*, 2017, **15**, 61-64.
15. J. T. G. Hamilton, C. D. Murphy, M. R. Amin, D. O'Hagan, D. B. Harper, *J. Chem. Soc. Perkin Trans. 1*, 1998, 759-768.
16. M. W. McDonough, W. A. Wood, *J. Biol. Chem.,* 1961, **236**, 1220-1224.
17. J. B. Lambert, L. G. Greifenstein, *J. Am. Chem. Soc.,* 1973, **95**, 6150-6152.
18. J. B. Lambert, L. G. Greifenstein, *J. Am. Chem. Soc.,* 1974, **96**, 5120-5124.
19. R. E. Hill, A. Iwanow, B. G. Sayer, W. Wysocka, I. D. Spencer, *J. Biol. Chem.*, 1987, **262**, 7463-7471.
20. J. Nieschalk, J. T. G. Hamilton, C. D. Murphy, D. B. Harper, D. O'Hagan, *Chem. Comm.*, 1997, 799-800.
21. L. F. Fieser, M. Fieser, *Reagents for Organic Synthesis*, 1967, 581-595.
22. Y-R. Chen, F. W. Larimer, E. H. Serpersu, F. C. Hartman, *J. Biol. Chem.,* 1999, **22**, 2132-2136.
23. H. Pang, K. H. Schram, D. L. Smith, S. P. Gupta, L. B. Townsend, J. A. McCloskey, *J. Org. Chem.*, 1982, **47**, 3923-3932.
24. G. Zhang, S. L. Richardson, Y. Mao, R. Huang, *Org. Biomol. Chem.*, 2015, **13**, 4149-4154.
25. J. B. Epp, T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293-295.
26. M. J. Robins, S. Sarker, V. Samano, S. F. Wnuk, *Tetrahedron*, 1997, **53**, 447-456.
27. S. Archer, P. F. Juranka, J. H. Ho, V. L. Chan, *J. Cell. Physio.*, 1985, **124**, 226-232.

4. Exploring production of nucleocidin 78a in cell-free extracts of *S. calvus*

4.1. Synthesis of putative substrates

As previously described in **Chapter 3**, cell-free extract experiments of *Streptomyces calvus* were performed concomitantly alongside pulse feeding experiments using various isotopically labelled substrates. These experiments revealed that glycerol is incorporated into the ribose ring moiety of nucleocidin **78a**. Moreover, pulse feeding experiments with (2*R*)-[1-$^{2}H_{2}$]-glycerol **116e** suggested that a hydrogen atom is lost at the carbon C-5' prior to or concominant with fluorine introduction.[1] In order to further elucidate the biosynthetic pathway of nucleocidin **78a** in *Streptomyces calvus*, the synthesis of additional putative substrates was investigated for use in cell-free extract experiments. Subsequently, each putative substrate that was synthesised shared a relative structural homology with either adenosine **111a** or nucleocidin **78a** as shown in Figure 1.

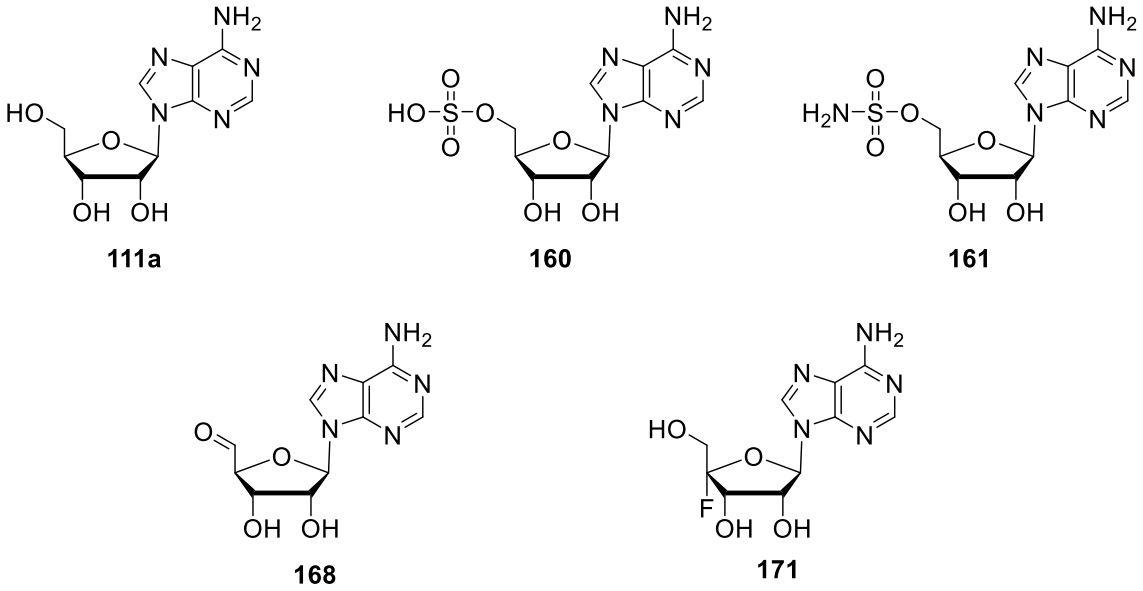


Figure 1. Structures of the putative intermediates in the biosynthesis of nucleocidin **78a**.

As mentioned previously, adenosine **111a** was thought to be a potential early intermediate in the metabolic pathway of nucleocidin **78a**. The synthesis of adenosine sulfate **160** and sulfamoyladenosine **161** was achieved, and each was used in separate cell-free extract experiments of *S. calvus*. Adenosine sulfate **160** and sulfamoyladenosine **161** were chosen as potential biosynthetic precursors for these experiments due to their structural similarity to that of nucleocidin **78a**. The aldehyde **168** was considered as a putative substrate of the fluorinating enzyme, which would explain the loss of deuterium atom observed during the previous feeding experiments with [1,1,2,3,3-$^{2}H_{5}$]-glycerol **116c** and (2*R*)-[1-$^{2}H_{2}$]-glycerol **116e**. The proposed mechanism for the fluorination of the aldehyde **168** is described in detail in the section **4.1.3**. Finally, the synthesis of fluoroadenosine **171** was performed and cell-free extracts were carried out in order to determine whether it is a fluorinated intermediate that can be biotransformed within the cell to nucleocidin **78a**. The syntheses of these putative substrates are described below.

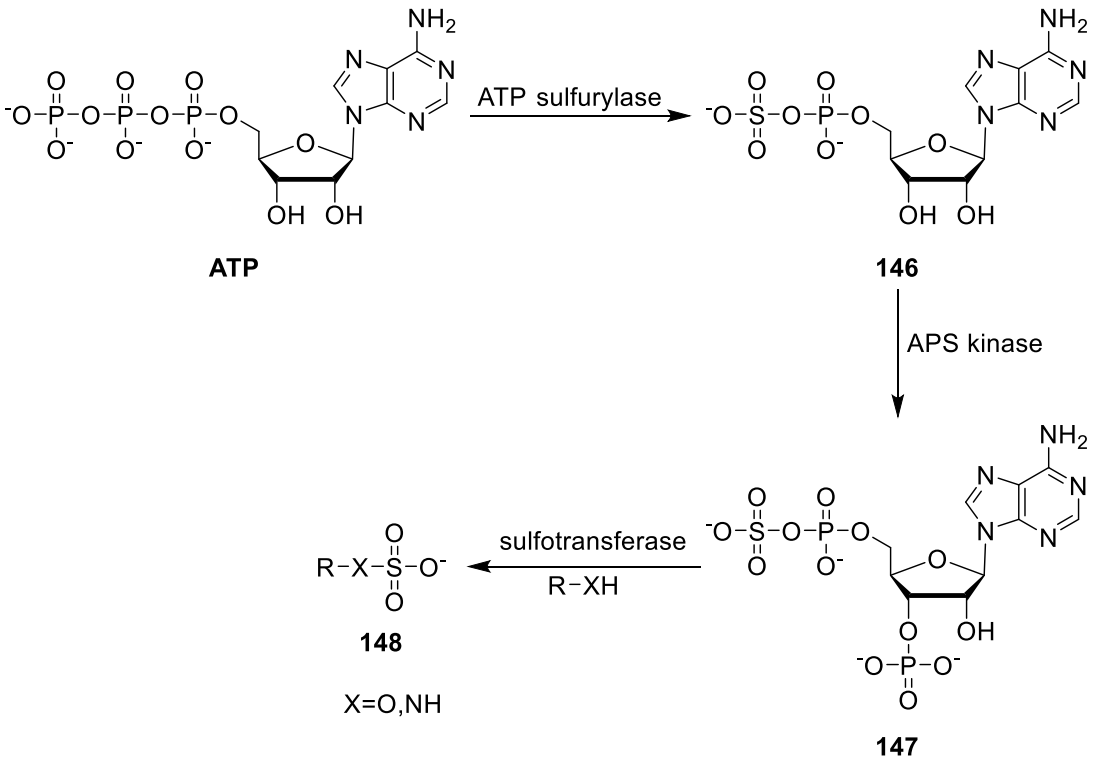
4.1.1. Synthesis of adenosine sulfate 160

4.1.1.1. Sulfation in nature

Sulfation is a major industrial process which is used to manufacture a wide range of products such as pigments, drugs, detergents, pesticides, *etc*.[2,3] In nature, sulfation is also very important and is utilised in a number different roles, for instance to generate reservoirs of bioactive molecules and to enhance the elimination of potential toxic agents.[4] Indeed, sulfation is used in nature to regulate the bioactivity of biomolecules. Steroid sulfates for instance, which are not active in themselves, are hydrolysed by a sulfatase enzyme in order to become active.[4] Toxic xenobiotic molecules are another example of a class of compound which can be derivatised through sulfation.[5]

During sulfation, the sulfate group must first be activated *via* the formation of adenosyl phosphosulfate **146** by an ATP sulfurylase enzyme. Phosphorylation of adenosyl

phosphosulfate **146** with ATP affords phosphoadenosine phosphosulfate **147**, which is a sulfate group donor. Sulfotransferase enzymes are able to catalyse the transfer of the sulfate group from phosphoadenosine phosphosulfate **147** to an acceptor, such as an alcohol or amine, as shown in Scheme 1.[6,7]



Scheme 1. Biological sulfation *via* formation of adenosyl phosphosulfate **146**.

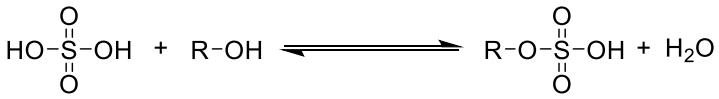
4.1.1.2. Sulfating reagents

The chemical introduction of a sulfate group into molecules has become the focus of many investigations. After sulfation, the physico-chemical properties of the compounds radically change. For instance, most sulfated molecules are water soluble which can make purification using conventional organic solvents more difficult. The stability of the sulfate group is also very low in acidic or under high temperature conditions.

Sulfur trioxide (SO_3) is a strong electrophilic reagent which quickly reacts with any molecule containing an electron donating group. In order to alleviate this issue of high reactivity, sulfur

trioxide is diluted or complexed with different agents such as H_2O, giving sulfuric acid **149** (H_2SO_4), or hydrochloric acid, affording chlorosulfonic acid **150**.

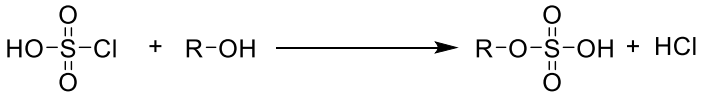
In the beginning of the 20th century, sulfation reactions were usually carried out with sulfuric acid **149**, as shown in Scheme 2.[8] Indeed, sulfation of alcohols was performed using sulfuric acid **149**, but the yield is limited due to the release of H_2O which forms during the reaction.[9] The subsequent use of Dean-Stark apparatus increased the yield by removing the water.



149 **151** **152**

Scheme 2. Sulfation of alcohols with sulfuric acid **149**.

In 1854, chlorosulfonic acid **150** was prepared using phosphorus pentachloride and concentrated sulfuric acid by Williamson *et al*.[10] Three years later, chlorosulfonic acid **150** was produced using an alternative method by reacting hydrochloric acid with sulfur trioxide.[11] Many different methods of production of chlorosulfonic acid **150** were then developed, such as *via* the reaction of carbon tetrachloride with sulfuric acid **149** or the action of phosphorus trichloride on sulfuric acid **149**. Chlorosulfonic acid **150** is a strong acid and extremely hygroscopic due to its propensity to react with H_2O to form sulfuric acid **149** and hydrochloric acid. Sulfation of alcohols carried out with chlorosulfonic acid **150** released hydrochloric acid, as shown in Scheme 3.[12]

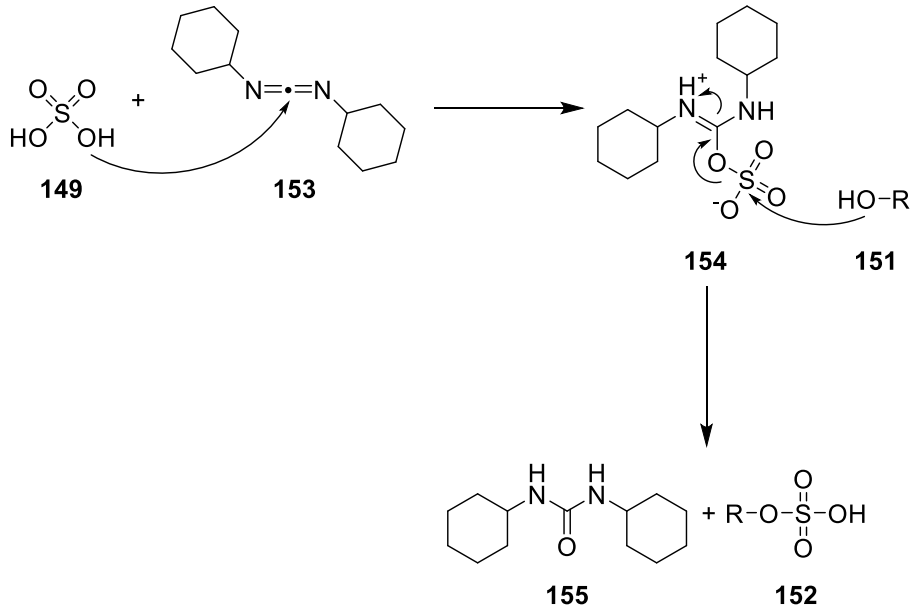


150 **151** **152**

Scheme 3. Sulfation of alcohols with chlorosulfonic acid **150**.

In 1966, another sulfating method was established for the sulfation of alcohols with sulfuric acid **149** and dicyclohexylcarbodiimide **153** (DCC).[13] The ratio and the order of addition of the

reactants were found to be important for achieving the desired sulfated product. Indeed, the alcohol, sulfuric acid **149** and DCC **153** were allowed to react in a 1:1:5 ratio. Moreover, the substrate was added to a solution of DCC **153** in DMF, which was followed by the addition of sulfuric acid **149**. At first the mechanism was not clearly elucidated, but it was thought that sulfuric acid **149** reacts with DCC **153** forming the protonated DCC-H_2SO_4 intermediate **154**. Finally, nucleophilic attack of the alcohol onto the sulfur atom would form the sulfate **152** releasing dicyclohexylurea **155**, as shown in scheme 4.

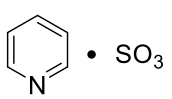


Scheme 4. Mechanism of sulfation with sulfuric acid **149** and DCC **153**.

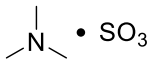
Although this reaction was in general successful, the relative insolubility of dicyclohexylurea **155** rendered the purification of the product rather difficult and despite the fact that the major products with such reagents are mono sulphates, many side products were also observed.[14] Moreover, this mixture of sulfuric acid **149** and DCC **153** can only generate sulfates with specific substrates which are not sensitive to acidic conditions.

In the 1980s, a new type of sulfating reagents was developed by complexing sulfur trioxide with amines such as trimethylamine, triethylamine, pyridine, *etc.*[15] Due to the difficulty of handling sulfur trioxide as a gas or liquid, these sulfur trioxide-amine complexes were

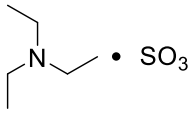
preferred over sulfur trioxide. These organic complexes can be formed by bubbling sulfur trioxide through a solution of the desired organic base.[16] The complexes are solids at room temperature and somewhat stable at high temperatures.



sulfur trioxide pyridine complex 156



sulfur trioxide trimethylamine complex 157



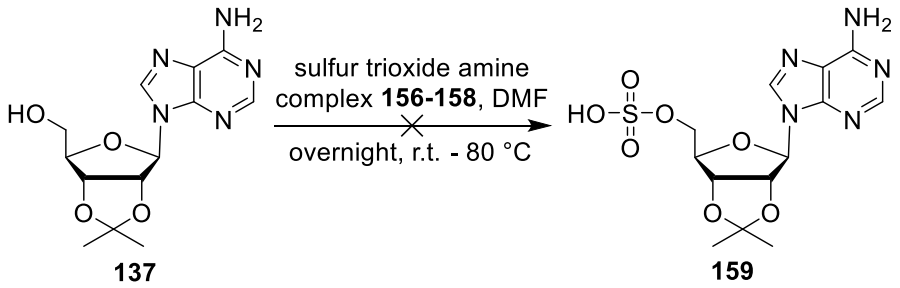
sulfur trioxide triethylamine complex 158

Figure 2. Sulfur trioxide amine complexes.

These sulfur trioxide amine complexes allowed for an easier purification when compared to the previous sulfating reagents. The sulfation of acidic hydroxyl groups was found to be optimal with use of the sulfur trioxide pyridine complex **156**. However, sulfation of hindered substrates was easier with the less bulky sulfur trioxide amine complexes such as **157** and **158**.

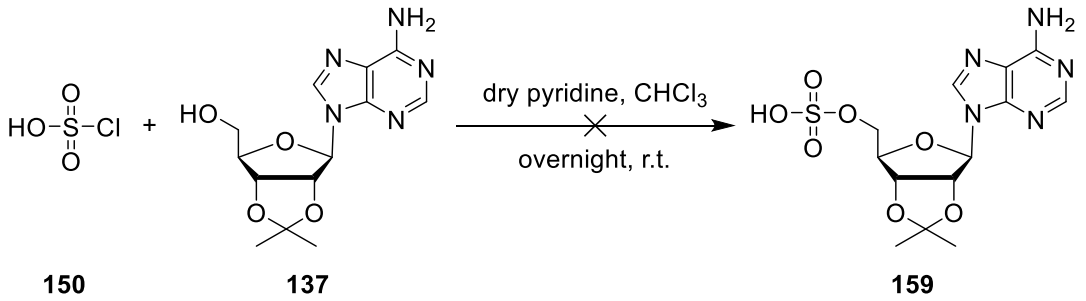
4.1.1.3. Preparation of adenosine sulfate 160

Sulfation of acetonide protected adenosine analogues using sulfur trioxide pyridine complex **156** has been previously described in a patent.[17] Since acetonide adenosine **137** was acquired previously during the synthesis of [5',5'-$^{2}H_2$]-adenosine **111c**, as mentioned in **Chapter 3**, sulfation of this molecule was carried out with sulfur trioxide pyridine complex **156**. Unfortunately, this sulfation reaction was not successful since only the starting material **137** was recovered. The lack of reactivity in this instance may be due to steric hindrance. Therefore, the sulfating reagent was changed to sulfur trioxide trimethylamine complex **157** and then to sulfur trioxide triethylamine complex **158**, but in both cases a complex mixture was obtained which could not be purified by column chromatography. Increasing the temperature and the reaction time also did not show any improvement.



Scheme 5. Sulfation of acetonide **137** with different sulfur trioxide amine complexes.

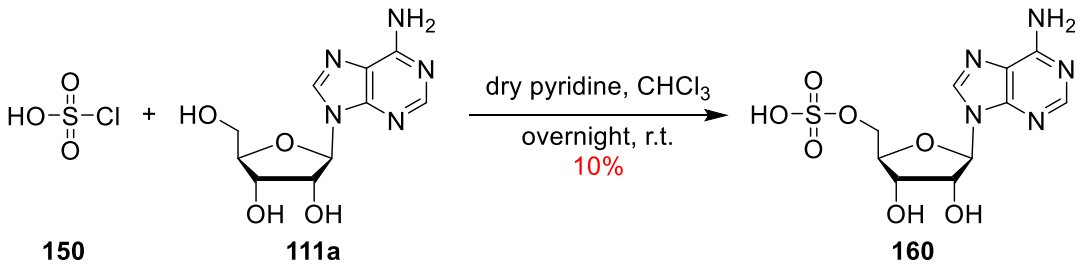
In 1956, sulfation of acetonide **137** using chlorosulfonic acid **150** followed by deprotection of the acetal with acetic acid was described.[18] As such this procedure was followed, but a complex mixture was obtained and the product could not be purified by ion exchange column or HPLC.



Scheme 6. Sulfation of acetonide **137** with chlorosulfonic acid **150**.

A method for the sulfation of adenosine **111a** using chlorosulfonic acid **150** was developed by Egami *et al*.[19] Adenosine **111a** was treated with chlorosulfonic acid **150** in a mixture of dry pyridine and chloroform. After eluting through a short Discovery® DSC-NH_2 SPE ion exchange column, the residue was then purified by semi-preparative HPLC, which was aided by the water solubility of adenosine sulfate **160**. The use of a semi-preparative HPLC allowed for the isolation of the desired product **160** in high purity, which is essential for subsequent cell-free extract experiments. A reverse phase C-18 column was used with a flow rate of 10 mL/min and an isocratic mobile phase of 95% buffer A (0.05% TFA/H_2O) and 5% buffer B (0.05% TFA/ACN). The adenosine sulfate **160** possessed a retention time of 7.4 minutes. The desired

fraction was collected, concentrated and freeze-dried affording a pure sample in 10% yield, as shown in Scheme 7.



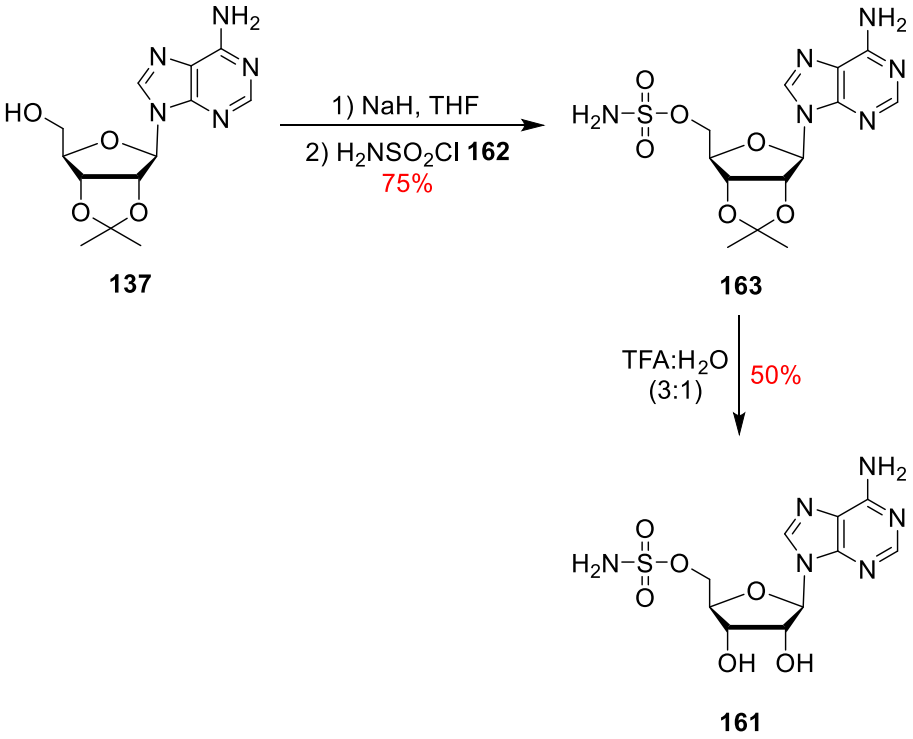
150 111a 160

Scheme 7. Sulfation of adenosine **111a** with chlorosulfonic acid **150**.

4.1.2. Synthesis of sulfamoyl adenosine 161

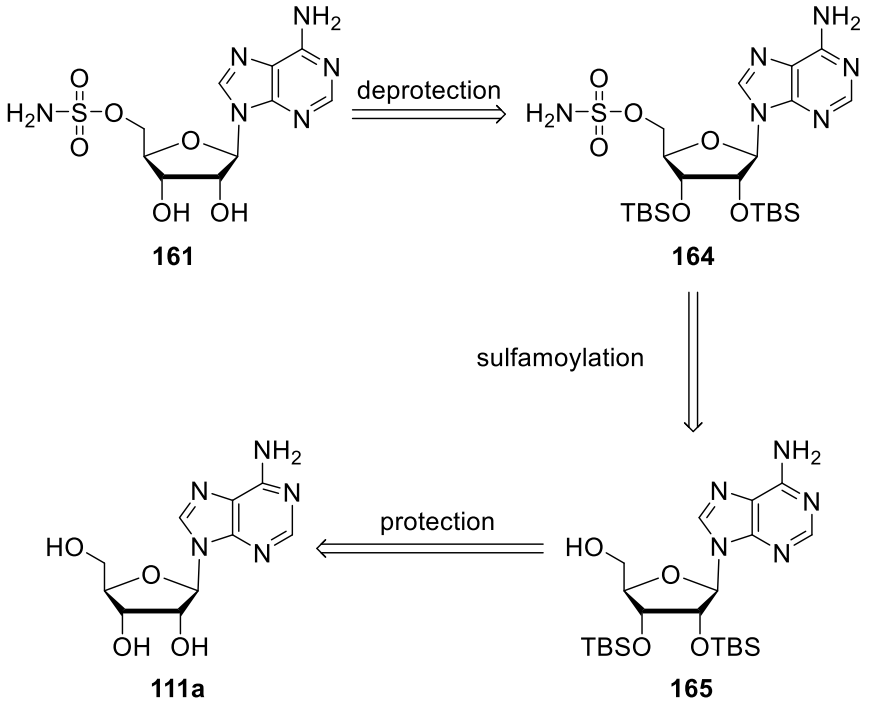
In nature, the mechanism and enzymes responsible for the sulfamoylation reaction are yet to be elucidated. However, it is commonly thought that sulfamoylation in nature occurs in two distinct steps. It initiates with a sulfation reaction, as described previously in section **4.1.1.1**, which is followed by the replacement of a hydroxyl group of the sulfonate with an amine group, performed by amidinotransferase enzymes, to generate the sulfamoyl group.[20]

The synthesis of sulfamoyl adenosine **161** was described in the literature by Lukkarila *et al* as shown in Scheme 8.[21] Sulfamoylation of the protected acetonide **137** was achieved using sodium hydride and sulfamoyl chloride **162** (75%). Final deprotection of the protected sulfamoyl adenosine **163** was carried out using a mixture of TFA and H_2O in a 3:1 ratio affording sulfamoyl adenosine **161**. The overall yield of these two reactions is 37.5%, primarily due to the relatively low yield of the last step (50%). This low yield can be explained by the harsh conditions (TFA/ H_2O) destabilising the sulfamoyl group.



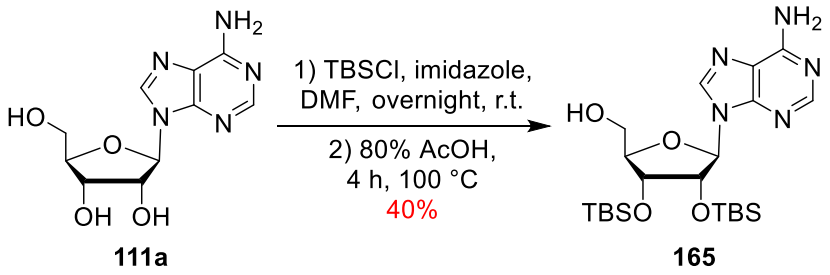
Scheme 8. Synthesis of sulfamoyl adenosine **161** by Lukkarila *et al*.[21]

The procedure developed by Lukkarila *et al*. was not followed directly, but modified in order to increase the overall yield of sulfamoyl adenosine **161**, as shown in Scheme 9. Instead of preparing acetonide **137**, it was envisaged that the secondary hydroxyl groups of adenosine **111a** would be TBS protected giving adenosine **165**. Sulfamoylation using a solution of amidosulfonyl chloride **162** in dry ACN would generate **164** and final deprotection would be achieved under milder conditions using TBAF **53** rather than a mixture TFA and H_2O. TBAF **53** should not interfere with the sulfamoyl group of sulfamoyl adenosine **164**.



Scheme 9. Retrosynthetic strategy to sulfamoyl adenosine **161**.

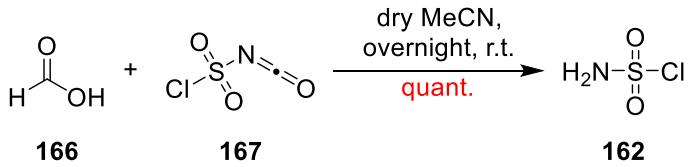
Protection of adenosine **111a** was carried out using TBSCl and imidazole in DMF.[22] After work-up, the residue was dissolved in 80% AcOH and stirred for 4 h at 100 °C. This second step allows for deprotection of any primary silyl ether. In the literature, the TBS protected adenosine **165** was used without further purification.[22] However, in this case the reaction mixture was purified by flash column chromatography to afford **165** in relatively low yield (40%).



Scheme 10. Synthesis of the TBS protected adenosine **165**.

In order to perform the sulfamoylation of the protected adenosine **165**, amidosulfonyl chloride **162** was required. This was prepared as a 2 M solution of amidosulfonyl chloride **162** in dry ACN following the procedure developed by Townsend *et al*.[23] Formic acid **166** was added to

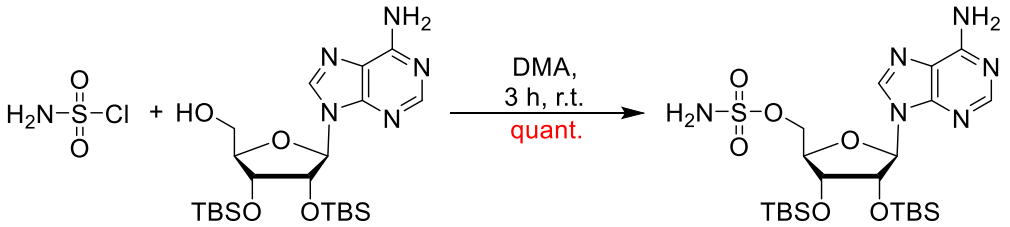
a solution of chlorosulphonyl isocyanate **167** in dry ACN allowing for the formation of amidosulfonyl chloride **162** and the release of carbon dioxide and carbon monoxide. The reaction was considered to have proceeded quantitatively and as such the reaction mixture was used without further purification.



166 **167** **162**

Scheme 11. Synthesis of a 2 M solution of amidosulfonyl chloride **162** in dry ACN.

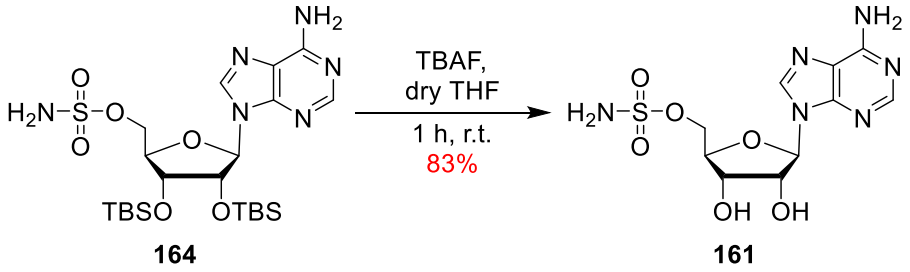
A cold solution of amidosulfonyl chloride **162** in dry ACN was added to a solution of the protected adenosine **165** in DMA.[24] The reaction was monitored by TLC and was stirred until the conversion was complete. Therefore, the reaction time was increased until complete conversion observed by TLC. The product was then purified by flash column chromatography affording the desired product **164** in quantitative yield.



162 **165** **164**

Scheme 12. Synthesis of the TBS protected sulfamoyl adenosine **164**.

Removal of the silyl ether protecting groups was achieved using TBAF **53** in THF.[25] The resultant sulfamoyl adenosine **161** was water soluble and was purified by semi-preparative HPLC due to its water solubility. A reverse phase C-18 column was used with a flow rate of 10 mL/min and an isocratic mobile phase of 95% buffer A (0.05% TFA/H_2O) and 5% buffer B (0.05% TFA/ACN). Sulfamoyl adenosine **161** possessed a very short retention time of 5.6 minutes due to its high polarity. The fraction containing the desired product **161** was collected, concentrated and freeze-dried affording a pure sample in 83% yield, as shown in Scheme 13.



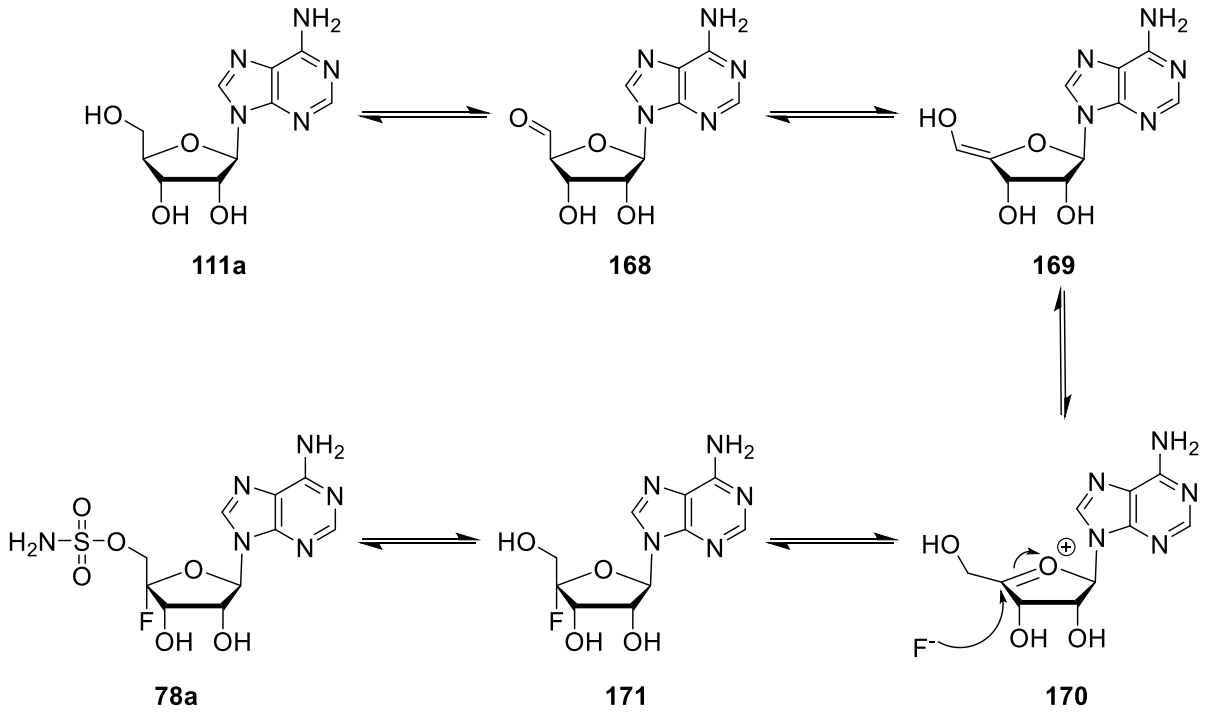
164 **161**

Scheme 13. Synthesis of sulfamoyl adenosine **161**.

In conclusion, a new synthetic route affording sulfamoyl adenosine **161** was successfully developed. The overall yield of this procedure is better (33%) than the synthesis described by Lukkarila *et al*.[21] This can be explained by the fact that the final deprotection with TBAF **53** is performed under milder conditions, which do not destabilise the sulfamoyl group of sulfamoyl adenosine **161**, unlike the use of a mixture of TFA and H_2O.

4.1.3. Synthesis of aldehyde 168

Previous pulse feeding experiments using [1,1,2,3,3-$^2H_5$]-glycerol **116c** and (2*R*)-[1-$^2H_2$]-glycerol **116e**, described in **Chapter 3**, showed a loss of a deuterium atom upon their incorporation into nucleocidin **78a**.[1] In order to explain this loss of deuterium, one of the working hypotheses was that an oxidation of the primary hydroxyl group of adenosine **111a** afforded the aldehyde **168**. This aldehyde emerged therefore as a putative substrate for the fluorinating enzyme in *Streptomyces calvus*. Aldehyde **168** has the potential to undergo tautomerisation, followed by a rearrangement to generate the oxonium species **170**. Fluoride could then react with oxonium **170** affording the fluoroadenosine **171**. Finally, fluoroadenosine **171** would be further metabolised generating nucleocidin **78a**, as shown in Scheme 14.



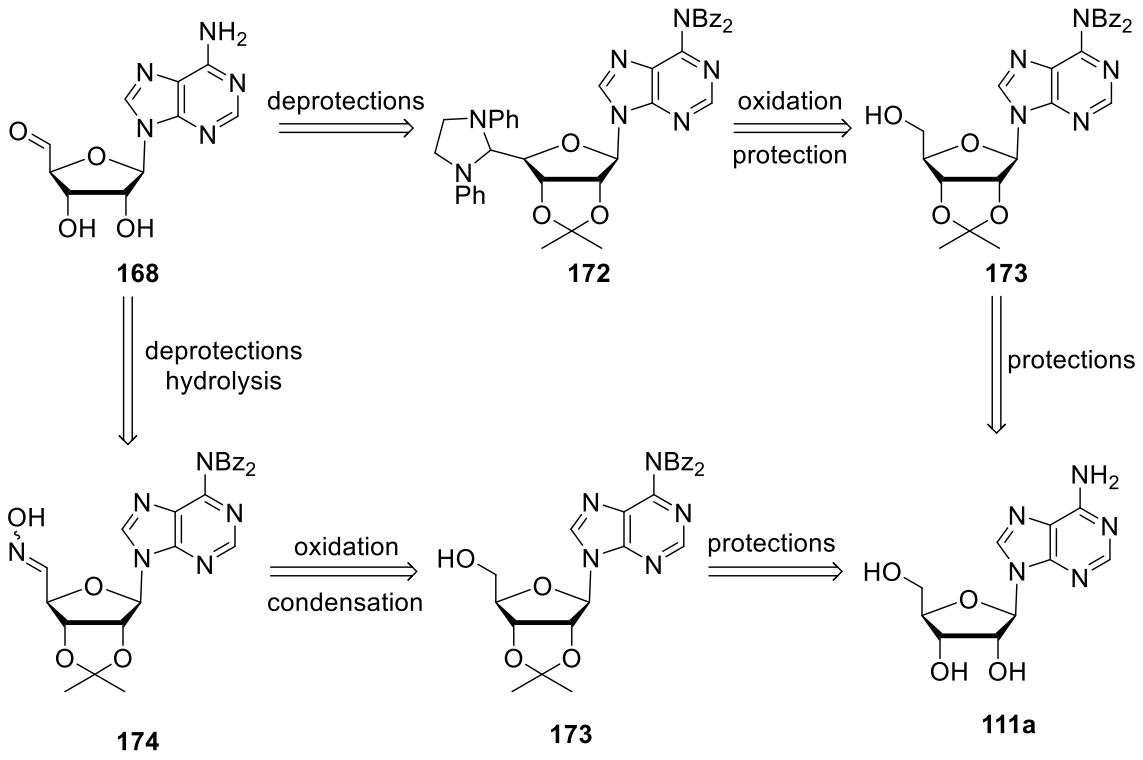
Scheme 14. Putative metabolic pathway of nucleocidin **78a**.

In order to test this working hypothesis, a synthesis of the aldehyde **168** was developed. Cell-free extract experiments with this aldehyde **168** were then explored to determine whether this putative intermediate is a substrate of the fluorination enzyme.

Two different strategies for the synthesis of the aldehyde **168** were reported by Liu *et al*.[26] and by Wnuk *et al*.[27], as illustrated in Scheme 15. Both strategies begin with the conversion of adenosine **111a** to the acetonide **168**, in which the primary hydroxyl-group is subjected to Moffat oxidation affording the aldehyde intermediate. Due to the instability of the aldehyde, the carbonyl group was then directly protected using *N,N'*-diphenylethylenediamine according to Liu *et al*. to give aminal **172**,[26] while Wnuk *et al*. reacted this aldehyde intermediate with hydroxylamine forming a stable oxime **174**.[27] Finally, hydrolysis of the oxime **174** and deprotection of the aminal **174** lead to the desired aldehyde **168**.

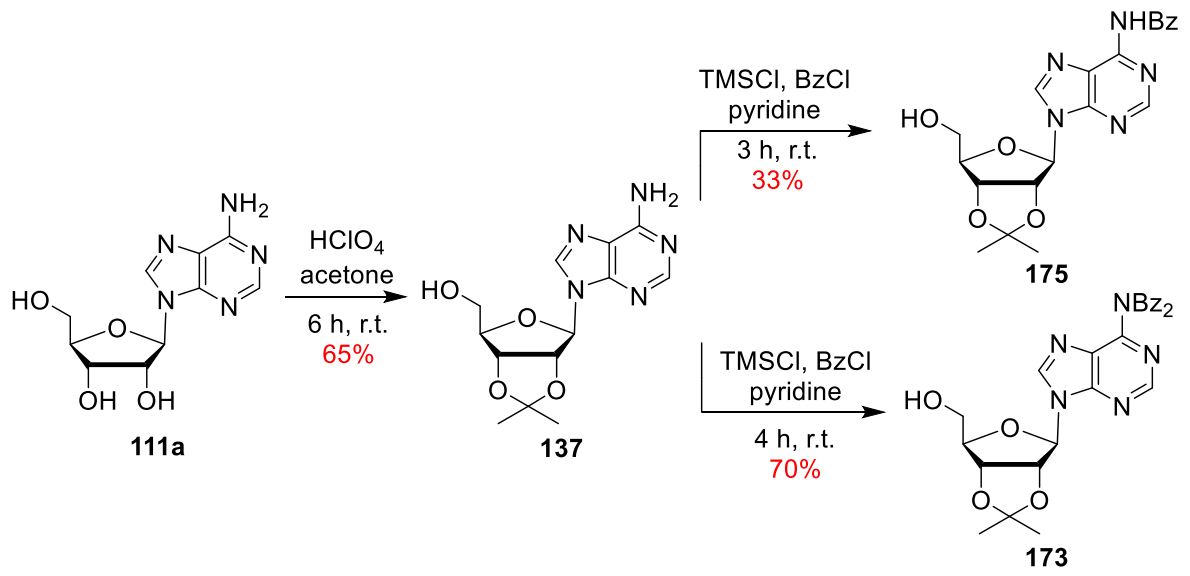
Both synthetic strategies offer access to the putative intermediate **168** in modest yield and in the same number of synthetic steps. However, Wnuk *et al*.'s method gives oxime **174**, which

could be potentially difficult to deal with,[28] and for this reason it was decided to first pursue Liu *et al*.'s synthetic strategy.



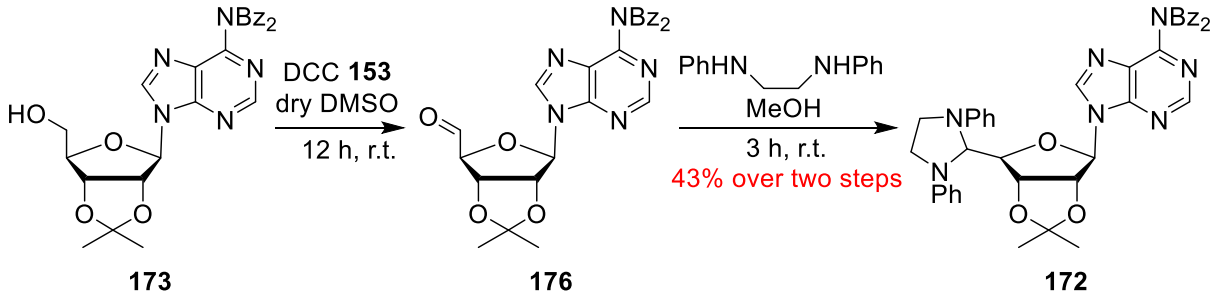
Scheme 15. Retrosynthetic strategies to aldehyde **168**.[26,27]

Treatment of adenosine **111a** with perchloric acid in acetone afforded acetonide **137** in good yield. For the protection of the amine, two methods were attempted. *O*-TMS protection of the primary hydroxyl group in acetonide **137** was achieved with TMSCl in pyridine for 1 h, followed by benzoylation of the amino-group with benzoyl chloride affording the monobenzoyladenosine **175** in 33% yield. Purification of **175** by column chromatography proved to be inefficient and time consuming. Due to the poor yield obtained, dibenzoylation of the amino-group was then explored using an excess of benzoyl chloride to afford the dibenzoylated amino-moiety in 70% yield. In this event, purification of protected adenosine **173** was straightforward; therefore the dibenzoylation protection strategy was favoured, as shown in Scheme 16.



Scheme 16. Protection of adenosine **111a** and benzoylation of acetonide **137**.

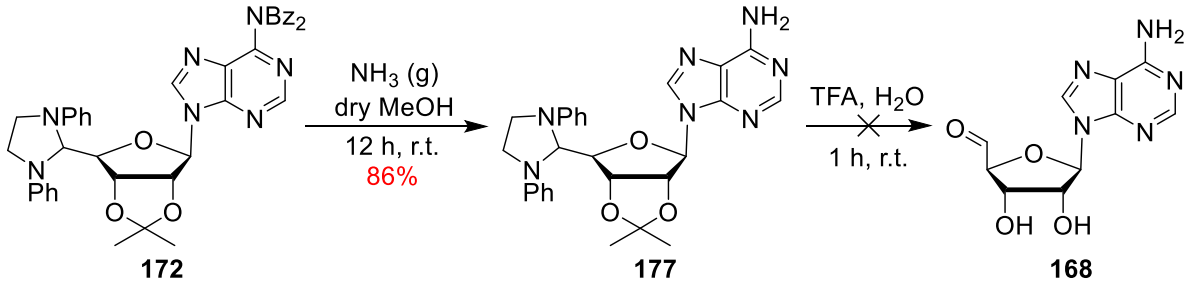
The protected adenosine **173** was subjected to Moffat oxidation (using *N,N'*-dicyclohexylcarbodiimide **153** in dry DMSO) to afford the unstable aldehyde **176**, which was immediately protected by adding *N,N'*-diphenylethylenediamine in methanol, to furnish the aminal **172** in 43% yield over 2 steps, as shown in Scheme 17.[29-31]



Scheme 17. Oxidation of protected adenosine **173**, followed by direct protection of the aldehyde **176**.

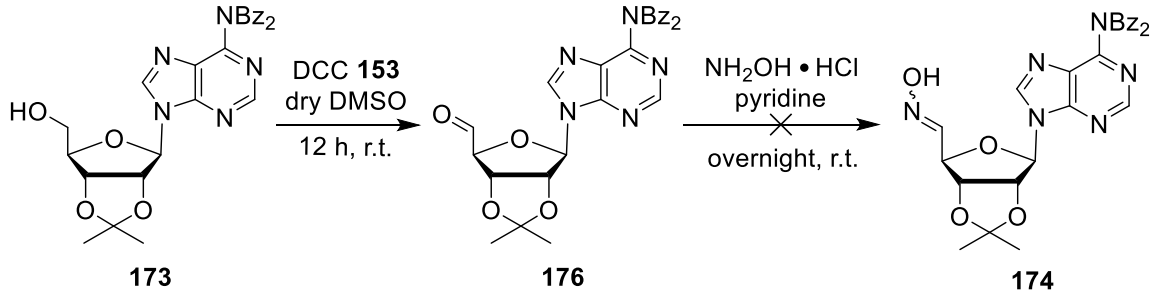
For deprotection of amine group, the aminal **172** was dissolved in dry methanol and ammonia gas was bubbled into the reaction mixture. After purification by column chromatography, free amine **177** was obtained in good yield as illustrated in Scheme 18. Treatment of **177** with a mixture of trifluoroacetic acid and H_2O to simultaneously deprotect the diol and the aldehyde, gave a complex mixture as revealed by ^{1}H-NMR. This mixture of trifluoroacetic acid and H_2O

proved to be too harsh for the deprotection of **177**; therefore, other acid reagents such as hydrochloric acid and sulfuric acid **149** were investigated, but these ultimately presented the same unsuccessful outcome.



Scheme 18. Deprotection of protected aldehyde **172** and attempt to synthesise free aldehyde **168**.

In view of the difficulty in the final deprotection, the procedure developed by Wnuk *et al*. was explored.[27] The protected adenosine **173** underwent Moffat oxidation affording aldehyde **176**, which was immediately reacted with hydroxylamine hydrochloride. Unfortunately, the desired product **174**, detected by mass spectrometry analysis, was too unstable and could not be isolated.

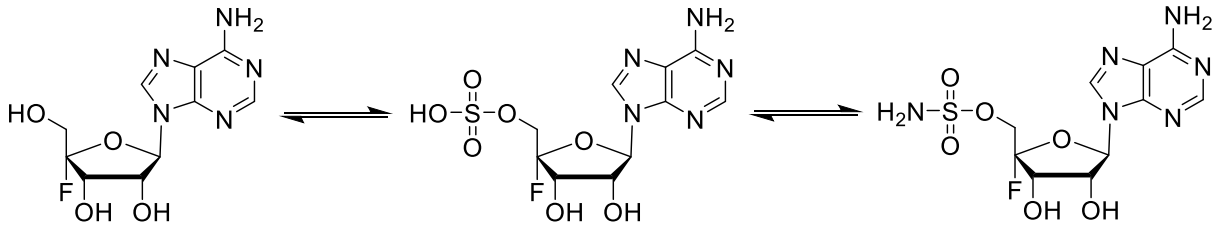


Scheme 19. Oxidation of protected aldehyde **173** and attempt to synthesise oxime **174**.

4.1.4. Synthesis of 4'-fluoroadenosine 171

4.1.4.1. Preparation of fluoroadenosine 171

4'-Fluoroadenosine **171** emerged as a possible advanced intermediate in the biosynthesis of nucleocidin **78a**, which would require sulfation, and then amination, as shown in Scheme 20.



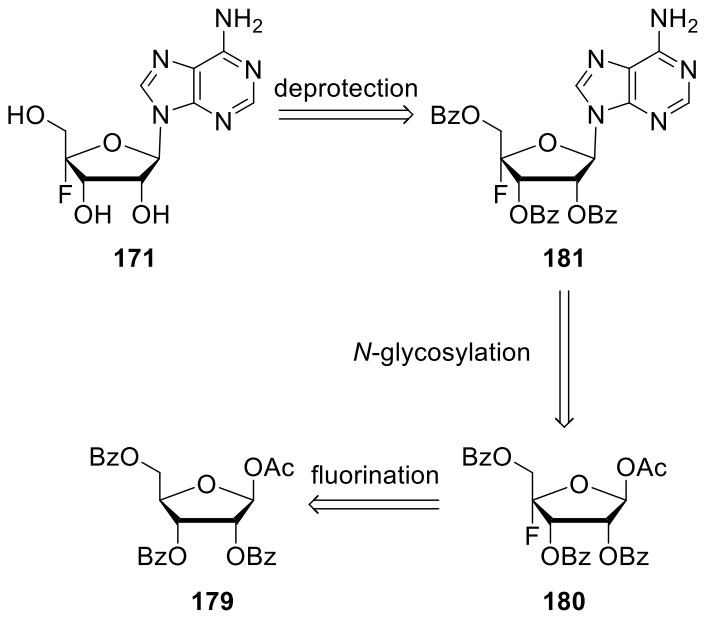
Scheme 20. Putative metabolic pathway of 4'-fluoroadenosine **171** to nucleocidin **78a**.

In order to be able to test this hypothesis, the synthesis of 4'-fluoroadenosine **171** was carried out. The synthesis of this fluorinated adenosine **171** has been described in the literature.[32-34] In the procedures developed to date, the protected ribose **179** was sequentially photobrominated *via N*-bromosuccinimide under UV irradiation and then fluorinated affording the fluorinated ribose **180**.

In 1964, the formation of nucleosides *via* the silyl-Hilbert-Johnson reaction, also known as the Vorbrüggen reaction, was described under mild conditions. In order to synthesise nucleosides by this method, a protected sugar is reacted with a silylated heterocycle in the presence of a Lewis acid such as trimethylsilyl triflate (TMSOTf).[35,36] However, silylated heterocycles are rather unstable due to hydrolysis and as a result are difficult to handle. Therefore, a one-step protocol was developed for the silylation of the heterocycle followed by nucleoside formation.[37,38]

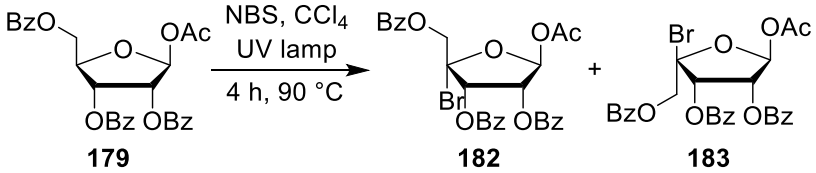
A modified Vorbrüggen reaction was devised using the fluorinated ribose **180**, adenine, *N*,*O*-bis(trimethylsilyl)trifluoro acetamide (BSTFA) and trimethylsilyl triflate, which gave the

protected fluoroadenosine **181**.[39] The deprotection of the protected fluoroadenosine **181** was then achieved using an excess of sodium carbonate.



Scheme 21. Retrosynthetic strategy to 4'-fluoroadenosine **171**.

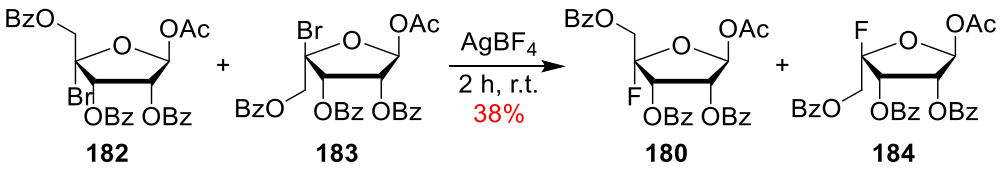
It was decided to pursue the method developed by Lee *et al*. for the synthesis of 4'-fluoroadenosine **171**.[33] As such, the commercially available ribose **179** was first treated with *N*-bromosuccinimide in carbon tetrachloride under UV irradiation affording the brominated intermediates **182** and **183**, as shown in Scheme 22.



Scheme 22. Photobromination of the protected ribose **179**.

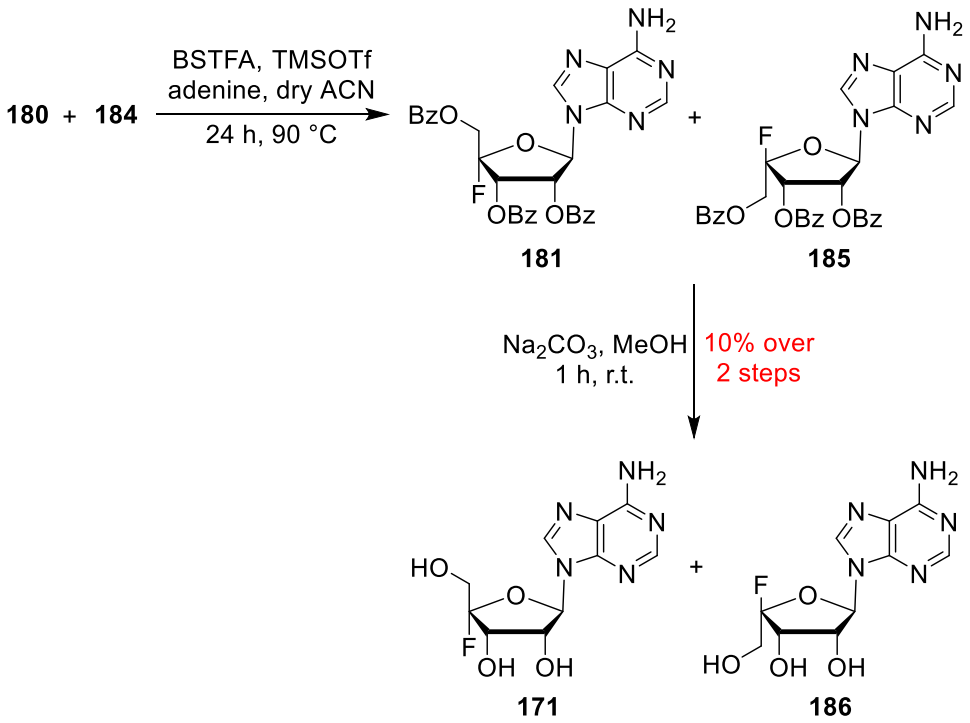
Due to the instability of the bromides **182** and **183**, this product mixture was immediately treated with a solution of $AgBF_4$, as illustrated in Scheme 23. Purification by column chromatography afforded an inseparable mixture of α-fluororibose **180** and β-fluororibose **184**

in a 1:1 ratio. Separation of these two products by HPLC with a normal phase column could not be performed because of their similar retention times.



Scheme 23. Fluorination of the bromide mixture of **182** and **183**.

The 4-fluororibose mixture of **180** and **184** underwent *N*-glycosylation giving a complex mixture which could not be purified by column chromatography, or by HPLC with a normal phase column or a reverse phase column. Therefore, the product mixture of **181** and **185** was immediately treated with sodium carbonate in methanol. Purification of the reaction mixture by HPLC afforded fluoroadenosines **171** and **186** as an inseparable mixture in a 1:1 ratio. A reverse phase C-18 column was used with a flow rate of 10 mL/min and an isocratic mobile phase of 95% buffer A (0.05% TFA/H_2O) and 5% buffer B (0.05% TFA/ACN). The inseparable mixture of **171** and **186** had a retention time of 8.4 minutes. The fraction containing the 4'-fluoroadenosines **171** and **186** was collected, concentrated and freeze-dried affording a sample of the two isomers, as shown in Figure 3. However, the yield over the two last steps was low (only 10%), as shown in Scheme 24, and it was found that this reaction could not be easily scaled up.



Scheme 24. Synthesis of the fluoroadenosine **171** and **186**.

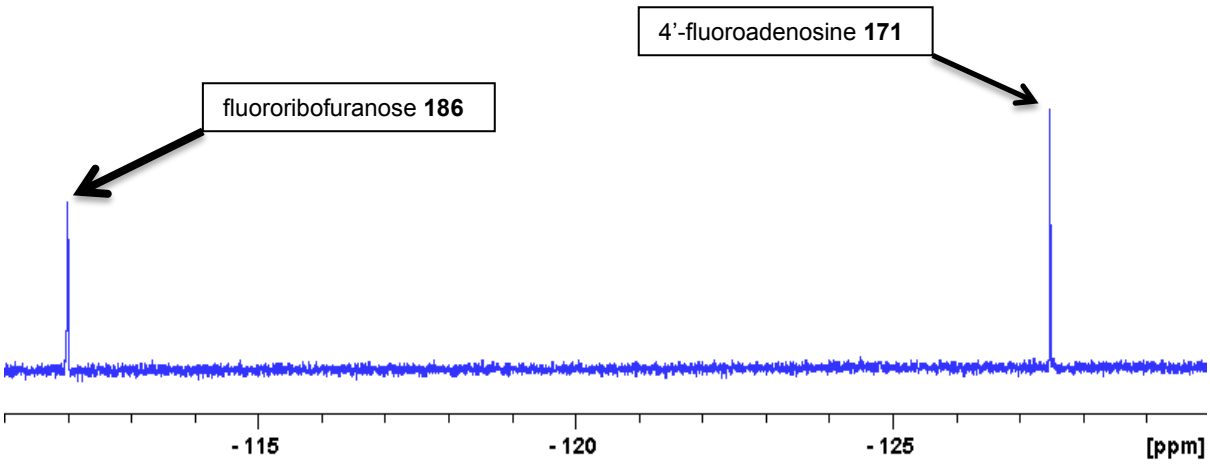
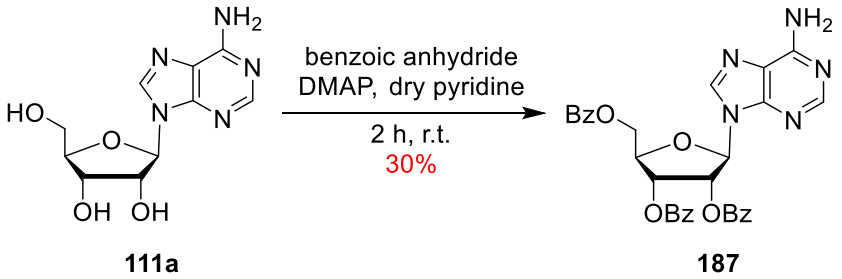


Figure 3. ^{19}F-NMR spectrum of the fluoroadenosine **171** and **186**.

4.1.4.2. Optimisation

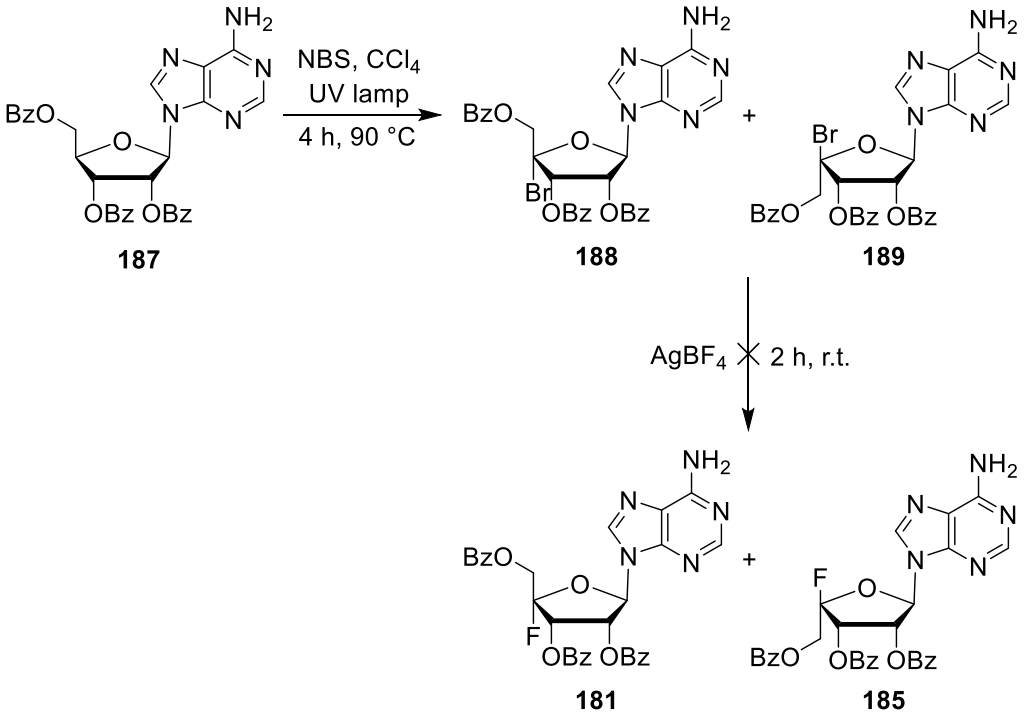
In order to optimise the overall yield of the synthesis of fluoroadenosine **171**, fluorination at the carbon C-4' was attempted of the protected adenosine **187** instead of at the protected ribose **179** stage. This new strategy does not require the *N*-glycosylation reaction, a step resulting in a low yield and difficult product purification. The synthesis of the tribenzoyladenosine **187** was achieved from adenosine **111a** using benzoic anhydride in dry pyridine. The product was purified in a 30% yield, as illustrated in Scheme 25.[40]



111a **187**

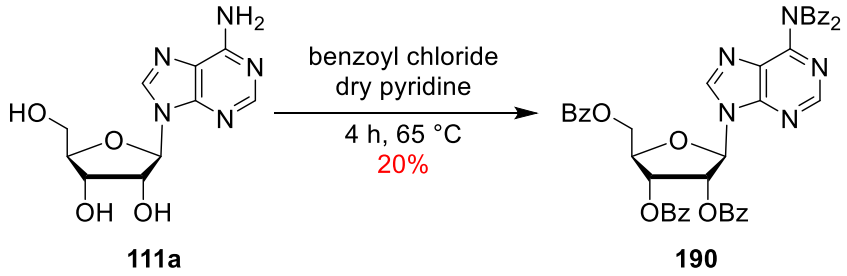
Scheme 25. Synthesis of the tribenzoyladenosine **187**.

Photobromination of the tribenzoyladenosine **187** resulted in an unpurified mixture of the bromides **188** and **189**, as confirmed by a mass spectrometer. Unfortunately, the fluorination reaction with a solution of $AgBF_4$ was unsuccessful, giving a complex mixture of nonfluorinated compounds, as shown in Scheme 26.



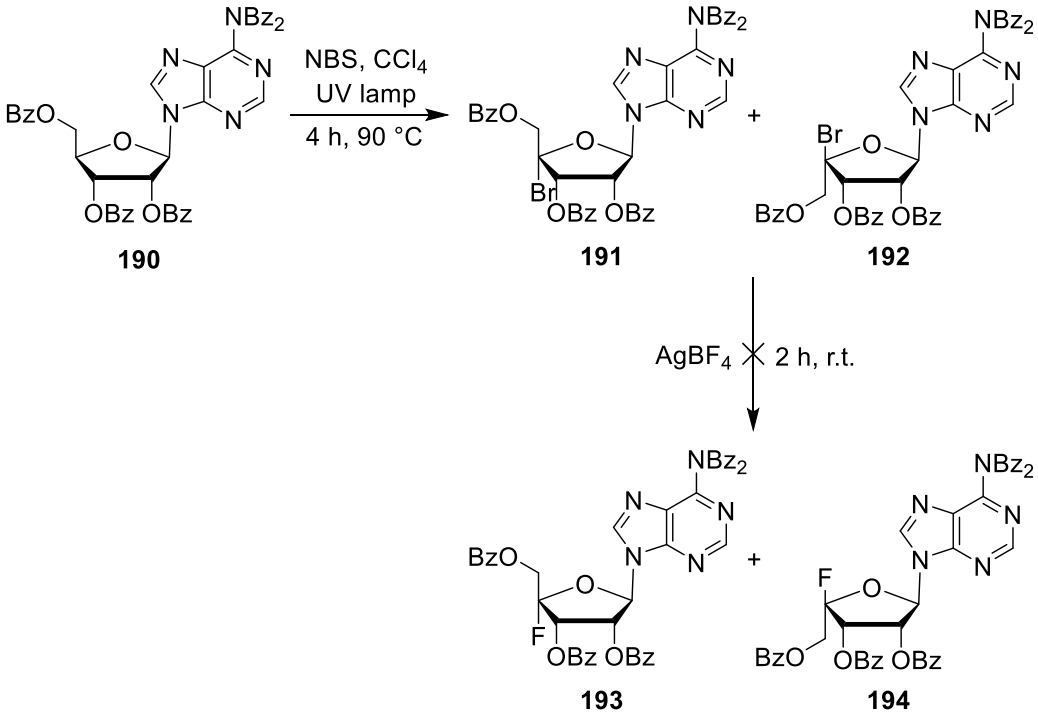
Scheme 26. Photobromination of the tribenzoyladenosine **187**.

In the light of this, an approach including the protection of adenosine **111a** was explored. Therefore, the synthesis of the pentabenzoyladenosine **190** was carried out using benzoyl chloride in dry pyridine, as illustrated in Scheme 27.[41]



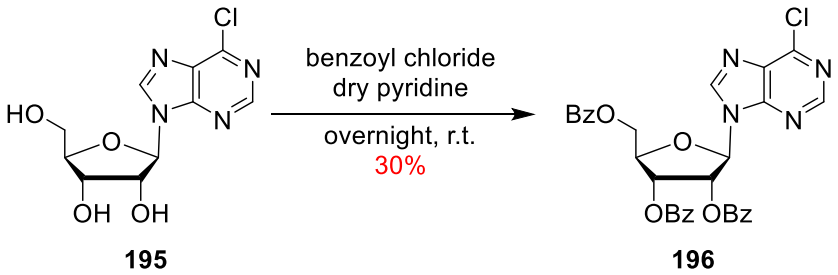
Scheme 27. Synthesis of the pentabenzoyladenosine **190**.

As with the tribenzoyladenosine **190** substrate, photobromination of the pentabenzoyladenosine **190** gave the unpurified bromide mixture **191** and **192**. Immediate fluorination of the bromide mixture **191** and **192** with $AgBF_4$ gave a complex mixture, but there were no organic fluorine signals in the ^{19}F-NMR.



Scheme 28. Photobromination of the pentabenzoyladenosine **190**.

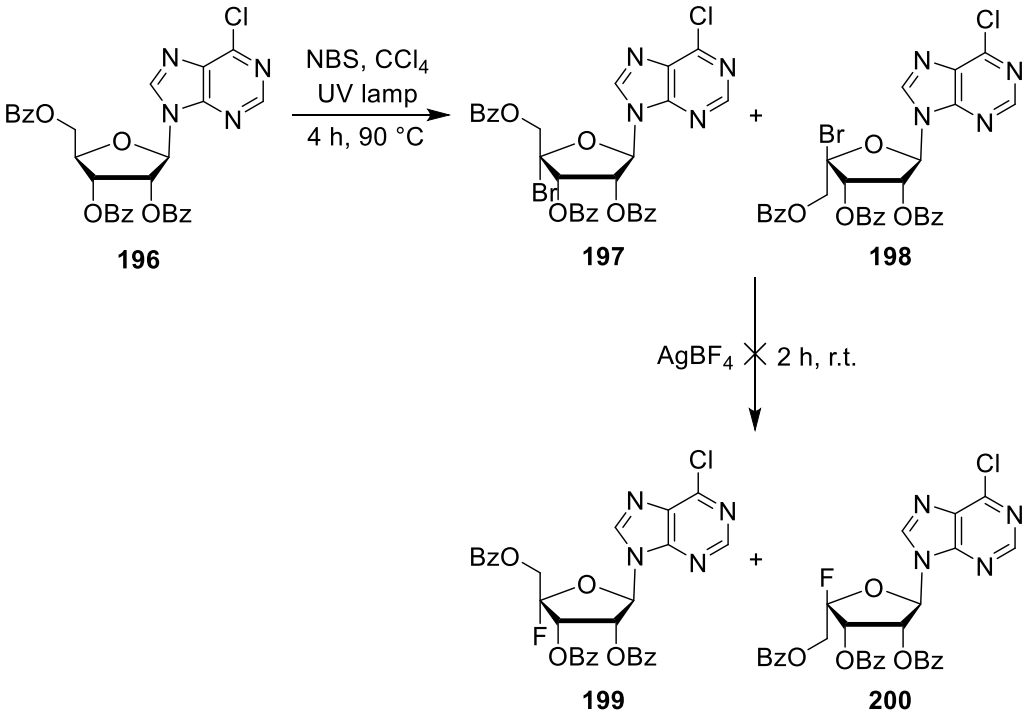
It was then decided to explore the synthesis of fluoroadenosine **171** *via* an alternative route involving chloropurine riboside **195**. This was chosen as it was anticipated that treatment with ammonia gas would remove the benzoyl groups and generate the adenine base at a later stage in the synthesis. Firstly, the synthesis of tribenzoylriboside **196** was carried out using chloropurine riboside **195** and benzoyl chloride in dry pyridine, as shown in Scheme 29.[42]



Scheme 29. Synthesis of the tribenzoylriboside **196**.

After photobromination of the tribenzoylriboside **196**, mass spectrometry indicated that bromoriboses **197** and **198** had formed. This mixture of **197** and **198** was treated with $AgBF_4$ and this gave a complex mixture with only traces of the anticipated fluororibosides **199** and

200 as judged by ^{19}F-NMR. To conclude, a synthesis of fluoroadenosine **171** was achieved as a 1:1 mixture with its isomer **186**. However, radical bromination followed by fluorination does not seem to function in the presence of the adenine base attached to the carbon C-1' of the ribose ring. Separation of fluorinated sugars **171** and **186** proved to be very challenging despite the use of reverse phase and normal phase columns.



Scheme 30. Photobromination of the tribenzoylriboside **196**.

4.2. Cell-free extract experiments

After the synthesis of several putative substrates for the fluorinating enzyme, it was necessary to carry out cell-free extract experiments with adenosine **111a**, adenosine sulfate **160**, sulfamoyladenosine **161** and 4'-fluoroadenosine **171**. In 2002, cell-free extract experiments were achieved with resting cells of *Streptomyces cattleya* and these experiments allowed the identification of SAM **27a** as the substrate for the fluorinase enzyme. Therefore, the same procedure as developed by O'Hagan *et al*. for the cell-free extract experiments of *Streptomyces cattleya* was followed for *Streptomyces calvus*.[43-45]

After 7 days of fermentation, cultures of *Streptomyces calvus* were centrifuged. The cells were washed with Tris–HCl buffer, disrupted by ultrasoniaction and the cell debris removed by centrifugation. The supernatant was then used as the cell free extract for incubation experiments with a mixture of substrates and co-factors at 37 °C. After incubation, the protein was precipitated by heating and then removed by centrifugation. The supernatant was freeze-dried and analysed by ^{19}F-NMR.

In all experiments, the buffer was supplemented with magnesium chloride which is used to stabilise ATP. ATP was added to the cell-free extracts as an energy source for the biosynthesis of nucleocidin **78a** and as potential source of a substrate. Cell-free extracts were also incubated with co-factors such as NAD^+ and $NADP^+$ to accommodate the redox-enzyme oxidations within the cell-free extract. When necessary, the buffer of the cell-free extract was supplemented with ammonia or hydrated magnesium sulfate in order to facilitate enzymatic sulfamoylation of adenosine **111a** or 4'-fluoroadenosine **171**. In the case of the cell-free extract experiments with adenosine sulfate **160**, only ammonia was added to facilitate the sulfamoyl group synthesis.

Two buffers were explored; Tris-HCl and phosphate buffer. Unfortunately, none of these experiments revealed the formation of nucleocidin **78a** or any fluorinated intermediates by ^{19}F-NMR or mass spectrometry.

	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Tris-HCl or phosphate buffer [20 mM, pH 7.0] + $MgCl_2$ [10 mM]	1260 µL	1290 µL	1320 µL	1260 µL
NAD^+ [100 mM]	30 µL	30 µL	30 µL	30 µL
$NADP^+$ [100 mM]	30 µL	30 µL	30 µL	30 µL
ATP [100 mM]	30 µL	30 µL	30 µL	30 µL
NH_3 [100 mM]	30 µL	30 µL		30 µL
$MgSO_4.7H_2O$ [100 mM]	30 µL			30 µL
KF [100 mM]	30 µL	30 µL	30 µL	30 µL
adenosine [50 mM]	60 µL			
adenosine sulfate [50 mM]		60 µL		
sulfamoyladenosine [50 mM]			60 µL	
4'-fluoroadenosine [50 mM]				60 µL

Figure 4. Incubation of cell-free extracts of *S. calvus*.

4.3. Conclusion

Despite the high number of sulfating reagents available, it proved to be more difficult than expected to synthesise adenosine sulfate **160**. In this study, adenosine **111a** underwent sulfation using chlorosulfonic acid **150** in dry pyridine and chloroform to afford adenosine sulfate **160**. Due to its water solubility and the high purity required for cell-free extract experiments, the desired product **160** was purified through a combination of ion exchange column and HPLC, sacrificing yield for purity.

The synthesis of the sulfamoyladenosine **161** was described in the literature by Lukkarila *et al*.[21] In place of the acetal protection of adenosine **111a**, the 2'-OH and 3'-OH of adenosine **111a** was TBS protected in order to increase the overall yield. This strategy proved to be successful and the yield of the final deprotection was essentially quantitative.

The synthesis of aldehyde **168** has been described in the literature by Liu *et al*.[26] and by Wnuk *et al*.[27] Unfortunately, due to its instability, the purification of carboxaldehyde **204** proved unsuccessful by HPLC and it could not be obtained for these studies.

4'-Fluoroadenosine **171** is a candidate advanced intermediate in the biosynthesis of nucleocidin **78a**. Therefore, the synthesis of fluoroadenosine **171** was achieved as described in the literature.[32-34] Purification by HPLC proved to be very difficult and gave an inseparable mixture of α and β fluoroadenosine **171** and **186** in a 1:1 ratio.

Cell-free extract experiments were carried out with the previous synthesised putative substrates **160**, **161**, **171** and adenosine **111a** following the procedures developed by O'Hagan.[43-45] Unfortunately, none of these cell-free extract experiments suggested the formation of nucleocidin **78a** or any fluorinated intermediates by ^{19}F-NMR and mass spectrometry. However, these results do not disprove any of these compounds as putative substrates since no control experiment could be performed on the integrity of the biosynthetic

enzymes in the cell-free extract. The proteins of *Streptomyces calvus*' cells may lose their activity during the preparation of the cell-free extract experiments.

4.4. References

1. A. Bartholomé, J. E. Janso, U. Reilly, D. O'Hagan, *Org. Biomol. Chem.*, 2017, **15**, 61-64.
2. A. Lanteri, *J. Am. Oil. Chem. Soc.*, 1978, **55**, 128-133.
3. E. A. Knaggs, M. J. Nepras, *Kirk-Othmer Encyclopedia of Chemical Technology*, 3rd Ed., Wiley Interscience , New-York.
4. R. Hobkirk, *Trends Endocrinol. Metab.*, 1993, **4**, 69-74.
5. M. J. Zamek-Gliszczynski, K. A. Hoffmaster, K. Nezasa, M. N. Tallman, K. L. Brouwer, *Eur. J. Pharm. Sci.*, 2006, **27**, 447-486.
6. C. N. Falany, *The FASEB Journal*, 1997, **11**, 206-216.
7. E. Chapman, M. D. Best, S. R. Hanson, C. H. Wong, *Angew Chem. Int. Ed. Engl.*, 2004, **43**, 3526-3548.
8. E. E. Gilbert, *Sulfonation and related reactions*, New York: Interscience publishers, Chapters 1 & 6.
9. N. C. Deno, M. S. Newman, *J. Am. Chem. Soc.*, 1950, **72**, 3852-3856.
10. A. W. Williamson, *Proc. R. Soc. London*, 1854, **7**, 11-15.
11. A. Williamson, *J. Chem. Soc.*, 1857, **10**, 97-101.
12. P. Sosis, L. J. Dringoli, *J. Am. Oil Chemists' Soc.*, 1970, **47**, 229-232.
13. R. O. Mumma, *Lipids*, 1966, **1**, 221-223.
14. R. O. Mumma, C. P. Hoiberg, *J. Chem. Eng. Data*, 1971, **16**, 492-494.
15. J. P. Dusza, J. P. Joseph, S. Bernstein, *Steroids*, 1985, **45**, 303-315.
16. V. Nair, S. Bernstein, *Orgn. Prepr. Proc. Int. Briefs*, 1987, **19**, 466-467.
17. K. William, *Method of providing ocular neuroprotection*, US2014/0275128 A1.
18. G. Huber, *Eur. J. Inorg. Chem.*, 1956, **89**, 2853-2862.
19. F. Egami, N. Takahashi, *Bull. Chem. Soc. Jpn.*, 1955, **28**, 666-668.
20. C. Zhao, J. Qi, W. Tao, L. He, W. Xu, J. Chan, Z. Deng, *PLoS ONE*, 2014, **9**, 1-15.

21. J. L. Lukkarila, S. R. da Silva, M. Ali, V. M. Shahani, G. Wei Xu, J. Berman, A. Roughton, S. Dhe-Paganon, A. D. Schimmer, P. T. Gunning, *ACS Med. Chem. Lett.*, 2011, **2**, 577-582.

22. S. Lun, H. Guo, J. Adamson, J. S. Cisar, T. D. Davis, S. S. Chavadi, J. D. Warren, L. E. N. Quadri, D. S. Tan, W. R. Bishal, *Antimicrob. Agents Chemother.*, 2013, **57**, 5138-5140.

23. L. B. Townsend, R. S. Tipson, *John Wiley & Sons*, New-York, 1978, 765-769.

24. P. Van de Vijver, T. Ostrowski, B. Sproat, J. Goebels, U. Rutgeerts, A. Van Aerschot, M. Waer, P. Herdewijn, *J. Med. Chem.*, 2008, **51**, 3020-3029.

25. D. S. Tan, L. E. Quadri, J-S. Ryu, J. S. Cisar, J. A. Ferreras, X. Lu, *Anti-microbial agents and uses thereof*, WO 2006/113615 A2.

26. S. Liu, S. F. Wnuk, C. Yuan, M. J. Robbins, R. T. Borchardt, *J. Med. Chem.*, 1993, **36**, 883-887.

27. S. F. Wnuk, C-S. Yuan, R. T. Borchardt, J. Balzarini, E. De Clercq, M. J. Robins, *J. Med. Chem.*, 1997, **40**, 1608-1618.

28. J. Kalia, R. T. Raines, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 7523-7526.

29. K. E. Pfitzner, J. G. Moffatt, *J. Am. Chem. Soc.*, 1963, **85**, 3027-3028.

30. T. T. Tidwell, *Org. React.*, 1990, **39**, 297-572.

31. T. V. Lee, *Comp. Org. Syn.*, 1991, **7**, 291-303.

32. R. J. Ferrier, S. R. Haines, *J. Chem. Soc., Perkin Trans. 1*, 1984, 1675-1681.

33. S. Lee, C. Uttamapinant, G. L. Verdine, *Org. Lett.*, 2007, **9**, 5007-5009.

34. G. L. Verdine, S. Lee, *Synthesis of nucleosides*, WO 2009/058800 A2.

35. E. Z. Wittenburg, *Chem.*, 1964, **4**, 303-304.

36. W-B. Choi, L. J. Wilson, S. Yeola, D. C. Liotta, R. F. Schinazi, *J. Am. Chem. Soc.*, 1991, **113**, 9377-9379.

37. H. Vorbrüggen, B. Bennua, *Tetrahedron Lett.*, 1978, **19**, 1339-1342.

38. H. Vorbrüggen, B. Bennua, *Chem. Ber.*, 1981, **114**, 1279-1286.

39. G. E. Wright, L. W. Dudycz, *J. Med. Chem.*, 1984, **27**, 175-181.

40. T. Wada, T. Moriguchi, M. Sekine, *J. Am. Chem. Soc.*, 1994, **116**, 9901-9911.

41. G. A. Brown, E. D. Savory, J. V. A. Ouzman, A. M. Stoddart, *Improved synthesis of 2-substituted adenosines*, WO 2005/056571 A1.

42. S. Kozai, S. Takamatsu, K. Izawa, T. Maruyama, *Tetrahedron Lett.*, 1999, **40**, 4355-4358.

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

44. C. Schaffrath, *Ph.D. Thesis*, University of St Andrews, 2002.

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

5. Thesis conclusions

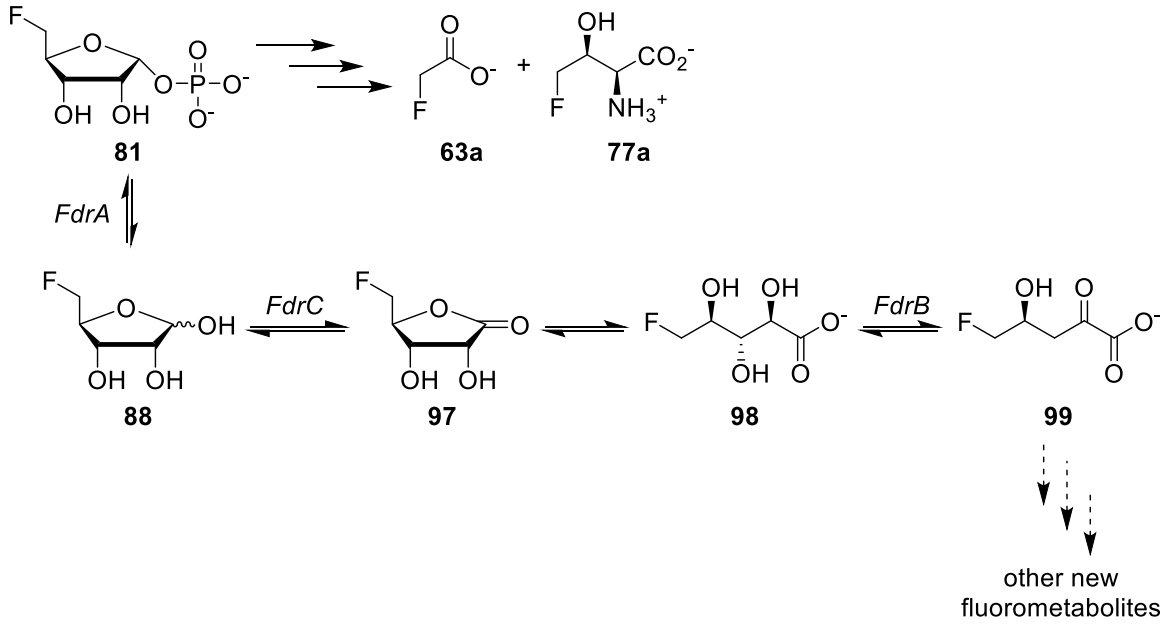
5.1. New fluorinated natural product from *Streptomyces* sp. MA37

In 2014, four different bacteria were each found to possess a gene with high homology to the gene encoding the fluorinase enzyme in Streptomyces cattleya.[1,2] Among these four bacteria, a soil bacterium named *Streptomyces* sp. MA37 was isolated. Fermentations of *S.* sp. MA37 in the presence of inorganic fluoride showed the production of fluoroacetate **63a** and fluorothreonine **77a**, two fluorometabolites also observed in *Streptomyces cattleya*.

Genome studies proved that both soil bacteria share the same metabolic pathway for the biosynthesis fluoroacetate **63a** and fluorothreonine **77a**.[1] Production of seven new unidentified fluorometabolites was also observed, each in lower concentrations than both fluoroacetate **63a** and fluorothreonine **77a**. Moreover, when analysed by ^{19}F-NMR, each of these new fluorinated natural products share the same multiplicity with different coupling constants to fluorothreonine **77a**.

Therefore, it was suggested that these new fluorometabolites have the same metabolic origin as fluoroacetate **63a** and 4-fluoro-L-threonine **77a** but within this biosynthetic pathway there exists a branch point creating two competing metabolic pathways. Genome studies of *S.* sp. MA37 showed that several genes share high homology with genes responsible for the biosynthesis of chlorosalinosporamide A **11** in *Salinospora tropica*.[3]

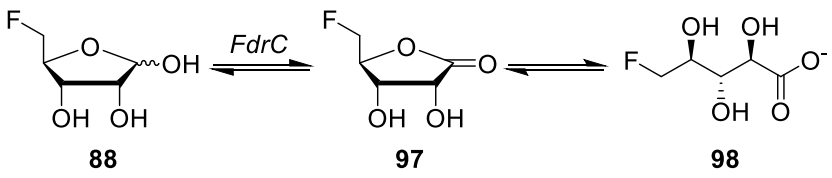
In accordance to the metabolic pathway of salinosporamide A **11**, a related hypothetical metabolic pathway was proposed to explain the production of these new fluorometabolites, as depicted in Scheme 1.



Scheme 1. Proposed metabolic pathway of new fluorometabolites in *S.* sp. MA37.

Over-expression of the gene *Fdr C*, encoding for a chain dehydrogenase was carried out in *E. coli* by Dr. Ma. In order to perform the enzyme assays, the synthesis of the substrate FDR **88** was conducted following the procedure developed by Li *et al*.[4] The absolute stereochemistry of the enzyme product was confirmed by total synthesis of FHPA **98**. GC-MS of the enzyme product and the synthetic **98** showed that they both shared the same retention time and fragmentation pattern, proving that they are the same compound.[5]

An additional spiking experiment confirmed that FHPA **98** is a fluorinated natural product produced by *Streptomyces* sp. MA37. Importantly, it is the first novel fluorometabolite to be identified in over two decades. However, further investigation is required for the identification of the other fluorinated natural products.



Scheme 2. Enzymatic synthesis of FHPA **98** as a new fluorometabolite.

5.2. Metabolism studies of nucleocidin 78a in *Streptomyces calvus*

Nucleocidin **78a** was the first fluorinated natural product isolated from the soil bacterium *Streptomyces calvus*. Since its initial isolation in 1957, the production of nucleocidin **78a** has been lost in publically available strains.[6-9] Due to its particular structure, with the fluorine atom attached at the carbon C-4' of the ribose ring, the originally isolated fluorinase enzyme cannot reasonably be involved in the fluorination of nucleocidin **78a**. Therefore, the discovery and isolation of the second native fluorinating enzyme became the focus of this study.

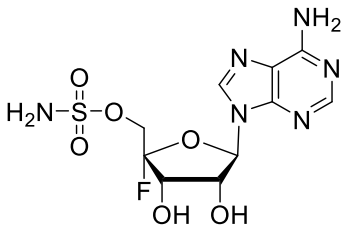
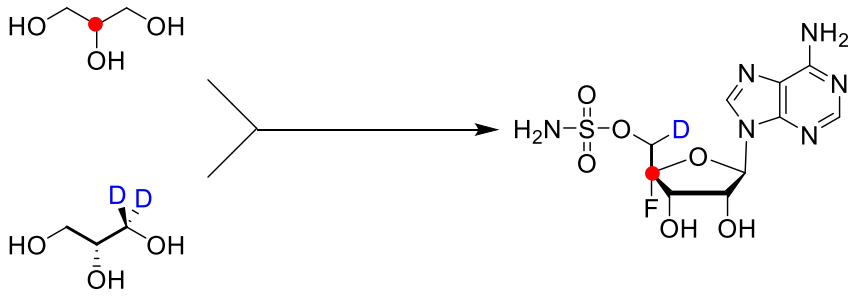


Figure 1. Structure of nucleocidin **78a**.

A strain of *Streptomyces calvus* T-3018 which was known to produce nucleocidin **78a** in very low yield was provided by the Pfizer company. The presence of nucleocidin **78a** could be observed by ^{19}F-NMR after a *n*-butanol extraction of the liquid medium. Identification of nucleocidin **78a** was carried by both LC-MS and spiking experiment with a pure sample provided by Pfizer.

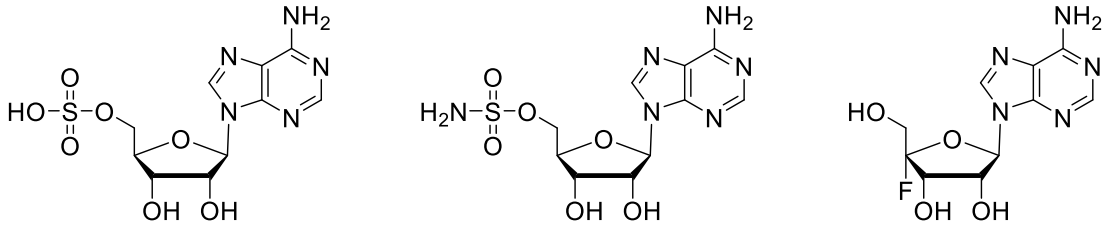
Due to its structural similarity, adenosine **111a** was thought to be an advanced intermediate in the biosynthesis of nucleocidin **78a**. Therefore, the biosynthesis of adenosine **111a** from glycerol *via* the pentose phosphate pathway **116a** was investigated. Pulse feeding experiments with [2-^{13}C]-glycerol **116b** proved that the carbon C-4' of nucleocidin **78a** is derived from the carbon C-2 of glycerol **116a**. Incubation of [1,1,2,3,3-$^{2}H_5$]-Glycerol **116c** and (2*R*)-[1-$^{2}H_2$]-glycerol **116e** in separate *Streptomyces calvus* fermentations showed incorporation with a loss of a deuterium atom at the carbon C-5' of nucleocidin **78a**. This loss suggested that the fluorination mechanism

might be more complex than initially expected. A summary of the results of the feeding experiments performed is described in Scheme 3.[10]



Scheme 3. Incorporation of labelled glycerols **116b** and **116c** into nucleocidin **78a**.[10]

Concurrently, the synthesis of putative intermediates in the metabolic pathway was achieved, as shown in Figure 2. Adenosine sulfate **160** and sulfamoyladenosine **161** were synthesised due to their high structural similarities with nucleocidin **78a**. The synthesis of fluoroadenosine **171** was also performed, because it was thought to be an advanced fluorinated intermediate in the biosynthesis of nucleocidin **78a**. Unfortunately, production of nucleocidin **78a** or new fluorinated intermediates could not be observed during any cell-free extract experiments.



160 **161** **171**

Figure 2. Structure of putative intermediates for cell-free extract experiments.

Recent genome studies of a gene cluster which is thought to be involved in the biosynthesis of nucleocidin **78a** revealed the presence of a radical-SAM enzyme.[8,9] These radical-SAM enzymes are known to be sensitive to oxygen, which could explain the lack of nucleocidin production in the previous cell-free extract experiments. Knock-out experiments and over-expression of the gene encoding for this radical-SAM enzyme is currently in progress.

5.3. References

1. H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyeremeh, Y. Yu, T. Shepherd, J. H. Naismith, D. O'Hagan, *ChemBioChem*, 2014, **15**, 364-368.
2. S. Huang, L. Ma, M. H. Tong, Y. Yu, D. O'Hagan, H. Deng, *Org. Biomol. Chem.*, 2014, **12**, 4828-4831.
3. R. H. Feling, G. O. Buchanan, T. J. Mincer, C. A. Kauffman, P. R. Jensen, W. Fenical, *Angew. Chem. Int. Ed.*, 2003, **42**, 355-357.
4. X-G. Li, S. Dall'Angelo, L. F. Schweiger, M. Zanda, D. O'Hagan, *Chem. Commun.*, 2012, **48**, 5247-5249.
5. L. Ma, A. Bartholome, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan, *Chem. Sci.*, 2015, **6**, 1414-1419.
6. S. O. Thomas, V. L. Singleton, J. A. Lowery, R. W. Sharpe, L. M. Pruess, J. N. Porter, J. H. Mowat, N. Bohonos, *Antibiotics Annual*, 1957, 716-721.
7. K. Fukuda, T. Tamura, Y. Segawa, Y. Mutaguchi, K. Inagaki, *Actinomycetologica,* 2009, **23**, 51-55.
8. L. Kalan, A. Gessner, M. N. Thaker, N. Waglechner, X. Zhu, A. Szawiola, A. Bechthold, G. D. Wright, D. L. Zechel, *Chemistry & Biology*, 2013, **20**, 1214-1224.
9. X. M. Zhu, S. Hackl, M. N. Thaker, L. Kalan, C. Weber, D. S. Urgast, E. M. Krupp, A. Brewer, S. Vanner, A. Szawiola, G. Yim, J. Feldmann, A. Bechtold, G. D. Wright, D. L. Zechel, *ChemBioChem*, 2015, **16**, 2498-2506.
10. A. Bartholomé, J. E. Janso, U. Reilly, D. O'Hagan, *Org. Biomol. Chem.*, 2017, **15**, 61-64.

6. Experimental

6.1. General experimental

Air and moisture sensitive reactions were carried out under an argon atmosphere in flame-dried glassware. Room temperature refers to 18-25 °C. All evaporations and concentrations were performed under reduced pressure (*in vacuo*). All reagents (from Acros UK, Alfa Aesar UK, Fischer UK, Fluorochem UK or Sigma Aldrich UK) were of synthetic grade and were used without further purification. Anhydrous solvents (Et_2O, THF) were obtained from MBraun MB SPS-800 solvent purification system by passage through two drying columns and dispensed under an argon atmosphere. Anhydrous MeOH and ACN were distilled from calcium hydride in a recycling still. Anhydrous pyridine was distilled from potassium hydroxide using standard protocols.[1]

The course of reactions was monitored by thin-layer chromatography (TLC) using aluminium plates coated with silica gel ($60F_{245}$ Merck). TLC plates were observed under UV light (254 nm and 266 nm) before being stained with anisaldehyde-sulfuric acid or alkaline potassium permanganate and developed by heating. Column chromatography was performed on Merck Geduran silica gel 60 (250-400 mesh) under a positive pressure of compressed air eluting with solvents (reported as *v/v*) as supplied. Reverse phase column chromatography was performed using Extract Clean C_{18}-HC prepacked cartridges.

NMR spectra were recorded at 298 K on Bruker Avance 300, Avance II 400, Avance 500 or Avance III HD 500 instruments. ^{1}H and ^{13}C-NMR spectra were recorded using deuterated solvent as the lock and the residual solvent as the internal standard. ^{19}F-NMR spectra were referenced to $CFCl_3$ as an external standard. Chemical shifts are reported in parts per million (ppm) and coupling constants (*J*) are reported in Hertz (Hz). The abbreviations for the multiplicity of the proton, carbon and fluorine signals are as follows: s singlet, d doublet, dd

doublet of doublets, ddd doublet of doublet of doublets, dddd doublet of doublet of doublet of doublets, dt doublet of triplets, q quartet, m multiplet, br s broad singlet. Compounds are numbered according to customary purine numbering. When necessary, resonances were assigned using two-dimensional experiments (COSY, HSQC, HMBC, TOCSY).

Optical rotations were measured with a Perkin Elmer 341 polarimeter in a 10.0 cm cell at the wavelength of the sodium D-line (λ = 589 nm). Specific rotations are reported in implied units of 10^{-1} deg cm^2 g^{-1} and concentrations (c) are reported in g/100 mL. High resolution electrospray ionisation mass spectra were obtained on a Micromass LCT or ThermoFisher Excalibur Orbitrap spectrometers operating in positive or negative mode, from solutions in $CHCl_3$, MeOH, ACN or H_2O by the Mass Spectrometry Service at the University of St Andrews. MALDI MS was acquired using a 4800 MALDI TOF/TOF Analyser (ABSciex) equipped with a Nd:YAG 355 nm laser and calibrated using a mixture of peptides by the Mass Spectrometry Service at the University of St Andrews.

Infrared spectra were acquired on a Perkin Elmer Spectrum GX FTIR apparatus spectrometer either as KBr pellets, or as a thin film between NaCl plates, or on a Shimadzu IRAffinity-1S spectrometer with a diamond ATR attachment. Absorption maxima are reported in units of wavenumbers (cm^{-1}). Melting points were recorded on an Electrothermal IA9100 melting point apparatus, or on a Griffin MPA 350.BM2.5 melting point apparatus and are uncorrected.

X-ray diffraction data were collected on a Mercury 70 diffractometer using graphite monochromated Mo-Kα radiation at 100 K by means of phi and omega scans. Structures were solved by direct methods and refined by full-matrix least squares techniques.

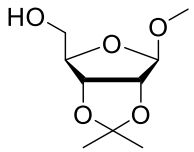
HPLC analyses/semi-preparations were performed using either a Varian Prostar (Varian 400 autosampler, Varian Prostar 230 solvent delivery system, Varian Prostar 235 UV-vis detector) or a Shimadzu Prominence (SIL-20A HT autosampler, CL-20AT ternary pump, DGU-20A3R solvent degasser, SPD 20A UV detector and CBM-20A controller module) with reverse phase

column. All solvents used were HPLC grade and were degassed prior to use by bubbling nitrogen through the solvents.

LC-MS analysis was performed on a Waters 2795 HPLC coupled in parallel to a Waters 2996 photodiode array detector and Micromass LCT TOF mass spectrometer in ESI positive mode using the column indicated in the individual experiment. Samples were freeze dried from frozen solutions in H_2O in a Christ Alpha 1-2 LO Plus freeze drier.

6.2. Compounds preparation

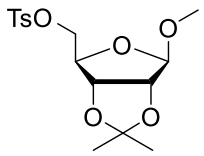
6.2.1. Methyl 2,3-*O*-isopropyliden-β-D-ribofuranoside 102



2,2-Dimethyloxypropane (20 mL, 16.3 mmol, 0.24 eq.) and perchloric acid 70% (4 mL, 66.6 mmol, 1 eq.) was added dropwise to a suspension of D-ribose **103** (10 g, 66.6 mmol, 1 eq.) in acetone (80 mL) cooled to 0 °C. The resulting reaction mixture was stirred for 2 h at room temperature and then MeOH (14 mL) was added and the mixture was stirred at room temperature for 16 h. A saturated solution of K_2CO_3 (10 mL) was added and the mixture was partitioned between H_2O (150 mL) and DCM (70 mL). The aqueous phase was extracted with DCM (3 × 70 mL). The combined organic phases were dried (Na_2SO_4), filtered and the filtrate was concentrated. The crude mixture was purified by column chromatography (PE/EtOAc 9:1) to afford methyl 2,3-*O*-isopropyliden-β-D-ribofuranoside **102** (5.9 g, 44%) as a yellow oil.

***R*$_f$** = 0.46 (EtOAc/PE 1:1); **[α]$^{20}_D$** : −73.5° (*c* 1.0, $CHCl_3$) (lit.,[2] −78.5°, *c* 1.0, $CHCl_3$); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.94 (1H, s, H-1), 4.80 (1H, d, *J* = 6.0 Hz, H-3), 4.55 (1H, d, *J* = 6.0 Hz, H-2), 4.40 (1H, t, *J* = 3.1 Hz, H-4), 3.68-3.56 (2H, m, CH_2-5), 3.40 (3H, s, OCH_3), 1.45 (3H, s, CH_3), 1.29 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 112.2 ($\underline{C}(CH_3)_2$), 110.0 (CH, C-1), 88.4 (CH, C-4), 85.9 (CH, C-2), 81.5 (CH, C-3), 64.0 (CH_2, C-5), 55.6 (CH_3, $O\underline{C}H_3$), 26.4 (CH_3), 24.7 (CH_3); **IR** ν_{max} (neat): 3458, 2988, 2932, 2836, 1457, 1380, 1275, 1207, 1163, 1097, 1044, 964, 870 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_9H_{16}O_5Na^+$ = 227.0895) found 227.0899 $[M+Na]^+$. Data are in agreement with literature.[2]

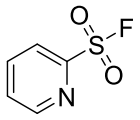
6.2.2. Methyl 2,3-*O*-(1-methylethylidene)-5-*O*-(*p*-toluenesulfonyl)-β-ribofuranoside 101



4-Toluenesulfonyl chloride (4.77 g, 25.0 mmol, 1.5 eq.) was added portionwise to a solution of methyl 2,3-*O*-isopropyliden-β-D-ribofuranoside **102** in dry pyridine (10 mL) at 0 °C and the reaction mixture was stirred for 3 h at room temperature. H_2O (4 mL) was carefully added to the reaction mixture and the stirring was continued for further 30 min. The mixture was extracted with $CHCl_3$ (3 × 15 mL) and the combined organic phases were washed with an aqueous solution of $CuSO_4$ (30 mL), saturated bicarbonate solution (30 mL), H_2O (30 mL) and dried (Na_2SO_4). The filtrate was concentrated *in vacuo* and the crude mixture was purified by column chromatography (PE/EtOAc 2:1) to afford methyl 2,3-*O*-(1-methylethylidene)-5-*O*-(*p*-toluenesulfonyl)-β-ribofuranoside **101** (4.92 g, 83%) as a colourless oil.

***R*$_f$** = 0.67 (EtOAc/PE 1:1); **[α]$^{20}_D$** : −49.5° (*c* 1.0, $CHCl_3$) (lit.,[2] −48.7°, *c* 1.0, $CHCl_3$); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 7.80 (2H, d, *J* = 8.3 Hz, ArH), 7.35 (2H, d, *J* = 8.3 Hz, ArH), 4.92 (1H, s, H-1), 4.58 (1H, d, *J* = 5.9 Hz, H-3), 4.52 (1H, d, *J* = 5.9 Hz, H-2), 4.30 (1H, t, *J* = 7.1 Hz, H-4), 4.04-3.97 (2H, m, CH_2-5), 3.22 (3H, s, OCH_3), 2.45 (3H, s, Ar-CH_3), 1.44 (3H, s, CH_3), 1.28 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 145.2 (Ar-C), 132.9 (Ar-C), 130.1 (Ar-CH), 128.1 (Ar-CH), 112.8 ($\underline{C}(CH_3)_2$), 109.6 (CH, C-1), 85.0 (CH, C-2), 83.7 (CH, C-4), 81.5 (CH, C-3), 69.3 (CH_2, C-5), 55.2 (CH_3, O$\underline{C}H_3$), 26.5 (CH_3), 25.0 (CH_3), 21.8 (CH_3, Ar-CH_3); **IR** ν_{max} (neat): 2952, 1596, 1450, 1375, 869 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{16}H_{22}O_7NaS^+$ = 381.0984) found 381.0980 $[M+Na]^+$. Data are in agreement with literature.[3]

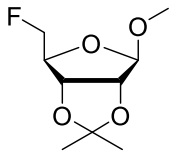
6.2.3. 2-Pyridinesulfonyl fluoride 104 (PyFluor)



A multi-neck flask fitted with an addition funnel and a thermometer was charged with sulfuric acid **149** (80 mL) and 2-mercaptopyridine (5.55 g, 50 mmol, 1 eq.) and left open to the atmosphere. The reaction mixture was cooled to 0 °C and 13% aqueous sodium hypochlorite (238 mL) was added dropwise over 4 h while stirring vigorously and maintaining an internal reaction temperature below 10 °C. Upon complete addition, the reaction mixture was extracted twice with ethyl acetate (2 × 125 mL). The combined organic extracts were concentrated to afford crude 2-pyridinesulfonyl chloride, which was immediately added to a mixture containing 1:3 (v/v) ACN:H_2O (80 mL) and potassium bifluoride (19.55 g, 250 mmol, 5 eq.). This solution was stirred vigorously for 20 min and was then extracted twice with EtOAc (75 mL). The organic extracts were dried over sodium sulfate and filtered through a 5 g silica plug, eluting with EtOAc. The filtrate was purified by column chromatography (PE/EtOAc 2:1) to afford 2-pyridinesulfonyl fluoride **104** (5.0 g, 62%) as a colourless crystalline solid.

***R*$_f$** = 0.44 (PE/EtOAc 2:1); **m.p.** 27-30 °C (lit.,[4] 23-26 °C); **^{1}H-NMR** ($CHCl_3$, 300 MHz) δ_H 8.85 (1H, d, *J* = 4.9 Hz, ArH), 8.15 (1H, d, *J* = 8.0 Hz, ArH), 8.08 (1H, tt, *J* = 7.8 Hz, *J* = 1.6 Hz, ArH), 7.70 (1H, ddd, *J* = 7.7 Hz, *J* = 4.8 Hz, *J* = 1.0 Hz, ArH); **^{13}C-NMR** ($CHCl_3$, 100 MHz) δ_C 151.9 (Ar-C, d, *J* = 30.5 Hz), 151.5 (Ar-CH, d, *J* = 1.1 Hz), 139.1 (Ar-CH), 130.0 (Ar-CH), 124.1 (Ar-CH, d, *J* = 2.2 Hz); **^{19}F-NMR** ($CHCl_3$, 376 MHz) δ_F -145.7; **HRMS** m/z (ES^+) (calculated $C_5H_4FNNaO_2S^+$ = 183.9839) found 183.9840 $[M+Na]^+$. Data are in agreement with literature.[4]

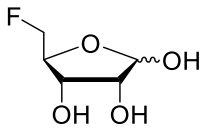
6.2.4. Methyl 5-deoxy-5-fluoro-2,3-*O*-isopropylidene-β-D-ribofuranoside 100



PyFluor **104** (0.81 g, 5.0 mmol, 1.1 eq.) and MTBD (1.3 mL, 9.0 mmol, 2 eq.) was added to a solution of 2,3-*O*-isopropyliden-β-D-ribofuranoside **102** (0.92 g, 4.5 mmol, 1 eq.) in toluene (12 mL). The reaction mixture was stirred at room temperature overnight then quenched by addition of a saturated aqueous solution of bicarbonate (15 mL) and the mixture was stirred for an additional 30 min. The organic phase was isolated and the aqueous layer was extracted with EtOAc (3 × 15 mL). The combined organic phases were dried ($MgSO_4$) and the solvent was evaporated. The crude mixture was then purified by column chromatography (PE/Et_2O 1:1) to afford methyl 5-deoxy-5-fluoro-2,3-*O*-isopropylidene-*β*-D-ribofuranoside **100** (0.76 g, 82%) as a colorless oil.

***R*$_f$** = 0.61 (Et_2O/PE 1:1); **[α]$^{20}_D$** : −90.0° (*c* 1.0, $CHCl_3$) (lit.,[5] −91.9°, *c* 1.0, $CHCl_3$); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.98 (1H, d, *J* = 2.4 Hz, H-1), 4.69 (1H, d, *J* = 6.0 Hz, H-3), 4.59 (1H, d, *J* = 6.0 Hz, H-2), 4.49-4.24 (3H, m, H-4 & CH_2), 3.32 (3H, s, OCH_3), 1.48 (3H, s, CH_3), 1.32 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 112.8 ($\underline{C}(CH_3)_2$), 109.4 (CH, C-1), 85.2 (CH, C-2), 84.6 (CH, d, *J* = 22.2 Hz, C-4), 83.1 (CH, d, *J* = 172.4 Hz, C-5), 81.2 (CH, d, *J* = 4.3 Hz, C-3), 55.0 (CH_3, $O\underline{C}H_3$), 26.5 (CH_3), 25.0 (CH_3); **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -225.2 – -225.7 (1F, m); **IR** ν_{max} (neat): 2940, 2837, 1457, 1384, 1212, 1088, 871 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_9H_{15}FNaO_4^+$ = 229.0847) found 229.0847 $[M+Na]^+$. Data are in agreement with literature.[3]

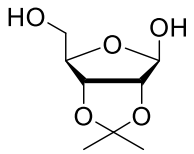
6.2.5. 5-Fluoro-5-deoxy-ribose 88 (FDR)



A solution of methyl 5-deoxy-5-fluoro-2,3-*O*-isopropylidene-*β*-D-ribofuranoside **100** (556 mg, 2.69 mmol, 1 eq.) in aqueous H_2SO_4 **149** (20 mM, 6 mL) was stirred at 98 °C for 2 h. The reaction mixture was cooled to room temperature and the pH was adjusted to pH 7 by addition of $BaCO_3$. The suspension was centrifuged and the supernatant was passed through a syringe filter (0.22 µm). The resultant solution was lyophilised for 16 h, affording 5-fluoro-5-deoxy-ribose **88** (360 mg, 88%) as a yellow oil.

^{1}H-NMR (D_2O, 300 MHz) δ_H 5.37 (0.3H, d, *J* = 4.0 Hz, α anomer, H-1), 5.23 (0.7H, d, *J* = 1.5 Hz, β anomer, H-1), other signals could not be interpreted due to overlapping; **^{13}C-NMR** (D_2O, 100 MHz) δ_C 101.5 (CH, β anomer, C-1), 96.8 (CH, α anomer, C-1), 83.7 (CH_2, d, *J* = 168.6 Hz, β anomer, C-5), 83.2 (CH_2, d, *J* = 167.2 Hz, α anomer, C-5), 81.3 (CH, d, *J* = 17.4 Hz, α anomer, C-4), 80.9 (CH, d, *J* = 18.2 Hz, β anomer, C-4), 75.3 (CH, β anomer, C-2), 71.0 (CH, α anomer, C-2), 69.8 (CH, d, *J* = 6.9 Hz, β anomer, C-3), 69.6 (CH, d, *J* = 6.5 Hz, α anomer, C-3); **^{19}F-NMR** (D_2O, 376 MHz) δ_F -229.4 (0.7F, dt, *J* = 47.3 Hz, 25.7 Hz, β anomer), -231.7 (0.3F, dt, *J* = 47.3 Hz, 25.7 Hz, α anomer); **HRMS** m/z (ES^-) (calculated $C_5H_8FO_4^-$ = 151.0412) found 151.0407 $[M-H]^-$. Data are in agreement with literature.[3]

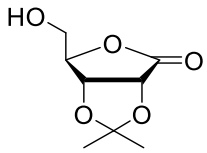
6.2.6. 2,3-*O*-Isopropylidene-D-ribose 110



2,2-Dimethoxypropane (2.5 mL, 20.6 mmol, 1.03 eq.) and 4-toluene sulfonic acid (0.19 g, 1.0 mmol, 0.05 eq.) were added to a cooled suspension of D-ribose **103** (3.0 g, 20.0 mmol, 1 eq.) in acetone (30 mL). After stirring at room temperature for 1 h, the clear mixture was neutralised with saturated aqueous sodium bicarbonate solution and filtered through pad of a celite. The filtrate was concentrated and the residue was dissolved in EtOAc (60 mL), and then washed with H_2O (15 mL). The aqueous layer was extracted with EtOAc (2 × 30 mL). The combined organic layer was dried ($MgSO_4$), filtered and the filtrate was concentrated. The product was purified by column chromatography (EtOAc) to afford 2,3-*O*-isopropylidene-D-ribose **110** (0.91 g, 24%) as a colourless oil.

***R*$_f$** = 0.38 (EtOAc); **[α]$^{20}_D$** : −29.8° (*c* 0.5, $CHCl_3$) (lit.,[7] −24.7°, *c* 1.1, $CHCl_3$); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 5.31 (1H, s, H-1), 4.73 (1H, d, *J* = 5.9 Hz, H-2), 4.47 (1H, d, *J* = 5.9 Hz, H-3), 4.30 (1H, t, *J* = 3.1 Hz, H-4), 3.68-3.55 (2H, m, CH_2), 1.38 (3H, s, CH_3), 1.21 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 112.2 ($\underline{C}(CH_3)_2$), 103.1 (CH, C-1), 88.0 (CH, C-4), 87.0 (CH, C-3), 81.8 (CH, C-2), 63.8 (CH_2, C-5), 26.5 (CH_3), 24.9 (CH_3); **IR** ν_{max} (neat): 3419, 2987, 2942, 1458, 1377, 1212, 1067, 870 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_8H_{14}NaO_5^+$ = 213.0733) found 213.0728 $[M+Na]^+$. Data are in agreement with literature.[6]

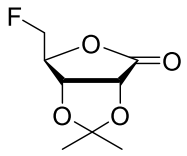
6.2.7. 2,3-*O*-Isopropylidene-D-ribono-1,4-lactone 109



K_2CO_3 (1.99 g, 14.4 mmol, 3 eq.) and iodine (3.65 g, 14.4 mmol, 3 eq.) were added to the solution of 2,3-*O*-isopropylidene-D-ribose **110** (0.91 g, 4.8 mmol, 1 eq.) in DCM (15 mL). The reaction mixture was stirred at room temperature for 4 h, then quenched by addition of a saturated aqueous solution of Na_2SO_3 (10 mL) and vigorous stirring for a further 30 min. The organic phase was isolated and the aqueous phase was extracted with EtOAc (3 × 20 mL). The combined organic phases were dried ($MgSO_4$) and solvent was evaporated. The crude mixture was then purified by column chromatography (EtOAc) to afford 2,3-*O*-isopropylidene-D-ribono-1,4-lactone **109** (0.88 g, 97%) as a white solid.

***R*$_f$** = 0.38 (EtOAc); **[α]$^{20}_D$** : −62.8° (*c* 0.25, $CHCl_3$) (lit.,[8] −62.3°, *c* 0.05, $CHCl_3$); **m.p.** 129-133 °C (lit.,[8] 136-137 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.84 (1H, d, *J* = 5.6 Hz, H-2), 4.78 (1H, d, *J* = 5.6 Hz, H-3), 4.63 (1H, br s, H-4), 4.00 (1H, d, *J* = 11.9 Hz, H-5), 3.82 (1H, d, *J* = 11.9 Hz, H-5), 1.48 (3H, s, CH_3), 1.39 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 175.5 (C=O), 113.3 ($\underline{C}(CH_3)_2$), 82.6 (CH, C-4), 78.4 (CH, C-3), 75.8 (CH, C-2), 62.2 (CH_2, C-5), 26.9 (CH_3), 25.6 (CH_3); **IR** ν_{max} (neat): 3460, 1763, 1377, 1221, 1196, 1153, 968, 854, 808, 773 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_8H_{12}NaO_5^+$ = 211.0577) found 211.0572 $[M+Na]^+$. Data are in agreement with literature.[8]

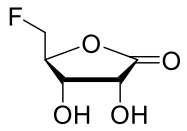
6.2.8. 5-Deoxy-5-fluoro-2,3-*O*-isopropylidene-D-ribono-1,4-lactone 108



PyFluor **104** (0.06 g, 0.36 mmol, 1.1 eq.) and MTBD (0.10 mL, 0.66 mmol, 2 eq.) were added to a solution of 2,3-*O*-isopropylidene-D-ribono-1,4-lactone **109** (62 mg, 0.33 mmol, 1 eq.) in toluene (2 mL). The reaction mixture was stirred at room temperature overnight, then quenched by addition of a saturated aqueous solution of bicarbonate (5 mL) and vigorous stirring for a further 30 min. The organic phase was isolated and the aqueous phase was extracted with EtOAc (3 × 10 mL). The combined organic phases were dried ($MgSO_4$) and the solvent evaporated. The crude mixture was concentrated and purified by column chromatography (PE/EtOAc 3:1) to afford 5-deoxy-5-fluoro-2,3-*O*-isopropylidene-D-ribono-1,4-lactone **108** (55 mg, 87%) as colourless needle-like crystals.

R_f = 0.55 (EtOAc/PE 1:1); **[α]$^{20}_D$** : −74.4° (*c* 0.25, $CHCl_3$) (lit.,[8] −82.8°, *c* 0.005, $CHCl_3$); **m.p.** 65-68 °C (lit.,[8] 60-62 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.82 (1H, d, *J* = 5.4 Hz, H-2), 4.79 (1H, dd, *J* = 5.6 Hz, 3.3 Hz, H-3), 4.75-4.59 (3H, m, H-4 & CH_2), 1.41 (3H, s, CH_3), 1.41 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 169.9 (C=O), 115.4 ($\underline{C}(CH_3)_2$), 82.6 (CH_2, d, *J* = 172.0 Hz, C-5), 80.4 (CH, d, *J* = 18.2 Hz, C-4), 77.4 (CH, d, *J* = 4.9 Hz, C-3), 75.3 (CH, d, *J* = 3.8 Hz, C-2), 26.9 (CH_3), 25.75 (CH_3); **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -235.5 (1F, dddd, *J* = 48.5 Hz, 45.3 Hz, 34.4 Hz, 3.3 Hz); **IR** ν_{max} (neat): 3545, 2995, 1788, 1383, 1269, 1055, 847 cm^{-1}; **HRMS** m/z (ES$^+$) (calculated $C_8H_{11}FNaO_4^+$ = 213.0534) found 213.0526 [M+Na]$^+$. Data are in agreement with literature.[8]

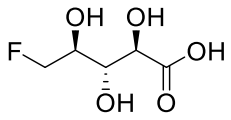
6.2.9. 5-Deoxy-5-fluoro-D-ribono-1,4-lactone 97



A solution of 5-deoxy-5-fluoro-2,3-*O*-isopropylidene-D-ribono-1,4-lactone **108** (120 mg, 0.63 mmol, 1eq.) in a mixture of TFA/H_2O (5 mL, 9:1) was stirred at room temperature for 5 h. The reaction mixture was co-evaporated five times with Et_2O (5 mL), affording 5-deoxy-5-fluoro-D-ribono-1,4-lactone **97** (90.5 mg, 96%) as white solid.

$[\alpha]^{20}_D$: +7.6° (*c* 0.25, H_2O); **m.p.** 81-85 °C; **^{1}H-NMR** (D_2O, 300 MHz) δ_H4.63-4.60 (1H, m), 4.58-4.48 (3H,m), 4.38 (1H,d, *J*= 5.5 Hz, H-2); **^{13}C-NMR** (D_2O, 100 MHz)δ_C 178.0 (C=O), 84.2 (CH, d, *J*= 17.8 Hz, C-4), 82.2 (CH_2, d, *J*= 168.5 Hz, C-5), 68.9 (CH, d, *J*= 2.1 Hz, C-2), 68.6 (CH, d, *J*= 5.1 Hz, C-3);**^{19}F-NMR** (D_2O, 376 MHz) δ_F -231.6 (1F, td, *J*= 46.7 Hz, 33.2 Hz); **IR** ν_{max} (neat): 3429, 3280, 1751, 1190, 1140, 966, 895, 781 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_5H_7FNaO_4^+$ = 173.0221) found 173.0219 $[M+Na]^+$.

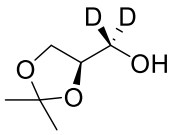
6.2.10. (2*R*,3*S*,4*S*)-5-Fluoro-2,3,4-trihydroxypentanoic acid 98



LiOH (27.8 mg, 0.66 mmol, 1.1 eq.) was added to a solution of 5-deoxy-5-fluoro-D-ribono-1,4-lactone **97** (90.5 mg, 0.6 mmol, 1 eq.) in H_2O (3 mL). The resultant mixture was stirred at room temperature for 72 h. The pH was then adjusted to 7 by addition of diluted HCl and the reaction mixture was concentrated, affording (2*R*,3*S*,4*S*)-5-fluoro-2,3,4-trihydroxypentanoic acid **98** (100 mg, 100%) as a white solid in quantitative yield.

$[\alpha]^{20}_D$: -4.0˚ (*c* 0.25, H_2O); **m.p.** 165-168 °C; **^{1}H-NMR** (D_2O, 300 MHz) δ_H4.61-4.41 (2H, m, CH_2), 4.04 (1H,d, *J*= 3.1 Hz, H-2), 3.94-3.82 (2H,m, H-3 & H-4); **^{13}C-NMR** (D_2O, 100 MHz) δ_C 177.9 (C=O), 85.0 (CH_2, d, *J*= 164.6 Hz, C-5), 73.4 (CH, C-2), 72.2 (CH, d, *J*= 7.3 Hz, C-3), 69.8 (CH, d, *J*= 17.8 Hz, C-4);**^{19}F-NMR** (D_2O, 376 MHz) δ_F -233.5 (1F, td, *J*= 47.4 Hz, 24.5 Hz); **IR** ν_{max} (neat): 3209, 1604, 1408, 1336, 1096, 1036, 949, 910, 810 cm^{-1}; **HRMS** m/z (ES^-) (calculated $C_5H_8FO_5^-$ = 167.0361) found 167.0358 $[M-H]^-$.

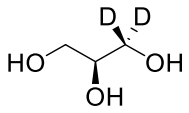
6.2.11. (*S*)-2,2-Dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane 132



A solution solution of (*S*)-methyl 2,2-dimethyl-1,3-dioxolane-4-carboxylate **131** (1.8 mL, 12.5 mmol, 1eq.) in anhydrous diethyl ether (10 mL) was added dropwise to a suspension of $LiAlD_4$ (1.0 g, 23.8 mmol, 1.9 eq.) in anhydrous diethyl ether (8.5 mL). The mixture was stirred under reflux for 1 h and then diluted with diethyl ether. After cooling the reaction mixture to 0 °C, H_2O (1 mL) was slowly added, followed by aqueous NaOH (15% w/v, 1 mL) and H_2O (3 mL) again. The solution was allowed to reach room temperature and was stirred for 15 min. The solution was dried over $MgSO_4$ and filtered to remove salts. The solvent was then evaporated affording (*S*)-2,2-dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane **132** (1.58 g, 94%) as a yellow oil.

***R*$_f$** = 0.44 (EtOAc/PE 1:1); **[α]$^{20}_D$** : -16.4° (*c* 1.0, $CHCl_3$) (lit.,[9] −14.9°, undiluted); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.21 (1H, t, *J* = 6.6 Hz, CH), 4.02 (1H, dd, *J* = 8.2 Hz, 6.6 Hz, CH_2), 3.77 (1H, dd, *J* = 8.2 Hz, 6.6 Hz, CH_2), 2.24 (1H, br s, OH), 1.42 (3H, s, CH_3), 1.35 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 109.5 ($\underline{C}(CH_3)_2$), 76.1 (CH), 65.8 (CH_2), 63.1 (CD_2, p, *J* = 21.3 Hz), 26.8 (CH_3), 25.4 (CH_3); **IR** ν_{max} (neat): 3454, 1395, 1380, 1049, 837 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_6H_{11}D_2O_3^+$ = 135.0985) found 135.0984 $[M+H]^+$. Data are in agreement with literature.[9]

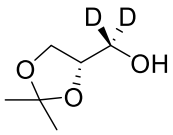
6.2.12. (*S*)-1,1-Dideuteroglycerol 116d



HCl (38%, 0.1 mL) was slowly added to a solution of (*S*)-2,2-dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane **132** (1.58 g, 11.8 mmol, 1 eq.) in H_2O (1.9 mL) and the reaction mixture was stirred at room temperature for 3 h. The solution was then concentrated affording (*S*)-1,1-dideuteroglycerol **116d** (0.87 g, 78%) as a yellow oil.

^{1}H-NMR (D_2O, 300 MHz) δ_H 3.71 (1H, dd, *J* = 6.6 Hz, 4.4 Hz, CH), 3.59 (1H, dd, *J* = 11.9 Hz, 4.4 Hz, CH_2), 3.49 (1H, dd, *J* = 11.9 Hz, 6.6 Hz, CH_2); **^{13}C-NMR** (D_2O, 100 MHz) δ_C 72.3 (CH), 62.8 (CH_2), 62.4 (CD_2, p, *J* = 22.7 Hz); **HRMS** m/z (ES^+) (calculated $C_3H_7D_2O_3^+$ = 95.0672) found 95.0672 $[M+H]^+$. Data are in agreement with literature.[9]

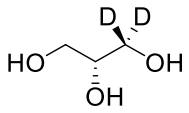
6.2.13. (*R*)-2,2-Dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane 134



A solution of (*R*)-methyl 2,2-dimethyl-1,3-dioxolane-4-carboxylate **133** (0.9 mL, 6.3 mmol, 1eq.) in anhydrous diethyl ether (5 mL) was added dropwise to a stirred suspension of $LiAlD_4$ (0.5 g, 12.6 mmol, 2 eq.) in anhydrous diethyl ether (4.5 mL). The mixture was stirred under reflux for 1 h and then diluted with diethyl ether. After cooling to 0 °C, H_2O (0.5 mL) was slowly added to the reaction, followed by a solution of aqueous NaOH (15% w/v, 1 mL) and H_2O (1.5 mL). The solution was allowed to reach room temperature and was stirred for 15 min. The solution was dried over $MgSO_4$ and filtered to remove salts. The solvent was then evaporated affording (*R*)-2,2-dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane **134** (0.5 g, 60%) as a colorless oil.

***R*$_f$** = 0.43 (EtOAc/PE 1:1); **[α]$^{20}_D$** : +15.9° (*c* 1.0, $CHCl_3$) (lit.,[9] −15.2°, undiluted); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 4.20 (1H, t, *J* = 6.6 Hz, CH), 4.01 (1H, dd, *J* = 8.2 Hz, 6.6 Hz, CH_2), 3.76 (1H, dd, *J* = 8.2 Hz, 6.6 Hz, CH_2), 2.29 (1H, br s, OH), 1.41 (3H, s, CH_3), 1.35 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 109.5 ($\underline{C}(CH_3)_2$), 76.1 (CH), 65.8 (CH_2), 63.1 (CD_2, p, *J* = 21.8 Hz), 26.8 (CH_3), 25.4 (CH_3); **IR** ν_{max} (neat): 3455, 1395, 1382, 1050, 837 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_6H_{11}D_2O_3^+$ = 135.0985) found 135.0984 $[M+H]^+$. Data are in agreement with literature.[9]

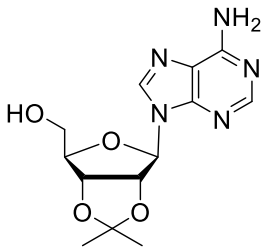
6.2.14. (*R*)-1,1-Dideuteroglycerol 116e



HCl (38%, 0.1 mL) was slowly added to a solution of (*R*)-2,2-dimethyl-4-hydroxy[$^2H_2$]methyl-1,3-dioxolane **134** (0.50 g, 3.7 mmol, 1 eq.) in H_2O (1.1 mL) and the reaction mixture was stirred at room temperature for 3 h. The solution was then concentrated affording (*R*)-1,1-dideuteroglycerol **116e** (0.25 g, 70%) as a colorless oil.

^{1}H-NMR (D_2O, 300 MHz) δ_H 3.71 (1H, dd, *J* = 6.5 Hz, 4.4 Hz, CH), 3.59 (1H, dd, *J* = 11.9 Hz, 4.4 Hz, CH_2), 3.49 (1H, dd, *J* = 11.9 Hz, 6.5 Hz, CH_2); **^{13}C-NMR** (D_2O, 100 MHz) δ_C 74.1 (CH), 64.0 (CH_2), 63.6 (CD_2, p, *J* = 22.5 Hz); **HRMS** m/z (ES^+) (calculated $C_3H_7D_2O_3^+$ = 95.0672) found 95.0672 $[M+H]^+$. Data are in agreement with literature.[9]

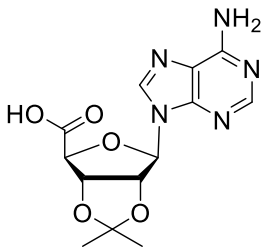
6.2.15. 2',3'-*O*-Isopropylidene-adenosine 137



Perchloric acid (1.13 mL, 18.71 mmol, 2 eq.) was added over 3 min to a suspension of adenosine **111a** (2.5 g, 9.35 mmol, 1 eq.) in acetone (50 mL) cooled to 0 °C. The reaction mixture was allowed to reach room temperature and stirred for 6 h. Aqueous ammonia (35% v/v, 25 mL) was added in 2.5 mL portions after the solution was cooled to 0 °C. The mixture was then concentrated to a white powder and stored at -20 °C overnight. The white solid was washed with diethyl ether (20 mL) and the residue was purified by silica gel column chromatography ($CHCl_3$/MeOH 4:1) to afford 2',3'-*O*-isopropylidene-adenosine **137** (1.86 g, 65%) as a white powder.

R_f = 0.66 ($CHCl_3$/MeOH 4:1); **[α]$^{20}_D$** : -64.2° (*c* 1.0, H_2O) (lit.,[10] −65°, *c* 1.0, H_2O); **m.p.** 214-215 °C (lit.,[11] 218-220 °C from DCM); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.32 (1H, s, H-8), 7.84 (1H, s, H-2), 5.85 (1H, d, *J* = 5.0 Hz, H-1'), 5.73 (2H, br s, NH_2), 5.24-5.18 (1H, m, H-2'), 5.14-5.09 (1H, m, H-3'), 4.56-4.53 (1H, m, H-4'), 4.02-3.94 (1H, m, H-5'), 3.85-9.74 (1H, m, H-5'), 1.65 (3H, s, CH_3), 1.38 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 156.0 (Ar-C), 152.8 (Ar-CH, C-8), 148.8 (Ar-C), 140.5 (Ar- CH, C-2), 121.5 (Ar-C), 114.3 ($\underline{C}(CH_3)_2$), 94.6 (CH, C-1'), 86.2 (CH, C-4'), 83.1 (CH, C-2'), 81.9 (CH, C-3'), 63.6 (CH_2, C-5'), 27.8 (CH_3), 25.4 (CH_3); **IR** ν_{max} (neat): 3720, 3110, 1688, 1605, 1210, 1090, 716 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{13}H_{17}N_5NaO_4^+$ = 330.1173) found 330.1173 $[M+Na]^+$. Data are in agreement with literature.[12]

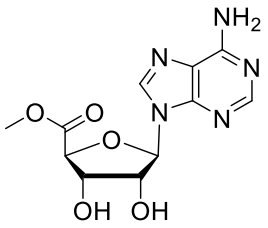
6.2.16. 2',3'-*O*-Isopropylideneadenosine-4'-dehydroxymethyl-4'-carboxylic acid 136



BAIB (2.3 g, 7.16 mmol, 2.2 eq.) and TEMPO (0.10 g, 0.65 mmol, 0.2 eq.) were added to a solution of 2',3'-*O*-isopropylidene-adenosine **137** (1.0 g, 3.25 mmol, 1 eq.) in H_2O (5 mL) and ACN (5 mL). The reaction mixture was stirred at room temperature for 3 h. The solvents were then removed by filtration and the solid was washed with Et_2O (3 × 20 mL) to afford 2',3'-*O*-isopropylideneadenosine-4'-dehydroxymethyl-4'-carboxylic acid **136** (0.75 g, 72%) as a white solid.

$[\alpha]^{20}_D$: -94.1° (*c* 1.0, $CHCl_3$); **m.p.** 251-255 °C (lit.,[13] 250-253 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.24 (1H, s, H-8), 8.08 (1H, s, H-2), 7.28 (2H, br s, NH_2), 6.33 (1H, s, H-1'), 5.54 (1H, dd, *J* = 6.0 Hz, 1.9 Hz, H-3'), 5.46 (1H, d, *J* = 6.0 Hz, H-2'), 4.68 (1H, d, *J* = 1.9 Hz, H-4'), 1.52 (3H, s, CH_3), 1.35 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 170.7 (CO_2H), 156.4 (Ar-C), 152.4 (Ar-CH, C-2), 149.1 (Ar-C), 140.4 (Ar- CH, C-8), 119.2 (Ar-C), 113.0 ($\underline{C}(CH_3)_2$), 89.5 (CH, C-1'), 85.4 (CH, C-3'), 83.8 (CH, C-2'), 83.4 (CH, C-4'), 26.5 (CH_3), 24.9 (CH_3); **IR** ν_{max} (neat): 3051, 1694, 1600, 1554, 1124, 1080 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{13}H_{16}N_5O_5^+$ = 322.1146) found 322.1146 $[M+H]^+$. Data are in agreement with literature.[13]

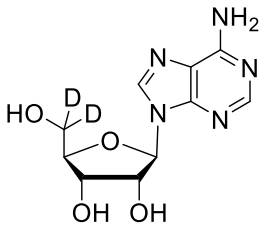
6.2.17. Adenosine-5'-carboxylic acid methyl ester 135



2',3'-*O*-Isopropylideneadenosine-4'-dehydroxymethyl-4'-carboxylic acid **136** (1.0 g, 3.1 mmol, 1 eq.) was suspended in MeOH (70 mL) and cooled to 0 °C. After 30 min stirring, $SOCl_2$ (1.1 mL, 15.6 mmol, 5 eq.) was added dropwise. The reaction mixture was stirred at room temperature overnight. Solvent and $SOCl_2$ were removed under reduced pressure and co-evaporated with toluene (3 × 20 mL). The residue was purified by silica gel column chromatography (DCM/MeOH 9:1) to afford adenosine-5'-carboxylic acid methyl ester **135** (0.74 g, 81%) as a yellow powder.

***R*$_f$** = 0.36 (DCM/MeOH 9:1); **[α]$^{20}_{D}$** : -24.7° (*c* 2.3, $CHCl_3$) (lit.,[14] −23°, *c* 2.34, HCl); **m.p.** 200-207 °C (lit.,[14] 220-222 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.59 (1H, s, H-8), 8.21 (1H, s, H-2), 6.21 (1H, d, *J* = 6.0 Hz, H-1'), 4.64-4.60 (2H, m, H-2' & H-3'), 4.46 (1H, dd, *J* = 4.6 Hz, 3.0 Hz, H-4'), 3.82 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 171.1 (C=O), 156.5 (Ar-C), 152.6 (Ar-CH, C-2), 148.9 (Ar-C), 139.9 (Ar- CH, C-8), 119.2 (Ar-C), 88.1 (CH, C-1'), 82.8 (CH, C-3'), 75.0 (CH, C-2'), 73.6 (CH, C-4'), 51.7 (CH_3); **IR** ν_{max} (KBr): 3431, 2921, 1734, 1667, 1601, 1584, 1485, 1334, 1301, 1284, 1200, 1099, 1068, 901, 660 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{11}H_{14}N_5O_5^+$ = 296.0989) found 296.0989 $[M+H]^+$. Data are in agreement with literature.[14]

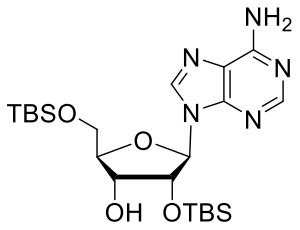
6.2.18. 5',5'-Dideuterioadenosine 111c



Adenosine-5'-carboxylic acid methyl ester **135** (0.65 g, 2.2 mmol, 1 eq.) was suspended in EtOH (16 mL) and cooled to 0 °C. $NaBD_4$ (0.46 g, 11.0 mmol, 5 eq.) was added in portions over 20 min. The reaction mixture was then heated to 110 °C and stirred for 2 d. The reaction mixture was cooled and concentrated, before being dissolved in H_2O (30 mL) and the pH adjusted to pH 6.5 with HCl (1 M). The resulting solution was concentrated and the residue was purified by silica gel column chromatography (EtOAc/^{i}PrOH/H_2O 7:2:1) to afford 5',5'-dideuterioadenosine **111c** (0.49 g, 82%) as a white powder.

***R*$_f$** = 0.28 (EtOAc/^{i}PrOH/H_2O 7:2:1); **[α]$^{20}_D$** : −61.4° (*c* 1.0, H_2O); **m.p.** 233-236 °C (lit.,[15] 234-236 °C from H_2O); **^{1}H-NMR** (DMSO-*d*$_6$, 300 MHz) δ_H 8.32 (1H, s, H-8), 8.18 (1H, s, H-2), 5.97 (1H, d, *J* = 6.5 Hz, H-1'), 4.75 (1H, t, *J* = 5.8 Hz, H-2'), 4.63 (2H, br s, NH_2), 4.33 (1H, dd, *J* = 5.0 Hz, *J* = 2.5 Hz, H-3'), 4.17 (1H, d, *J* = 2.5 Hz, H-4'); **^{13}C-NMR** (DMSO-*d*$_6$, 100 MHz) δ_C 156.2 (Ar-C), 152.1 (Ar-CH, C-8), 149.5 (Ar-C), 141.3 (Ar- CH, C-2), 119.0 (Ar-C), 88.9 (CH, C-1'), 85.1 (CH, C-4'), 72.8 (CH, C-3'), 70.1 (CH, C-2'), 62.9 (CH_2, p, *J* = 21.2 Hz, C-5'); **IR** ν_{max} (KBr): 3432, 2920, 1667, 1605, 1584, 1481, 1331, 1299, 1200, 1103, 1070, 900, 660 cm^{-1}; **HRMS** m/z (ES$^+$) (calculated $C_{10}H_{12}D_2N_5O_4^+$ = 270.1166) found 269.1098 [2%] and 270.1169 $[M+H]^+$.

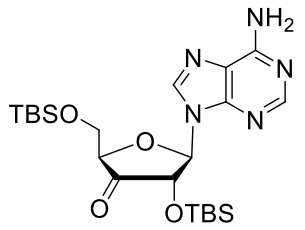
6.2.19. 2',5'-Bis-*O*-(*tert*-butyldimethylsilyl)-β-D-adenosine 141



TBSCl (6.8 g, 45.0 mmol, 3 eq.) was added to a solution of adenosine **111a** (4 g, 15.0 mmol, 1 eq.) in pyridine (32 mL). After 2 d stirring at room temperature, the reaction mixture was diluted with DCM (30 mL), washed with H_2O (3 × 20 mL), dried over $MgSO_4$, filtered and concentrated. The residue was purified by column chromatography (EtOAc/Et_2O 1:2) to afford 2',5'-bis-*O*-(*tert*-butyldimethylsilyl)-β-D-adenosine **141** (2.5 g, 34%) as a white solid.

***R*$_f$** = 0.38 (EtOAc/Et_2O 1:2); **m.p.** 175-178 °C (lit.,[16] 172-175 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.37 (1H, s, H-8), 8.24 (1H, s, H-2), 6.13 (1H, d, *J* = 5.1 Hz, H-1'), 5.65 (2H, s, NH_2), 4.67 (1H, t, *J* = 5.0 Hz, H-2'), 4.31 (1H, t, *J* = 3.6 Hz, H-3'), 4.23 (1H, q, *J* = 2.7 Hz, H-4'), 4.04 (1H, dd, *J* = 11.5 Hz, 2.7 Hz, H-5'), 3.88 (1H, dd, *J* = 11.5 Hz, 2.5 Hz, H-5'), 2.77 (1H, d, *J* = 4.2 Hz, OH-3'), 0.98 (9H, s, $(CH_3)_3$), 0.87 (9H, s, $(CH_3)_3$), 0.18 (3H, s, CH_3), 0.16 (3H, s, CH_3), -0.03 (3H, s, CH_3), -0.09 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 155.5 (Ar-C), 153.1 (Ar-CH, C-2), 150.0 (Ar-C), 138.9 (Ar-CH, C-8), 119.8 (Ar-C), 87.9 (CH, C-1'), 85.2 (CH, C-4'), 77.1 (CH, C-2'), 71.3 (CH, C-3'), 63.1 (CH_2, C-5'), 26.0 (CH_3, ($\underline{C}H_3)_3$), 25.6 (CH_3, ($\underline{C}H_3)_3$), 18.5 (Si$\underline{C}$-$(CH_3)_3$), 18.0 (Si$\underline{C}$-$(CH_3)_3$), -5.00 (CH_3, Si-$\underline{C}H_3$), -5.2 (CH_3, Si-$\underline{C}H_3$), -5.3 (CH_3, Si-$\underline{C}H_3$), -5.5 (CH_3, Si-$\underline{C}H_3$); **IR** ν_{max} (KBr): 3376, 3330, 3180, 2937, 2857, 1651, 1595, 1473, 1329, 1255, 1140, 1105, 835, 781 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{22}H_{42}N_5O_4Si_2^+$ = 496.2770) found 496.2757 $[M+H]^+$. Data are in agreement with literature.[17]

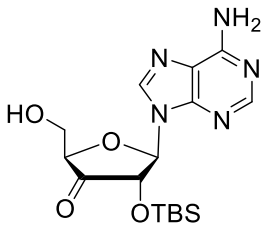
6.2.20. 9-[2',5'-Bis-*O*-(*tert*-butyldimethylsilyl)-β-D-erythro-pentofuranos-3'-ulosyl]adenine 140



Pyridine (1.98 mL, 24.0 mmol, 4 eq.) and Ac_2O (1.14 mL, 12.0 mmol, 2eq.) were sequentially added to an ice cold suspension of CrO_3 (1.2 g, 12.0 mmol, 2 eq.) in DCM (90 mL), and stirring at room temperature under argon atmosphere was continued until a homogeneous solution was obtained. A solution of 2',5'-bis-*O*-(*tert*-butyldimethylsilyl)-β-D-adenosine **141** (3 g, 6.0 mmol, 1 eq.) in DCM (90 mL) was added dropwise and stirring was continued for 2 h. The reaction mixture was poured into cold EtOAc (150 mL) and the resulting precipitate was filtered (glass microfibre filter). The filtrate was concentrated and the residue was purified by column chromatography (EtOAc) to afford 9-[2',5'-bis-*O*-(*tert*-butyldimethylsilyl)-β-D-erythro-pentofuranos-3'-ulosyl]adenine **140** (1.48 g, 50%) as a white solid.

R_f = 0.31 (EtOAc); **m.p.** 176-180 °C (lit.,[16] 177-178 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.39 (1H, s, H-8), 8.17 (1H, s, H-2), 6.16 (1H, d, *J* = 8.3 Hz, H-1'), 5.72 (2H, s, NH_2), 4.96 (1H, d, *J* = 8.3 Hz, H-2'), 4.35-4.31 (1H, m, H-4'), 4.00 (2H, d, *J* = 2.3 Hz, H-5'), 0.95 (9H, s, $(CH_3)_3$), 0.75 (9H, s, $(CH_3)_3$), 0.13 (3H, s, CH_3), 0.10 (3H, s, CH_3), -0.01 (3H, s, CH_3), -0.18 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 208.6 (C=O, C-3'), 155.4 (Ar-C), 153.4 (Ar-CH, C-2), 150.4 (Ar-C), 138.6 (Ar-CH, C-8), 119.8 (Ar-C), 85.0 (CH, C-1'), 82.4 (CH, C-4'), 77.9 (CH, C-2'), 62.5 (CH_2, C-5'), 25.9 (CH_3, ($\underline{C}H_3)_3$), 25.3 (CH_3, ($\underline{C}H_3)_3$), 18.3 (Si$\underline{C}$-$(CH_3)_3$), 18.0 (Si$\underline{C}$-$(CH_3)_3$), -4.8 (CH_3, Si-$\underline{C}H_3$), -5.4 (CH_3, Si-$\underline{C}H_3$), -5.6 (CH_3, Si-$\underline{C}H_3$), -5.7 (CH_3, Si-$\underline{C}H_3$); **IR** ν_{max} (KBr): 3330, 3188, 2931, 2859, 1788, 1651, 1595, 1475, 1308, 1250, 1114, 1057, 970 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{22}H_{40}N_5O_4Si_2^+$ = 494.2613) found 494.2602 $[M+H]^+$. Data are in agreement with literature.[18]

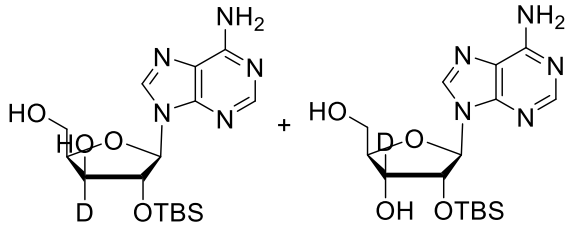
6.2.21. 9-[2-*O*-(*Tert*-butyldimethylsilyl)-β-D-erythro-pentofuran-3-ulosyl]adenine 139



9-[2',5'-Bis-*O*-(tert-butyldimethylsilyl)-β-D-erythro-pentofuranos-3'-ulosyl]adenine **140** (1 g, 2.0 mmol, 1 eq.) was dissolved in a mixture of TFA (24.3 mL) and H_2O (2.7 mL). After 30 min stirring at 0 °C, the yellow solution was poured into a mixture of EtOAc (60 mL), cyclohexane (40 mL), and aqueous saturated bicarbonate solution (100 mL). The phases were separated and the aqueous phase was extracted with a mixture of EtOAc (60 mL) and cyclohexane (40 mL). The combined organic phases were washed with H_2O (60 mL), dried over Na_2SO_4, filtered and concentrated to afford 9-[2-*O*-(*tert*-butyldimethylsilyl)-β-D-erythro-pentofuran-3-ulosyl]adenine **139** as viscous yellow oil, which was used in the following reaction without further purification.

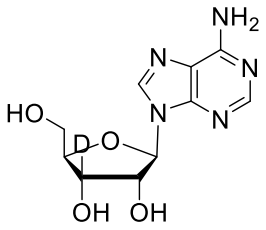
^{1}H-NMR ($CDCl_3$, 300 MHz) δ_H 8.68 (1H, s, H-8), 8.65 (1H, s, H-2), 6.21 (1H, d, *J* = 7.8 Hz, H-1'), 4.88 (1H, d, *J* = 7.8 Hz, H-2'), 4.50 (1H, t, *J* = 1.8 Hz, H-4'), 4.11 (2H, m, H-5'), 0.77 (9H, s, $(CH_3)_3$), 0.04 (3H, s, CH_3), -0.17 (3H, s, CH_3). Data are in agreement with literature.[18]

6.2.22. 2'-*O*-(*Tert*-butyldimethylsilyl)-[3'-^{2}H]-adenosine 138 and 2'-*O*-(*tert*-butyldimethylsilyl)-[3'-^{2}H]-β-D-xylofurannosyl-9 adenine 144



$NaBD_4$ (268 mg, 6.4 mmol, 2 eq.) was added carefully to a solution of 9-[2-*O*-(*tert*-butyldimethylsilyl)-β-D-erythro-pentofuran-3-ulosyl]adenine **139** (1.2 g, 3.2 mmol, 1 eq.) in AcOH (60 mL) cooled to 0 °C. The resulting solution was stirred overnight at room temperature. EtOAc (50 mL) was added to the reaction mixture and the phases were separated. The aqueous phase was extracted with EtOAc (3 × 50 mL), dried over $MgSO_4$, filtered and concentrated to afford a mixture of 2'-*O*-(*tert*-butyldimethylsilyl)-[3'-^{2}H]-adenosine **138** and 2'-*O*-(*tert*-butyldimethylsilyl)-[3'-^{2}H]-β-D-xylofurannosyl-9 adenine **144** in a ratio of 2:1, which was used in the following reaction without further purification.

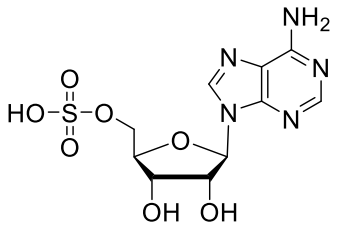
6.2.23. 3'-Deuterioadenosine 111d



TBAF **53** (1 M in dry THF, 4.8 mL, 4.8 mmol, 1.5 eq.) was added dropwise to a solution of 2'-*O*-(*tert*-butyldimethylsilyl)-[3'-^{2}H]-adenosine **138** and 2'-*O*-(*tert*-butyldimethylsilyl)-[3'-^{2}H]-β-D-xylofurannosyl-9 adenine **144** in dry THF (100 mL). The reaction mixture was stirred at room temperature for 2 h. The volatiles were removed *in vacuo* and the residue was purified by column chromatography (EtOAc/^{i}PrOH/H_2O 7:2:1) to afford 3'-deuterioadenosine **111d** (177 mg, 33% over 3 steps) as a white powder.

***R*$_f$** = 0.28 (EtOAc/^{i}PrOH/H_2O 7:2:1); **[α]$^{20}_{D}$** : −59.9° (*c* 1.0, H_2O); **m.p.** 233-237 °C; **^{1}H-NMR** (DMSO-d_6, 300 MHz) δ_H 8.35 (1H, s, H-8), 8.14 (1H, s, H-2), 7.36 (2H, s, NH_2), 5.88 (1H, d, *J* = 6.2 Hz, H-1'), 5.51 – 5.40 (2H, m, OH-2' & OH-5'), 5.19 (1H, s, OH-3'), 4.60 (1H, t, *J* = 5.5 Hz, H-2'), 3.96 (1H, t, *J* = 3.6 Hz, H-4'), 3.67 (1H, dt, *J* = 12.2 Hz, 4.0 Hz, H-5'), 3.55 (1H, ddd, *J* = 12.2 Hz, 7.2 Hz, 3.6 Hz, H-5'); **^{13}C-NMR** (DMSO-d_6, 100 MHz) δ_C 156.6 (Ar-C), 152.8 (Ar-CH, C-2), 149.6 (Ar-C), 140.4 (Ar-CH, C-8), 119.8 (Ar-C), 88.4 (CH, C-1'), 86.3 (CH, C-4'), 73.9 (CH, C-2'), 70.9 (CD, t, *J* = 21.3 Hz, C-3'), 62.1 (CH_2, C-5'); **IR** ν_{max} (KBr): 3430, 2920, 1666, 1605, 1580, 1486, 1330, 1300, 1200, 1100, 1069, 901, 660; **HRMS** m/z (ES^+) (calculated $C_{10}H_{13}DN_5O_4^+$ = 269.1103) found 268.0877 [4%] and 269.1097 $[M+H]^+$. Data are in agreement with literature.[18]

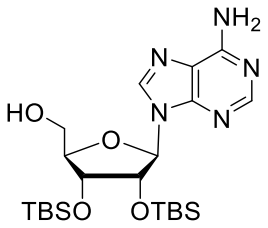
6.2.24. Adenosine sulfate 160



A solution of chlorosulfonic acid **150** (30 µL, 0.45 mmol, 1.2 eq.) in $CHCl_3$ (0.2 mL) was slowly added to a solution of adenosine **111a** (100 mg, 0.37 mmol, 1 eq.) in dry pyridine (5 mL) at 0 °C. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under reduced pressure, and the residue was passed through a short Discovery® DSC-NH_2 SPE ion exchange column. The white solid was purified by semi-preparative HPLC (Phenomenex Kingsorb C18 (250 × 21.20 mm, 5µ)), using the HPLC conditions: Mobile phase A: 0.05% TFA/H_2O, B: 0.05% TFA/ACN; an isocratic mobile phase of: 95% A and 5% B for 30 min. Under these conditions, the product eluted between 6.9 to 7.9 min. Evaporation of the solvents from this fraction, followed by freeze-drying afforded adenosine sulfate **160** (10 mg, 6%) as a white solid.

$[\alpha]^{20}_D$: -11.6° (*c* 0.5, H_2O); **m.p.** 98-101 °C; **^{1}H-NMR** (D_2O, 300 MHz) δ_H 8.34 (1H, s, H-8), 8.18 (1H, s, H-2), 6.26 (1H, d, *J* = 6.6 Hz, H-1'), 5.38 (1H, t, *J* = 6.2 Hz, H-2'), 5.16 (1H, dd, *J* = 5.6 Hz, *J* = 2.9 Hz, H-3'), 4.59-4.51 (1H, m, H-4'), 4.27 (2H, m, H-5'); **^{13}C-NMR** (D_2O, 100 MHz) δ_C 154.7 (Ar-C), 151.6 (Ar-CH, C-8), 149.2 (Ar-C), 140.4 (Ar-CH, C-2), 118.6 (Ar-C), 85.0 (CH, C-1'), 81.5 (CH, C-4'), 76.6 (CH, C-2'), 75.3 (CH, C-3'), 67.2 (CH_2, C-5'); **IR** ν_{max} (neat): 3400, 2955, 1686, 1601, 1568, 1499, 1349, 1318, 1288, 1199, 1142, 1101, 901, 660 cm^{-1}; **HRMS** m/z (ES^-) (calculated $C_{10}H_{12}N_5O_7S^-$ = 346.0463) found 346.0463 $[M-H]^-$. Data are in agreement with literature.[19]

6.2.25. 2',3'-*O*-Bis(*tert*-butyldimethylsilyl)-adenosine 165

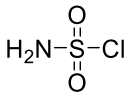


To a solution of adenosine **111a** (5.0 g, 18.7 mmol, 1.0 eq.) and imidazole (11.5 g, 168.3 mmol, 9.0 eq.) in DMF (25 mL) was added TBSCl (9.9 g, 65.5 mmol, 3.5 eq.) in DMF (25 mL). The resulting solution was stirred at room temperature overnight, then diluted with DCM (150 mL) and washed with saturated bicarbonate solution (3 × 80 mL). The combined organic layers were washed with brine (3 × 80 mL), dried over $MgSO_4$, filtered and concentrated. The resulting liquid was dissolved into 80% AcOH (25 mL) and stirred at 100 °C for 4 h. The reaction mixture was concentrated and the residue was co-evaporated with toluene (3 × 40 mL). The residue was purified by silica gel column chromatography (EtOAc) to afford 2',3'-*O*-bis(*tert*-butyldimethylsilyl)-adenosine **165** (3.7 g, 40%) as a white solid.

***R*$_f$** = 0.48 (EtOAc); **[α]$^{20}_D$** : -113.8° (*c* 1.0, $CHCl_3$); **m.p.** 242-247 °C (lit.,[20] 180 °C from EtOH); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.31 (1H, s, H-8), 7.80 (1H, s, H-2), 6.69 (1H, dd, *J* = 12.4 Hz, *J* = 1.6 Hz, H-4'), 5.75 (1H, d, *J* = 7.9 Hz, H-1'), 5.71 (2H, br s, NH_2), 5.01 (1H, dd, *J* = 8.0 Hz, *J* = 4.6 Hz, H-2'), 4.30 (1H, d, *J* = 4.6 Hz, H-3'), 4.13 (1H, s, OH), 3.91 (1H, d, *J* = 13.0 Hz, H-5'), 3.67 (1H, t, *J* = 12.6 Hz, H-5'), 0.92 (9H, s, $(CH_3)_3$), 0.71 (9H, s, $(CH_3)_3$), 0.10 (3H, s, CH_3), 0.08 (3H, s, CH_3), -0.17 (3H, s, CH_3), -0.65 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 153.0 (Ar-C), 148.9 (Ar-CH, C-8), 148.7 (Ar-C), 141.3 (Ar- CH, C-2), 119.2 (Ar-C), 91.2 (CH, C-1'), 89.8 (CH, C-4'), 74.2 (CH, C-3'), 74.0 (CH, C-2'), 63.2 (CH_2, C-5'), 26.0 ($(CH_3)_3$), 25.8 ($(CH_3)_3$), 18.8 ($\underline{C}(CH_3)_3$), 18.2 ($\underline{C}(CH_3)_3$), -4.4 (CH_3), -4.4 (CH_3), -4.5 (CH_3), -5.9 (CH_3); **IR** ν_{max} (neat): 3421, 3331, 3156, 2948, 2910, 1650, 1634, 1470, 1330, 1274, 1138, 1111, 821, 778

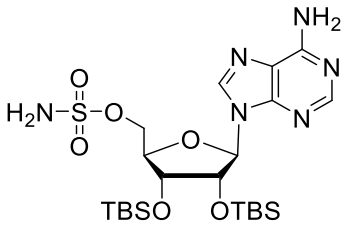
cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{22}H_{42}N_5O_4Si_2^+$ = 496.2770) found 496.2770 $[M+H]^+$. Data are in agreement with literature.[21]

6.2.26. Amidosulfonyl chloride 162



Formic acid **166** (0.75 mL, 20 mmol, 1 eq.) was added dropwise to a solution of chlorosulfonyl isocyanate **167** (1.74 mL, 20 mmol, 1 eq.) in dry ACN (10 mL) at 0 °C. The reaction mixture was stirred at room temperature overnight. The resulting solution of amidosulfonyl chloride **162** was used in the following reaction without any further purification.

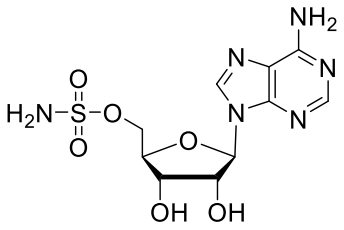
6.2.27. 2',3'-*O*-Bis(*tert*-butyldimethylsilyl)-5'-*O*-(sulfamoyl)adenosine 164



A chilled solution of amidosulfonyl chloride **162** (2 M in dry ACN, 0.12 mL, 0.23 mmol, 2.3 eq.) was added to an ice cold solution of 2',3'-*O*-bis(*tert*-butyldimethylsilyl)-adenosine **165** (50 mg, 0.1 mmol, 1 eq.). After stirring at room temperature for 1 h, the reaction mixture was cooled to 0 °C and quenched by subsequent additions of Et_3N (1 mL) and MeOH (4 mL). The volatiles were removed *in vacuo* and the residue was taken up into EtOAc (10 mL), extracted with aqueous saturated bicarbonate solution (2 × 10 mL), and brine (2 × 10 mL). The organic layer was dried over Na_2SO_4, filtered and evaporated. The residue was purified by silica gel column chromatography (EtOAc) to afford 2',3'-*O*-bis(*tert*-butyldimethylsilyl)-5'-*O*-(sulfamoyl)adenosine **164** (57 mg, quantitative) as a white solid.

R_f = 0.55 (EtOAc); **[α]$^{20}_{D}$** : -149.6° (*c* 1.0, $CHCl_3$); **m.p.** 118-130 °C; **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.36 (1H, s, H-8), 8.15 (1H, s, H-2), 7.65 (2H, br s, NH_2), 7.32 (2H, br s, NH_2), 5.95 (1H, d, *J* = 6.8 Hz, H-1'), 4.96 (1H, dd, *J* = 6.7 Hz, *J* = 4.5 Hz, H-2'), 4.42-4.36 (2H, m, H-3' & H-5'), 4.31 (1H, dd, *J* = 11.0 Hz, *J* = 5.8 Hz, H-5'), 4.17 (1H, td, *J* = 5.5 Hz, *J* = 1.9 Hz, H-4'), 0.92 (9H, s, $(CH_3)_3$), 0.70 (9H, s, $(CH_3)_3$), 0.14 (3H, s, CH_3), 0.12 (3H, s, CH_3), -0.09 (3H, s, CH_3), -0.38 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 156.1 (Ar-C), 152.7 (Ar-CH, C-2), 149.5 (Ar-C), 139.8 (Ar-CH, C-8), 119.3 (Ar-C), 87.0 (CH, C-1'), 82.8 (CH, C-4'), 73.7 (CH, C-2'), 72.5 (CH, C-3'), 68.0 (CH_2, C-5'), 25.7 ($(CH_3)_3$), 25.5 ($(CH_3)_3$), 17.8 ($\underline{C}(CH_3)_3$), 17.5 ($\underline{C}(CH_3)_3$), -4.7 (CH_3), -4.78 (CH_3), -4.84 (CH_3), -5.6 (CH_3); **IR** ν_{max} (neat): 2930, 2857, 1645, 1601, 1472, 1362, 1252, 1163, 995, 939, 835, 775 cm^{-1}; **HRMS** m/z (ES^-) (calculated $C_{22}H_{41}N_6O_6SSi_2^-$ = 573.2352) found 573.2351 $[M-H]^-$. Data are in agreement with literature.[22]

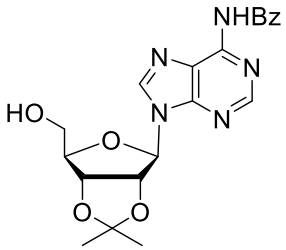
6.2.28. 5'-Sulfamoyladenosine 161



TBAF **53** (1 M in THF, 0.87 mL, 0.87 mmol, 2.5 eq.) was added dropwise to a solution of 2',3'-*O*-bis(*tert*-butyldimethylsilyl)-5'-*O*-(sulfamoyl)adenosine **164** (201 mg, 0.35 mmol, 1 eq.) in dry THF (10 mL). The reaction mixture was stirred at room temperature for 1 h, and the solvent was then evaporated. The residue was purified by semi-preparative HPLC, using this HPLC condition: Mobile phase A: 0.05% TFA/H_2O, B: 0.05% TFA/ACN; Gradient: 95% A and 5% B. Under these conditions, the product eluted between 5.2 to 6.0 min. Evaporation of the solvents from this fraction, followed by freeze-drying afforded 5'-sulfamoyladenosine **161** (100 mg, 83%) as a white solid.

$[\alpha]^{20}_D$: -31.9° (*c* 1.0, H_2O) (lit.,[23] -33.6° *c* 1.0, DMF); **m.p.** 109-120 °C (lit.,[24] 98-112 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.30 (1H, s, H-8), 8.21 (1H, s, H-2), 6.07 (1H, d, *J* = 5.1 Hz, H-1'), 4.66 (1H, t, *J* = 5.1 Hz, H-2'), 4.46-4.28 (4H, m, H-3' & H-4' & H-5'); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 157.4 (Ar-C), 154.0 (Ar-CH, C-8), 150.7 (Ar-C), 140.8 (Ar-CH, C-2), 120.4 (Ar-C), 89.9 (CH, C-1'), 83.7 (CH, C-4'), 75.6 (CH, C-2'), 71.9 (CH, C-3'), 70.0 (CH_2, C-5'); **IR** ν_{max} (neat): 3251, 1682, 1630, 1447, 1375, 1211, 1192, 1069, 853, 635 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{10}H_{14}N_6O_6S^+$ = 347.0768) found 347.0761 $[M+H]^+$. Data are in agreement with literature.[24]

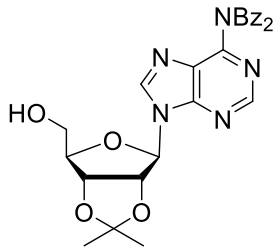
6.2.29. N^6-Monobenzyl-2',3'-*O*-isopropylidene-adenosine 175



2',3'-*O*-Isopropylidene-adenosine (200 mg, 0.65 mmol, 1 eq.) **137** was co-evaporated twice with anhydrous pyridine and dissolved in pyridine (4 mL) followed by addition of trimethylsilyl chloride (0.44 mL, 3.46 mmol, 5.3 eq.) at 0 °C. After 30 min stirring at room temperature, benzoyl chloride (0.10 mL, 0.86 mmol, 1.3 eq.) was added and stirring was continued for 3 h at room temperature. The reaction mixture was cooled down to 0 °C and diluted with H_2O (1.5 mL). The reaction mixture was treated with a concentrated aqueous ammonia solution (2 mL) and stirred for 30 min at room temperature. The mixture was concentrated and the residue was purified by column chromatography (EtOAc) to afford N^6-monobenzyl-2',3'-*O*-isopropylidene-adenosine **175** (88.8 mg, 33%) as a white powder.

***R*$_f$ =** 0.33 (DCM/MeOH 9:1); **[α]$^{20}_D$** : -92.8° (*c* 1.0, $CHCl_3$) (lit.,[25] -83.0° *c* 2.4, DCM); **m.p.** 136-138 °C (lit.,[25] 146-148 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 9.30 (1H, br s, NH), 8.74 (1H, s, H-8), 8.09 (s, 1H, H-2), 8.06-7.44 (5H, m, ArH), 5.95 (1H, d, *J* = 4.5 Hz, H-1'), 5.26-5.16 (1H, m, H-2'), 5.14-5.03 (1H, m, H-3'), 4.55-4.49 (1H, m, H-4'), 4.03-3.90 (1H, m, H-5'), 3.86-3.72 (1H, m, H-5'), 1.63 (3H, s, CH_3), 1.37 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 164.8 (C, C=O), 152.4 (CH, C-8), 150.6 (C, Ar), 150.4 (C, Ar), 142.7 (CH, C-2), 133.5 (C, Ar), 133.0 (C, ArH), 129.0 (C, ArH), 128.0 (C, ArH), 124.4 (C, Ar), 114.4 (C, $\underline{C}(CH_3)_2$), 94.1 (CH, C-1'), 86.4 (CH, C-4'), 83.3 (CH, C-2'), 81.7 (CH, C-3'), 63.3 (CH_2, C-5'), 27.7 (CH_3), 25.3 (CH_3); **IR** ν_{max} (neat): 3350, 1709, 1605, 1590, 1465, 1329, 1255, 1210, 1057, 701 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{20}H_{21}N_5NaO_5^+$ = 434.1435) found 434.1425 $[M+Na]^+$. Data are in agreement with literature.[26]

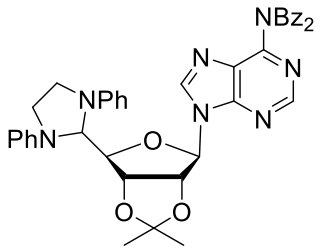
6.2.30. N^6,N^6-Dibenzoyl-2',3'-*O*-isopropylidene-adenosine 173



2',3'-*O*-Isopropylidene-adenosine **137** (2.0 g, 6.52 mmol, 1eq.) was co-evaporated twice with anhydrous pyridine and dissolved in pyridine (24 mL) followed by the addition of trimethylsilyl chloride (2.48 mL, 19.56 mmol, 3 eq.) at 0 °C. The reaction mixture was stirred for 2 h at room temperature. The mixture was cooled to 0 °C and benzoyl chloride (1.67 mL, 14.36 mmol, 2.2 eq.) was slowly added. The reaction mixture was stirred for 2 h at room temperature before being quenched with H_2O (8 mL) and concentrated. The mixture was reconstituted with $CHCl_3$ (60 mL) and washed with diluted H_2SO_4 **149** solution (2 × 40 mL), saturated aqueous sodium bicarbonate solution (40 mL), and H_2O (40 mL). The organic layer was dried over $MgSO_4$ and concentrated. The mixture was purified by column chromatography (EtOAc) to afford N^6,N^6-dibenzoyl-2',3'-*O*-isopropylidene-adenosine **173** (2.35 g, 70%) as a white foam.

R_f = 0.59 (EtOAc); **[α]$^{20}_{D}$** : -68.3° (*c* 0.60, $CHCl_3$); **m.p.** 124-125 °C (lit.,[11] 129-133 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.62 (1H, s, H-8), 8.19 (1H, s, H-2), 7.90-7.31 (10H, m, Ar-CH), 5.97 (1H, d, *J* = 4.6 Hz, H-1'), 5.26-5.14 (1H, m, H-2'), 5.07-4.99 (1H, m, H-3'), 4.52-4.48 (1H, m, H-4'), 3.98-3.85 (1H, m, H-5'), 3.81-3.67 (1H, m, H-5'), 1.63 (3H, s, CH_3), 1.36 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 172.2 (C=O), 152.6 (Ar-C), 151.9 (Ar-C), 151.8 (Ar-CH, C-8), 144.5 (Ar-CH, C-2), 133.9 (Ar-C), 133.2 (Ar-CH), 129.6 (Ar-CH), 128.9 (Ar-CH), 128.6 (Ar-C), 114.4 ($\underline{C}(CH_3)_2$), 93.9 (CH, C-1'), 86.3 (CH, C-4'), 83.3 (CH, C-2'), 81.6 (CH, C-3'), 63.2 (CH_2, C-5'), 27.6 (CH_3), 25.3 (CH_3); **IR** ν_{max} (neat): 2989, 1700, 1600, 1576, 1238, 1077, 830, 694, 641 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{27}H_{25}N_5NaO_6^+$ = 538.1697) found 538.1692 $[M+Na]^+$. Data are in agreement with literature.[11]

6.2.31. N^6,N^6-Dibenzoyl-5'-deoxy-2',3'-*O*-isopropylidene-5',5'-(*N*,*N*'-diphenylethylenediamino)adenosine 172

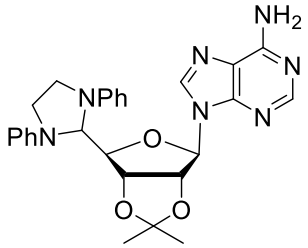


Dichloroacetic acid (96 µL, 1.17 mmol, 0.5 eq.) was added dropwise to a solution of N^6,N^6-dibenzoyl-2',3'-*O*-isopropylidene-adenosine **173** (1.2 g, 2.33 mmol, 1 eq.) and DCC **153** (1.44 g, 6.99 mmol, 3 eq.) in anhydrous DMSO (15 mL) at 0 °C. The mixture was allowed to warm to room temperature and was stirred overnight. A solution of oxalic acid (587 mg, 4.66 mmol, 2 eq.) in methanol (2.5 mL) was slowly added. After 30 min stirring at room temperature, the mixture was filtered and the residue was washed with cold methanol. *N*,*N*'-diphenylethylenediamine (569 mg, 2.68 mmol, 1.2 eq.) was added to the combined filtrate and washings and the resulting solution was gently stirred at room temperature for additional 3 h. H_2O (5 mL) was added and after storage at 0 °C crystals were obtained. The mother liquor was partitioned between H_2O (40 mL) and $CHCl_3$ (40 mL) and the organic phase was washed with H_2O (2 × 40 mL), dried over $MgSO_4$ and concentrated. The crude mixture was purified by column chromatography (EtOAc/PE 1:1) to afford N^6,N^6-dibenzoyl-5'-deoxy-2',3'-*O*-isopropylidene-5',5'-(*N*,*N*'-diphenylethylenediamino)adenosine **172** (708 mg, 43%) as a white solid.

***R*$_f$** = 0.41 (EtOAc/PE 1:1); **[α]$^{20}_D$** : +45.4° (*c* 1.10, $CHCl_3$); **m.p.** 139-141 °C (lit.,[27] 135-138 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.47 (1H, s, H-8), 7.96 (1H, s, H-2), 7.90-7.85 (4H, m, ArH), 7.54-7.47 (2H, m, ArH), 7.41-7.33 (4H, m, ArH), 7.24-7.12 (4H, m, ArH), 6.83-6.70 (4H, m, ArH), 6.64 (2H, d, *J* = 8.1 Hz, ArH), 6.13 (1H, d, *J* = 2.0 Hz, H-1'), 5.73 (1H, d, *J* = 2.5 Hz, N-CH-N), 5.19-5.12 (2H, m, H-2', H-3'), 4.63-4.59 (1H, m, H-4'), 3.74-3.45 (4H, m, N-CH_2-CH_2-N), 1.51 (3H, s, CH_3), 1.32 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 172.4 (C=O),

152.5 (Ar-CH, C-8), 146.4 (Ar-C), 143.8 (Ar-CH, C-2), 134.2 (Ar-C), 133.2 (Ar-CH), 129.6 (Ar-CH), 129.5 (Ar-CH), 129.4 (Ar-CH), 128.9 (Ar-C), 118.6 (Ar-C), 118.5 (Ar-C), 115.3 ($\underline{C}(CH_3)_2$), 113.8 (Ar-CH), 113.6 (Ar-CH), 88.7 (CH, C-1'), 87.1 (CH, C-4'), 83.8 (CH, C-2'), 80.2 (CH, C-3'), 73.4 (N-CH-N), 47.6 (CH_2), 47.2 (CH_2), 27.5 (CH_3), 25.8 (CH_3); **IR** ν_{max} (neat): 1710, 1602, 1501, 1238, 1077, 750, 511 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{41}H_{37}N_7NaO_5^+$ = 730.2748) found 730.2756 $[M+Na]^+$. Data are in agreement with literature.[27]

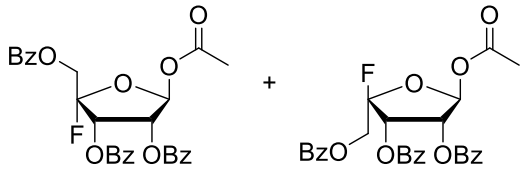
6.2.32. 5'-Deoxy-2',3'-*O*-isopropylidene-5',5'-(*N*,*N*'-diphenylethylenediamino)adenosine 177



A solution of N^6,N^6-dibenzoyl-5'-deoxy-2',3'-*O*-isopropylidene-5',5'-(*N*,*N*'-diphenylethylenediamino) adenosine **172** (540 mg, 0.76 mmol, 1 eq.) in anhydrous methanol (5 mL) was transferred into a heavy-walled tube. Ammonia was bubbled through the solution for 2 h at 0 °C. The tube was sealed and the mixture stirred at 60 °C for 18 h. The mixture was cooled to room temperature and carefully opened, followed by degassing with argon for 2 h before being concentrated. The product was purified by column chromatography (EtOAc) to afford 5'-deoxy-2',3'-*O*-isopropylidene-5',5'-(*N*,*N*'-diphenylethylenediamino)adenosine **177** (316 mg, 83%) as a white powder.

***R*$_f$** = 0.18 (EtOAc); **[α]$^{20}_D$** : +51.1° (*c* 1.10, $CHCl_3$); **m.p.** 249-250 °C (lit.,[27] 255-258 °C); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H 8.31 (1H, s, H-8), 7.60 (1H, s, H-2), 7.25-7.14 (4H, m, ArH), 6.88-6.73 (4H, m, ArH), 6.69-6.63 (2H, m, ArH), 6.11 (1H, d, *J* = 2.3 Hz, H-1'), 5.96 (2H, br s, NH_2), 5.78 (1H, d, *J* = 3.3 Hz, N-CH-N), 5.20-5.11 (2H, m, H-2', H-3'), 4.62-4.57 (1H, m, H-4'), 3.79-3.55 (4H, m, N-CH_2-CH_2-N), 1.50 (3H, s, CH_3), 1.32 (3H, s, CH_3); **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C 155.7 (Ar-C), 153.4 (Ar-CH, C-8), 146.8 (Ar-C), 139.5 (Ar-CH, C-2), 129.4 (Ar-C), 119.9 (Ar-C), 118.4 (Ar-CH), 114.9 ($\underline{C}(CH_3)_2$), 113.8 (Ar-CH), 113.6 (Ar-CH), 88.7 (CH, C-1'), 87.2 (CH, C-4'), 84.0 (CH, C-2'), 80.4 (CH, C-3'), 73.7 (N-CH-N), 47.9 (CH_2), 46.8 (CH_2), 27.5 (CH_3), 25.8 (CH_3); **IR** ν_{max} (neat): 3351, 1600, 1528, 1238, 1077, 741, 519 cm^{-1}; **HRMS** m/z (ES$^+$) (calculated $C_{27}H_{30}N_7O_3^+$ = 500.2405) found 500.2397 $[M+H]^+$. Data are in agreement with literature.[28]

6.2.33. 1-*O*-Acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose 180 and 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-α-L-lyxofuranose 184

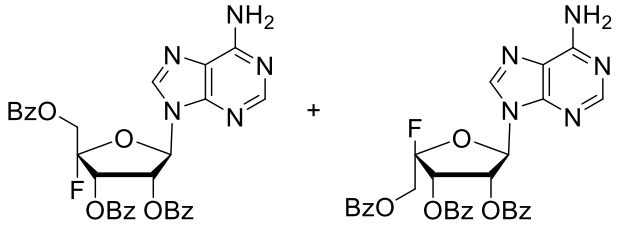


N-bromosuccinimide (534 mg, 3 mmol, 3 eq.) was added to a solution of 1-*O*-acetyl-2,3,5-tri-benzoyl-β-D-ribofuranose **179** (504 mg, 1 mmol, 1 eq.) in carbon tetrachloride (6 mL) and the mixture was stirred under irradiation with a sun lamp (275 W) at 90 °C for 5 h. The reaction mixture was diluted with DCM (10 mL) and washed with aqueous sodium thiosulfate (10 mL), aqueous sodium bicarbonate (10 mL) and H_2O (10 mL), and dried over anhydrous sodium sulfate. Concentration of the solvent gave a crude product mixture of the 4-bromo-β-D-ribofuranose **182** and 4-bromo-α-L-lyxofuranose **183**. In a round bottom flask silver tetrafluoroborate was prepared from the reaction of silver fluoride (189 mg, 1.5 mmol, 1.5 eq.) with borontrifluoride-diethyl ether (377 µL, 3.0 mmol, 3 eq.) in anhydrous diethyl ether (10 mL) for 1 h at room temperature. An ethyl acetate (5 mL) solution of the bromide mixture was added to the suspension of $AgBF_4$ and the mixture was stirred for 1 h at 0 °C. The reaction mixture was diluted with diethyl ether (20 mL), washed with aqueous saturated sodium bicarbonate (15 mL), H_2O (15 mL) and brine (15 mL), and dried over anhydrous sodium sulfate. The solvents were removed under reduced pressure and the residue was subjected to silica gel column chromatography (PE/EtOAc 2:1) to give an inseparable mixture of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-α-L-lyxofuranose **184** and 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose **180** (199 mg, 38%) in a ratio of 1:1 as a colourless gum.

1-*O*-Acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose 180; R_f = 0.34 (EtOAc/PE 2:1); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H signals were not interpreted due to overlapping; **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C signals were not interpreted due to overlapping; **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -114.89 (1F, ddd, J = 17.9 Hz, 10.5 Hz, 4.5 Hz); **HRMS** m/z (ES^+) (calculated $C_{28}H_{23}FNaO_9^+$ = 545.1218) found 545.1230 $[M+Na]^+$.

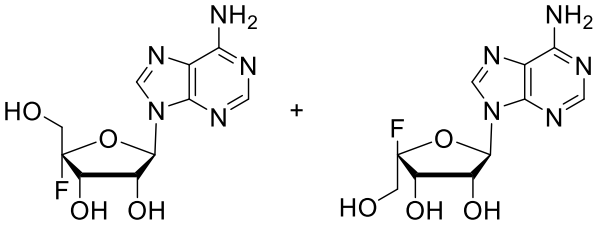
1-*O*-Acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-α-L-lyxofuranose 184; R_f = 0.34 (EtOAc/PE 2:1); **^{1}H-NMR** ($CDCl_3$, 300 MHz) δ_H signals were not interpreted due to the overlap; **^{13}C-NMR** ($CDCl_3$, 100 MHz) δ_C signals were not interpreted due to the overlap; **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -104.30 - -104.49 (1F, m); **HRMS** m/z (ES^+) (calculated $C_{28}H_{23}FNaO_9^+$ = 545.1218) found 545.1230 $[M+Na]^+$.

6.2.34. 9-(2',3',5'-Tri-*O*-benzoyl-4'-fluoro-β-D-ribofuranosyl)adenine 181 and 9-(2',3',5'-tri-*O*-benzoyl-4'-fluoro-α-L-lyxofuranosyl)adenine 185



To a suspension of adenine **82** (58 mg, 0.43 mmol, 5 eq.) in ACN (4 mL) was added *N*,*O*-bis(trimethylsilyl)trifluoro acetamide (228 µL, 0.86 mmol, 10 eq.) and the mixture was stirred for 2 h at 50 °C to give silylated adenine. The 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose **180**/1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-α-L-lyxofuranose **184** (45 mg, 0.086 mmol, 1 eq.) mixture and TMSOTf (72 µL, 0.43 mmol, 5 eq.) were then added to the silylated adenine solution and the resulting mixture was stirred overnight at 90 °C. The reaction was quenched by adding saturated aqueous sodium bicarbonate (8 mL). The aqueous phase was extracted with ethyl acetate (10 mL) and the combined organic layers were washed with brine (15 mL), dried over anhydrous sodium sulfate, filtered, and concentrated to afford a mixture of 9-(2',3',5'-tri-*O*-benzoyl-4'-fluoro-β-D-ribofuranosyl)adenine **181** and 9-(2',3',5'-tri-*O*-benzoyl-4'-fluoro-α-L-lyxofuranosyl)adenine **185**, which was used in the following reaction without further purification.

6.2.35. 4'-Fluoro adenosine 171 and 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose 186



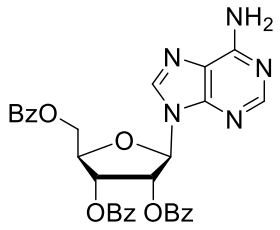
To a MeOH (5 mL) solution of 9-(2',3',5'-tri-*O*-benzoyl-4'-fluoro-β-D-ribofuranosyl)adenine **181** and 9-(2',3',5'-tri-*O*-benzoyl-4'-fluoro-α-L-lyxofuranosyl)adenine **185** (51 mg, 0.086 mmol, 1 eq.) was added sodium carbonate (91 mg, 0.86 mmol, 10 eq.) and the reaction mixture was stirred for 1 h at room temperature. The reaction mixture was filtered through sintered glass and concentrated. The mixture was passed through a short C-18 RP silica gel column (H_2O:ACN (9:1)). Fractions containing **171** and **186** were combined and concentrated *in vacuo*. The residues were purified by semi-preparative HPLC (Phenomenex Kingsorb C18 (250 × 21.20 mm, 5µ)), using the HPLC conditions: Mobile phase A: 0.05% TFA/ H_2O; B: 0.05% TFA/ACN; an isocratic mobile phase of: 95% A and 5% B for 25 min. Under these conditions, the product eluted between 8.1 to 8.7 min. Evaporation of the solvents from this fraction, followed by freeze-drying afforded a mixture of 4'-fluoro adenosine **171** and 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose **186** (2.5 mg, 10%) in a ratio of 1:1 as a white solid.

4'-Fluoro adenosine 171; **^{1}H-NMR** (DMSO-d_6, 300 MHz) δ_H 8.55 (1H, s, H-8), 8.26 (1H, s, H-2), 6.14 (1H, d, J = 4.5 Hz, H-1'), 4.40 (1H, dd, J = 14.9 Hz, J = 7.2 Hz, H-3'), 4.33 (1H, m, H-2'), 3.78 (2H, m, H-5'); **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -127.5 (1F, dt, J = 17.1 Hz, 6.6 Hz); **HRMS** m/z (ES^+) (calculated $C_{10}H_{13}FN_5O_4^+$ = 286.0946) found 286.0947 $[M+H]^+$. Data are in agreement with literature.[29]

1-*O*-Acetyl-2,3,5-tri-*O*-benzoyl-4-fluoro-β-D-ribofuranose 186; **^{1}H-NMR** (DMSO-d_6, 300 MHz) δ_H 8.44 (1H, s, H-8), 8.22 (1H, s, H-2), 6.56 (1H, d, J = 4.8 Hz, H-1'), 5.67 (1H, t,

J = 6.4 Hz, H-2'), 5.42 (1H, dd, *J* = 5.5 Hz, *J* = 3.0 Hz, H-3'), 4.51 (2H, m, H-5'); **^{19}F-NMR** ($CDCl_3$, 376 MHz) δ_F -111.8 - -112.2 (1F, m); **HRMS** m/z (ES^+) (calculated $C_{10}H_{13}FN_5O_4^+$ = 286.0946) found 286.0947 $[M+H]^+$.

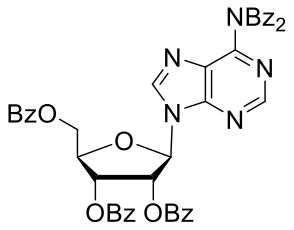
6.2.36. 2',3',5'-Tri-*O*-benzoyladenosine 187



Benzoic anhydride (2.26 g, 10.0 mmol, 5 eq.) and 4-(*N*,*N*-dimethylamino)pyridine (61 mg, 0.05 mmol, 0.025 eq.) were added to a suspension of adenosine **111a** (535 mg, 2.0 mmol, 1 eq.) in dry pyridine (20 mL). After 2 h stirring at room temperature, the mixture was concentrated to half-volume, diluted with $CHCl_3$ (30 mL) and washed three times with 5% $NaHCO_3$ (3 × 20 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated to dryness under reduced pressure. The residue was purified by silica gel column chromatography (EtOAc) to afford 2',3',5'-tri-*O*-benzoyladenosine **187** (345 mg, 30%) as a white solid.

***R*$_f$** = 0.42 (EtOAc); **[α]$^{20}_D$** : -72.4° (*c* 1.0, $CHCl_3$) (lit.,[30] -73.7° *c* 1.0, $CHCl_3$); **m.p.** 97-99 °C (lit.,[30] 90-94 °C from H_2O); **^{1}H-NMR** ($CHCl_3$, 300 MHz) δ_H 8.29 (1H, s, H-8), 8.13-7.92 (7H, m, ArH), 7.63-7.34 (9H, m, ArH), 6.42 (H, d, *J* = 5.1 Hz, H-1'), 6.38 (1H, t, *J* = 5.3 Hz, H-2'), 6.25 (1H, t, *J* = 5.3 Hz, H-3'), 5.61 (2H, br s, NH_2), 4.90 (1H, dd, *J* = 12.1 Hz, *J* = 3.3 Hz, H-5'), 4.85-4.80 (1H, m, H-4'), 4.71 (1H, dd, *J* = 12.1 Hz, *J* = 4.3 Hz, H-5'); **^{13}C-NMR** ($CHCl_3$, 100 MHz) δ_C 166.6 (C=O), 165.7 (C=O), 165.5 (C=O), 156.0 (Ar-CH, C-2), 154.2 (Ar-C), 153.9 (Ar-CH, C-8), 150.0 (Ar-C), 139.6 (Ar-CH), 139.4 (Ar-CH), 134.3 (Ar-CH), 133.6 (Ar-CH), 133.3 (Ar-CH), 130.6 (Ar-CH), 130.1 (Ar-CH), 130.1 (Ar-CH), 129.8 (Ar-CH), 129.3 (Ar-C), 128.9 (Ar-C), 128.5 (Ar-C), 119.3 (Ar-C), 87.1 (CH, C-1'), 80.8 (CH, C-2'), 75.1 (CH, C-4'), 71.5 (CH, C-3'), 63.9 (CH_2, C-5'); **IR** ν_{max} (KBr): 3200, 2920, 1729, 1640, 1595, 1328, 1231, 1047, 899, 658 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{31}H_{26}N_5O_7^+$ = 580.1827) found 580.1828 $[M+H]^+$. Data are in agreement with literature.[31]

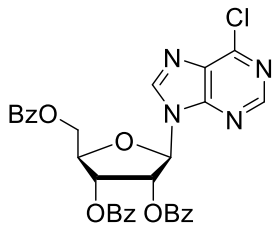
6.2.37. 6-*N*,*N*-2',3',5'-Tri-*O*-pentabenzoyladenosine 190



Adenosine **111a** (2.0 g, 7.5 mmol, 1 eq.) was suspended in dry pyridine (20 mL). To the suspension was added benzoyl chloride (6.1 mL, 52.3 mmol, 7 eq.), and the mixture was stirred at 65 °C for 4 h. Ethanol (20 mL) was added to the reaction mixture and the solvents were removed *in vacuo*. Residue was diluted with DCM (300 mL), washed H_2O (3 × 200 mL) and brine (200 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated to dryness under reduced pressure. The residue was purified by silica gel column chromatography (EtOAc/PE 1:1) to afford 6-*N*,*N*-2',3',5'-tri-*O*-pentabenzoyladenosine **190** (1.2 g, 20%) as a colorless foam.

***R*$_f$** = 0.39 (EtOAc/PE 1:1); **[α]$^{20}_D$** : -79.7° (*c* 1.0, $CHCl_3$); **m.p.** 182-185 °C (lit.,[32] 185-187 °C); **^{1}H-NMR** ($CHCl_3$, 300 MHz) δ_H 8.51 (1H, s, H-8), 8.22 (1H, s, H-2), 8.11-8.07 (2H, m, ArH), 8.02-7.97 (2H, m, ArH), 7.96-7.91 (2H, m, ArH), 7.87-7.82 (4H, m, ArH), 7.61-7.32 (15H, m, ArH), 6.47 (H, d, *J* = 5.2 Hz, H-1'), 6.39 (1H, t, *J* = 5.5 Hz, H-2'), 6.26 (1H, t, *J* = 5.3 Hz, H-3'), 4.90 (1H, dd, *J* = 12.1 Hz, *J* = 3.2 Hz, H-5'), 4.86-4-82 (1H, m, H-4'), 4.71 (1H, dd, *J* = 12.1 Hz, *J* = 4.3 Hz, H-5'); **^{13}C-NMR** ($CHCl_3$, 100 MHz) δ_C 171.4 (C=O), 167.7 (C=O), 166.6 (C=O), 165.7 (C=O), 165.5 (C=O), 155.1 (Ar-CH, C-2), 154.0 (Ar-C), 153.6 (Ar-CH, C-8), 150.2 (Ar-C), 139.6 (Ar-CH), 139.4 (Ar-CH), 134.3 (Ar-CH), 134.1 (Ar-CH), 134.0 (Ar-CH), 133.6 (Ar-CH), 133.3 (Ar-CH), 132.6 (Ar-CH), 131.9 (Ar-CH), 131.1 (Ar-CH), 130-128 signals were not interpreted due to the overlap, 127.5 (Ar-C), 119.6 (Ar-C), 86.9 (CH, C-1'), 80.8 (CH, C-2'), 73.7 (CH, C-4'), 71.2 (CH, C-3'), 63.2 (CH_2, C-5'); **IR** ν_{max} (neat): 1730, 1655, 1601, 1357, 1230, 1048, 902, 666 cm^{-1}; **HRMS** m/z (ES$^+$) (calculated $C_{45}H_{33}N_5NaO_9^+$ = 810.2170) found 810.2169 $[M+Na]^+$. Data are in agreement with literature.[33]

6.2.38. 6-Chloro-9-(2',3',5'-tri-*O*-benzoyl-β-D-pentofuranosyl)purine 196



Benzoyl chloride (1.05 mL, 8.2 mmol, 4.1 eq.) was added to a suspension of 6-chloropurine riboside **195** (573 mg, 2.0 mmol, 1 eq.) in dry pyridine (12 mL) and the mixture was stirred at room temperature overnight. Ethanol (12 mL) was added to the reaction mixture and the solvents were removed *in vacuo*. The residue was diluted in DCM (100 mL) and washed with brine (3 × 100 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated to dryness under reduced pressure. The residue was purified by silica gel column chromatography (EtOAc/PE 1:2) to afford 2',3',5'-tri-*O*-benzoyladenosine **196** (345 mg, 30%) as a white solid.

***R*$_f$** = 0.36 (EtOAc/PE 1:2); **[α]$^{20}_D$** : -65.5° (*c* 0.8, $CHCl_3$) (lit.,[34] -63.8° *c* 0.085, $CHCl_3$); **m.p.** 115-119 °C (lit.,[34] 114-115 °C); **^{1}H-NMR** ($CHCl_3$, 300 MHz) δ_H 8.60 (1H, s, H-8), 8.27 (1H, s, H-2), 8.10-7.89 (7H, m, ArH), 7.64-7.51 (3H, m, ArH), 7.50-7.33 (5H, m, ArH), 6.45 (H, d, *J* = 5.1 Hz, H-1'), 6.41 (1H, t, *J* = 5.2 Hz, H-2'), 6.25 (1H, t, *J* = 5.2 Hz, H-3'), 4.94 (1H, dd, *J* = 12.1 Hz, *J* = 3.2 Hz, H-5'), 4.88-4-82 (1H, m, H-4'), 4.70 (1H, dd, *J* = 12.1 Hz, *J* = 4.1 Hz, H-5'); **^{13}C-NMR** ($CHCl_3$, 100 MHz) δ_C 167.1 (C=O), 166.4 (C=O), 166.0 (C=O), 156.1 (Ar-CH, C-2), 154.7 (Ar-C), 154.1 (Ar-CH, C-8), 151.0 (Ar-C), 147.2 (Ar-CH), 144.9 (Ar-CH), 135.2 (Ar-CH), 134.0 (Ar-CH), 133.7 (Ar-CH), 132.8 (Ar-CH), 130.5 (Ar-CH), 130.1 (Ar-CH), 130.0 (Ar-CH), 129.3 (Ar-C), 128.7 (Ar-C), 128.5 (Ar-C), 121.3 (Ar-C), 80.9 (CH, C-1'), 75.7 (CH, C-2'), 70.5 (CH, C-4'), 69.2 (CH, C-3'), 65.3 (CH_2, C-5'); **IR** ν_{max} (neat): 1749, 1604, 1585, 1550, 1288, 1118, 742 cm^{-1}; **HRMS** m/z (ES^+) (calculated $C_{31}H_{23}ClN_4NaO_7^+$ = 621.1147) found 602.1147 $[M+Na]^+$. Data are in agreement with literature.[35]

6.3. Biological methods

6.3.1. General method

Reagents were obtained from commercial sources (Acros Organic, Sigma, Fisher) and used without further purification.

All microbiological works were carried out using standard sterile techniques under a Gallenkamp laminar flowhood. Experiments involving cell-free extract solutions were carried out at 4 °C or on ice. Glassware, equipment and consumable for biological works were sterilised by autoclaving, flaming or spraying with 70% ethanol, as appropriate prior to use. Media were sterilised by autoclaving. Cell cultures were incubated in a temperature controlled Gallenkamp orbit incubator or an Innova 2000 platform shaker. Centrifugation was carried out on a Beckman Avanti centrifuge. Micro-centrifugation was carried out on a Hettich Mikro 200 bench-top centrifuge. Sonication was performed with a Vibra Cell apparatus from Sonics & Materials Inc.

Optical density was measured at 600 nm using a Jenway 6300 Spectrophotometer. Protein concentrations were determined by a Nanodrop ND1000 Spectrophotometer.

^{19}F-NMR spectra were recorded with proton-decoupling in 10% D_2O in H_2O at 298 K on a Burker Avance 500 MHz spectrometer and calibrated with an external reference ($CFCl_3$).

6.3.2. Cultures of *Streptomyces calvus*

Streptomyces calvus was grown on solid agar plates composed of soluble starch (10 g), dipotassium phosphate (1 g), magnesium sulfate VSP (1 g), sodium chloride (1 g), ammonium sulfate (2 g), calcium carbonate (2 g), ferrous sulfate (1 mg), manganous chloride (1 mg), zinc sulfate (1 mg), agar (20 g) and deionised H_2O (1 L). The medium ISP4 was sterilised by autoclave before use. The plates were maintained at 30 °C where the bacteria mature after a

period of 20 days. The spores were collected by means of sterilised cotton swabs and stored at -80 °C in a 50% glycerol solution (Seed culture).[36]

A mass of the mycelium of *S. calvus* was obtained by inoculating a sterilised, defined medium (100 mL) with the spores obtained above (Seed culture, use 1 to 5 µl per flask) and allowed to grow at 28 °C for 12 days in a 500 mL conical flask with shaking at 180 rpm. The defined medium is composed of tap H_2O (1 L), corn steep liquor (12.5 g), mannitol (10 g), sodium chloride (2 g), Hoagland's salt solution (1 mL), magnesium sulfate (0.25 g), monopotassium phosphate (1.5 g), potassium fluoride (7.5 mL, 0.5 M) and diammonium phosphate (2 g).

Hoagland's salt solution contains deionised H_2O (1 L), H_3PO_3 (0.611 g), $MnCl_2.4H_2O$ (0.389 g), $CuSO_4$ (0.056 g), $ZnSO_4.7H_2O$ (0.056 g), $Al_2(SO_4)_3$ (0.056 g), $NiSO_4.6H_2O$ (0.056 g), $Co(NO_3)_2.6H_2O$ (0.056 g), $(NH_4)_6Mo_7O_{24}.4H_2O$ (0.056 g), TiO_2 (0.056 g), LiCl (0.028 g), $SnCl_2.2H_2O$ (0.028 g), KI (0.028 g) and KBr (0.028 g), and was sterilised by autoclaving.

After 18 days of fermentation, the cells were discarded by centrifugation and the supernatant was extracted with *n*-butanol (20 mL). The organic layer was concentrated under reduced pressure. The extract was analysed by ^{19}F-NMR (3600 scans).

6.3.3. Pulse feeding experiments

Cultures of *S. calvus* (100 mL) were shaken at 30 °C and labelled glycerol was added after 4 days; the same quantity of labelled glycerol was added every two days for a total of 6 additions. The final concentration of labelled glycerol was between 8-10 mM. After 21 days of fermentation, the cells were discarded by centrifugation and the supernatant was extracted into *n*-butanol (20 mL). The organic layer was concentrated under reduced pressure. The extract was analysed by ^{19}F-NMR (3600 scans) to detect nucleocidin.

6.3.4. Cell-free extract experiments

The cells of *Streptomyces* sp. MA 37 were harvested by centrifugation (20 min, 13 000 rpm, 4 °C) after an 8 day fermentation. The cell pellets were washed three times using Tris–HCl buffer (20 mM, pH 7.5) supplemented with 10 mM $MgCl_2$ to remove remnant culture media. The cells were re-suspended in the same buffer (0.1 g wet-cell weight per mL). The cells were disrupted by ultrasonication (60% duty cycle for 30–60 s). Cell debris were removed by centrifugation (30 min, 13 000 rpm, 4 °C) and the resultant clear supernatant was used as the cell free extract for incubation experiments. The cell free extracts (1 mL) were supplemented with or without 5-FDR at 37 °C for 6 hours. At the end of the incubation period, protein was precipitated by heating the vial to 90 °C for 3 min and the protein was then removed by centrifugation. The supernatant was collected for ^{19}F-NMR analysis.[37]

The cells of *Streptomyces calvus* were harvested by centrifugation (20 min, 13 000 rpm, 4 °C) after 7 day fermentation. The cell pellets were washed three times using Tris–HCl buffer (20 mM, pH 7.0) supplemented with 15 mM $MgCl_2$ or phosphate buffer (15 mM, pH 7.0) supplemented with 15 mM $MgCl_2$ to remove remnant culture media. The cells were re-suspended in the same buffer used previously (0.1 g wet-cell weight per mL). The cells were disrupted by ultrasonication (60% duty cycle for 30–60 s). Cell debris was removed by centrifugation (30 min, 13 000 rpm, 4 °C) and the resultant clear supernatant was used as the cell free extract for incubation experiments. The cell free extracts (3 mL) were supplemented with a mix of substrates and co-factors at 37 °C overnight. At the end of the incubation period, protein was precipitated by heating the vial to 90 °C for 3 min and the protein was then removed by centrifugation. The supernatant was collected for ^{19}F-NMR analysis.

6.4. References

1. D. D. Perrins, W. L. Armarego, *Purification of laboratory chemicals*. 3rd Ed., Pergamon Press, Oxford.
2. N. A. Ivanova, Z. R. Valiullina, O. V. Shitikova, M. S. Miftakhov, *Russ. J. Org. Chem.*, 2007, **43**, 742-746.
3. X-G. Li, S. Dall'Angelo, L. F. Schweiger, M. Zanda, D. O'Hagan, *Chem. Commun.*, 2012, **48**, 5247-5249.
4. M. K. Nielsen, C. R. Ugaz, W. Li, A. G. Doyle, *J. Am. Chem. Soc.*, 2015, **137**, 9571-9574.
5. N. F. Taylor, P. W. Kent, *J. Chem. Soc.*, 1958, 872-875.
6. J. H. Cho, D. L. Bernard, R. W. Sidwell, E. R. Kern, C. K. Chu, *J. Med. Chem.*, 2006, **49**, 1140-1148.
7. S. Pedatella, A. Guaragna, D. D'Alonzo, M. De Nisco, G. Palumbo, *Synthesis*, 2006, **2**, 305-308.
8. P. Nasomjai, D. O'Hagan, A. M. Z. Slawin, *Beilstein J. Org. Chem.*, 2009, **5**, 37.
9. R. E. Hill, A. Iwanow, B. G. Sayer, W. Wysocka, I. D. Spencer, *J. Biol. Chem.*, 1987, **262**, 7463-7471.
10. J. Baddiley, A. R. Todd, *J. Chem. Soc.*, 1947, 648-651.
11. S. Thompson, S. A. McMahon, J. H. Naismith, D. O'Hagan, *Bioorg. Chem.*, 2016, **64**, 37-41.
12. G. Zhang, S. L. Richardson, Y. Mao, R. Huang, *Org. Biomol. Chem.*, 2015, **13**, 4149-4154.
13. O. Moukha-Chafiq, R. C. Reynolds, *Nucleosides, Nucleotides Nucleic Acids*, 2014, **33**, 53-63.
14. R. N. Prasad, A. Fung, K. Tietje, H. Stein, H. D. Brondyk, *J. Med. Chem.*, 1976, **19**, 1180-1186.

15. H. Pang, K. H. Schram, D. L. Smith, S. P. Gupta, L. B. Townsend, J. A. McCloskey, *J. Org. Chem.*, 1982, **47**, 3923-3932.
16. C. Klinchan, Y-L. Hsu, L-C. Lo, W. Pluempanupat, P. Chuawong, *Tetrahedron Lett.*, 2014, **55**, 6204-6207.
17. V. Samano, M. J. Robins, *J. Org. Chem.*, 1991, **56**, 7108-7113.
18. M. J. Robins, S. Sarker, V. Samano, S. F. Wnuk, *Tetrahedron*, 1997, **53**, 447-456.
19. F. Egami, N. Takahashi, *Bull. Chem. Soc. Jpn.*, 1955, **28**, 666-668.
20. R. Takasawa, I. Yoshikawa, K. Araki, *Org. Biomol. Chem.*, 2004, **2**, 1125-1132.
21. S. Lun, H. Guo, J. Adamson, J. S. Cisar, T. D. Davis, S. S. Chavadi, J. D. Warren, L. E. N. Quadri, D. S. Tan, W. R. Bishai, *Antimicrob. Agents Chemother.*, 2013, **57**, 5138-5140.
22. P. Van de Vijver, T. Ostrowski, B. Sproat, J. Goebels, O. Rutgeerts, A. Van Aerschot, M. Waer, P. Herdewijn, *J. Med. Chem.*, 2008, **51**, 3020-3029.
23. D. A. Shuman, R. K. Robins, M. J. Robins, *J. Am. Chem. Soc.*, 1969, **91**, 3391-3392.
24. J. L. Lukkarila, S. R. da Silva, M. Ali, V. M. Shahani, G. Wei Xu, J. Berman, A. Roughton, S. Dhe-Paganon, A. D. Schimmer, P. T. Gunning, *ACS Med. Chem. Lett.*, 2011, **2**, 577-582.
25. A. R. Maguire, W-D. Meng, S. M. Roberts, A. J. Willetts, *J. Chem. Soc., Perkin Trans. 1*, 1993, **15**, 1795-1808.
26. A. M. Jawalekar, N. Meeuwenoord, J. G. O. Cremers, H. S. Overkleeft, G. A. Van der Marel, F. P. J. T. Rutjes, F. L. Van Delft, *J. Org. Chem.*, 2008, **73**, 287-290.
27. E. T. Jarvi, J. R. McCarthy, S. Mehdi, D. P. Matthews, M. L. Edwards, N. J. Prakash, T. L. Bowlin, P. S. Sunkara, P. Bey, *J. Med. Chem.*, 1991, **34**, 647-656.
28. S. Liu, S. F. Wnuk, C. Yuan, M. J. Robbins, R. T. Borchardt, *J. Med. Chem.*, 1993, **36**, 883-887.
29. T. Wada, T. Moriguchi, M. Sekine, *J. Am. Chem. Soc.*, 1994, **116**, 9901-9911.

30. N. Shimomura, T. Matsutani, T. Mukaiyama, *Bull. Chem. Soc. Jpn.*, 1994, **67**, 3100-3106.
31. S. Lee, C. Uttamapinant, G. L. Verdine, *Org. Lett.*, 2007, **9**, 5007-5009.
32. M. Smith, D. H. Rammler, I. H. Goldberg, H. G. Khorana, *J. Am. Chem. Soc.*, 1962, **84**, 430-440.
33. G. A. Brown, E. D. Savory, J. V. A. Ouzman, A. M. Stoddart, *Improved synthesis of 2-substituted adenosines*, WO 2005/056571 A1.
34. V. Yadav, C. K. Chu, R. H. Rais, O. N. Al Safarjalani, V. Guarcello, F. N. M. Naguib, M. H. El Kouni, *J. Med. Chem.*, 2004, **47**, 1987-1996.
35. S. Kozai, S. Takamatsu, K. Izawa, T. Maruyama, *Tetrahedron Lett.*, 1999, **40**, 4355-4358.
36. K. K. J. Chan, D. O'Hagan, *Meth. Enzymol.*, 2012, **516**, 219-235.
37. L. Ma, A. Bartholome, M. H. Tong, Z. Qin, Y. Yu, T. Shepherd, K. Kyeremeh, H. Deng, D. O'Hagan, *Chem. Sci.*, 2015, **6**, 1414-1419.