Expanding the terpenome: Complementary approaches to novel terpenoids

A thesis submitted to Cardiff University
for the degree of Doctor of Philosophy by:

Daniel John Grundy

Supervisor: Rudolf K. Allemann

2015

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed …… (candidate) Date ……

Abstract

Terpenes and terpenoids make up one of the largest and most structurally diverse families of natural products with known compounds numbering in the tens of thousands. A significant proportion of terpenoids are secondary metabolites, which display a wide range of biological activities, and consequently many have found uses as therapeutic drugs, most notably paclitaxel, the anti-cancer agent; and artemisinin, a key drug in anti-malarial therapies. Due to the challenges in synthesising these highly complex compounds in satisfactory yields the majority of the supply of these drugs is still extracted from natural sources or produced semi-synthetically from naturally extracted compounds.

Terpene synthases convert the linear isoprenyl diphosphates, precursors of all terpenoids, into complex hydrocarbon skeletons, which are often naturally further derivatised by P450 cytochromes. Recently there have been advances in exploiting terpene synthases for the production of terpenoid precursors, but despite these advances selectively derivatising the hydrocarbon skeleton remains the major barrier to the conversion of these skeletons to therapeutically viable compounds.

This thesis focuses on exploring methods of generating simple terpenoids from terpene synthases. The project is divided into two parts both designed to investigate methods of generating simple terpenoids. The first part focuses on germacradien-4-ol synthase (GdolS), a recently discovered bacterial sesquiterpene synthase, which incorporates water in the cyclisation cascade to generate a single terpene alcohol. This enzyme was characterised and the mechanism probed using substrate analogues and single-point mutations to elucidate how GdolS is able to incorporate water into its product while still protecting the highly reactive carbocation intermediates. Our data indicated that germacradien-4-ol synthase catalyses the loss of the diphosphate group and ring closure in a stepwise manner, followed by a 1,3-hydride shift to generate a long-lived allylic-carbocation intermediate. We propose this carbocation is then quenched by water ingress into the active site due to loop movement.

The second part of the project involved the synthesis of a range of farnesyl diphosphate analogues designed to offer means of trapping carbocation intermediates in the catalysis by a range of sesquiterpene synthases. These analogues were primarily epoxy-FDPs, designed to produce simple terpenoids with hydroxy- and epoxy-functionalities. While a number of the prepared analogues were turned over by the tested sesquiterpene synthases two, 10,11-epoxy-FDP and 10-hydroxy-11-ene-FDP, were able to intercept the farnesyl cation upon loss of the diphosphate group, to generate a novel macrocyclic ether.

Acknowledgments

First and foremost I would like to thank my supervisor, Professor Rudolf Allemann, for giving me the opportunity to undertake this exciting work, and for his continued support and advice throughout my Ph.D.

Special thanks also go to Dr Veronica Gonzalez for her help and patience in teaching me all the protein production and molecular biology techniques in my first year and for her continued help and helpful discussions; and to Dr Juan Faraldos for his direction and insight, and long talks. I would also like to thank Dr Robert Mart for his help and advice in the lab and Dr Robert Jenkins and the technical support staff for their help. Further thanks go to Dr David Miller for his patience in proofreading this thesis, and his advice along the way.

I would like to thank all the members of the Allemann group who have made my studies so enjoyable, with special mentions to Will Dawson, Oscar Cascon and Sarah Adams; But also Mel, Ryan, Roger, Antonio, Marianna, Chris and everyone else.

My sincere gratitude to my parents and family, without whose support I would not be where I am today. Finally a special thank you to my wife, Charlotte, for putting up with me even when I'm tired and irritable. I could not have done it without your love and support.

I would also like to thank my examiners, Prof. Nigel Richards and Dr Rebecca Goss for making my *viva* surprisingly enjoyable, and for insightful discussions on a Friday afternoon.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

ADS	Amorphadiene synthase
AT-AS	Aristolochene synthase from *Aspergillus terreus*
ATP	Adenonsine triphosphate
9-BBN	9-Borabicyclo[3.3.1]nonane
BN-PAGE	Blue-native polyacrylamide gel electrophoresis
CD	Circular dichroism
CHES	*N*-Cyclohexyl-2-aminoethanesulfonic acid
CMP	Cytidyl monophosphate
d. H2O	Deionised water
DCM	Dichloromethane
DCS	δ-Cadinene synthase
DHFDP	2,3-Dihydro farnesyl diphosphate
DHP	Dihydropyran
DIBAL-H	Diisobutyl aluminium hydride
DMADP	Dimethylallyl diphosphate
DNA	Deoxyribonucleic acid
DOX	Deoxyxylulose phosphate
DP	Diphosphate
E. coli	*Escherichia coli*
EBFS	(*E*)-β-Farnesene synthase
FDP	Farnesyl diphosphate
GAS	Germacrene A synthase
GC-FID	Gas chromatography flame ionisation detection
GC-MS	Gas chromatography mass spectrometry
GdolS	Germacradien-4-ol synthase
GDP	Geranyl diphosphate
GDS	Germacrene d synthase
GGDP	Geranylgeranyl diphosphate
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HMG-CoA	Hydoxymethylglutaryl coenzyme A
HPLC	High pressure liquid chromatography
IDP	Isopentyl diphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KS-HcS	Hedycaryol synthase from *Kitasatospora setae*
LB	Luria-Bertani media
LDP	Linalyl diphosphate
LiHMDS	Lithium bis(trimethylsilyl)amide
*m*CPBA	3-Chloroperbenzoic acid
1.1.1 MeCN	Acetonitrile
MEP	Methylerythritol phosphate
MES	2-(*N*-Morpholino)ethanesulfonic acid
MVA	Mevalonic acid
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form

NBS	*N*-Bromosuccinimide
NDP	Nerolidyl diphosphate
NIST	National Institute of Standards and Technology
NMO	*N*-methylmorpholine-*N*-oxide
NMR	Nuclear magnetic resonance spectroscopy
NTA	Nitriolotriacetic acid
OD_{600}	Optical density (600 nm)
PCR	Polymerase chain reaction
PPTS	Pyridinium *para*-toluenesulfonate
PR-AS	Aristolochene synthase from *Penicillum roquefortii*
*p*TsOH	*para*-Toluenesulfonic acid
1.1.2 RT	1.1.3 Room temperature
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP-SdS	Selinadiene synthase from *Streptomyces pristinaespiralis*
TIC	Total ion chromatogram
TEAP	Bis(triethylammonium) hydrogen phosphate
THF	Tetrahydrofuran
THP	Tetrahydropyran
TLC	Thin layer chromatography
TPAP	Tetrapropylammonium perruthenate
TS	Trichodiene synthase
WT	Wild type

Throughout this work, amino acids are expressed as their triple or single letter code:

Amino Acid	3-letter code	1-Letter code
L-Alanine	Ala	A
L-Arginine	Arg	R
L-Asparagine	Asn	N
L-Aspartic acid	Asp	D
L-Cysteine	Cys	C
L-Glutamic acid	Glu	E
L-Glutamine	Gln	Q
L-Glycine	Gly	G
L-Histidine	His	H
L-Isoleucine	Ile	I
L-Leucine	Leu	L
L-Lysine	Lys	K
L-Methionine	Met	M
L-Phenylalanine	Phe	F
L-Proline	Pro	P
L-Serine	Ser	S
L-Threonine	Thr	T
L-Tryptophan	Trp	W
L-Tyrosine	Tyr	Y
L-Valine	Val	V

1 Introduction

1.1 Terpenes

Terpenes and terpenoids comprise one of the largest and most structurally diverse families of natural products, with known compounds numbering in the tens of thousands.[1,2] More than half of the known terpenoids are biosynthesised by plants and some, such as the phytol side-chain of chlorophyll[3] and the phytosterols[4] of the plant wall, are the products of primary metabolism, which are vital to the plant for growth and development. A greater number of terpenoids are formed by secondary metabolism and have a wide array of biological functions such as communication pheromones both within insect species[5] and from plants to insects,[6] and for phytoalexins for plant defence and pigments.

The wide variety of bioactivities found within terpenes and terpenoids means that they have found uses in therapeutics for the treatments of a range of conditions, most notably paclitaxel (Taxol™, **1**),[7] the anti-cancer agent and artemisinin (**2**), one of the most important drugs in antimalarial combination therapies,[8] Figure 1.1.

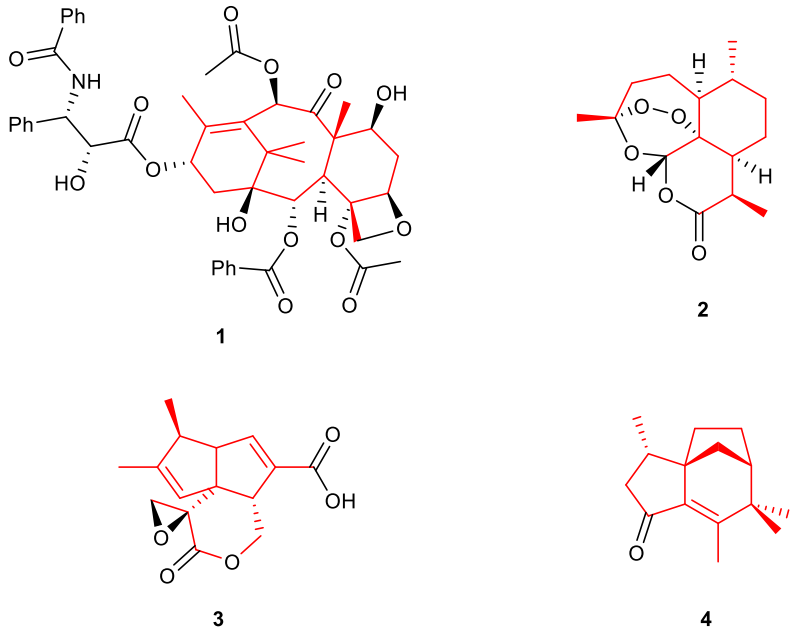


Figure 1.1 Paclitaxel (**1**), artemisinin (**2**), pentalenolactone (**3**) and albaflavenone (**4**) with the hydrocarbon terpene skeletons higlighted in red.

Terpenoids are also widely found in fungi[9] and over the last decade, through the use of genome mining, a large number of putative terpene synthase genes have been identified in bacteria, particularly in *Streptomyces*,[10,11] hinting at a wealth of unidentified terpenoids still to be discovered. Many of these putative

terpene synthase genes are within gene clusters that also contain regions encoding P450 cytochromes and other enzymes, which in a number of cases have been shown to be those responsible for the transformation of the hydrocarbon terpene product to oxidised terpenoids, such as with the *Streptomyces* antibiotics pentalenolactone[12] (**3**) and albaflavenone (**4**).[13]

1.2 Biosynthesis

All terpenes and terpenoid derivatives are ultimately derived from two isomeric five-carbon precursors, isopentyl diphosphate (IDP, **5**) and dimethylallyl diphosphate (DMADP, **6**).[14] These are the biologically activated counterparts of isoprene (**7**), which was postulated to make up all terpenes through regular; head to tail, or irregular, head to head, condensations in the 'empirical isoprene rule'. Further research led Ruzicka to postulate the 'biogenetic isoprene rule',[14] which invoked the formation of simple precursors which were then cyclised.[14] Consequently, terpenes are classified according to the number of five-carbon 'isoprene' (**5**) units they contain, Table 1.1; each class of terpenes are made from their own universal precursor formed from the sequential condensation of IDP (**5**) with the starting molecule DMADP (**6**).

Table 1.1 Classification of terpenes

Prefix-	C_x	Precursor	Examples
Hemi-	C_5	dimethylallyl diphosphate	Isoprene
Mono-	C_{10}	geranyl diphosphate	Limonene
Sesqui-	C_{15}	farnesyl diphosphate	Amorphadiene
Di-	C_{20}	geranylgeranyl diphosphate	Taxadiene
Tri-	C_{30}	(2x) farnesyl diphosphate	Steroids, cholesterol
Tetra-	C_{40}	(2x) geranylgeranyl diphosphate	Carotenoids

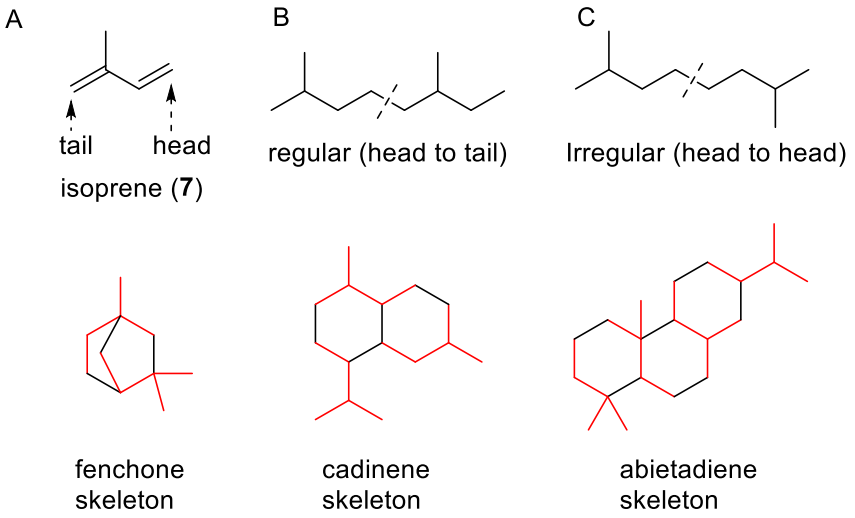
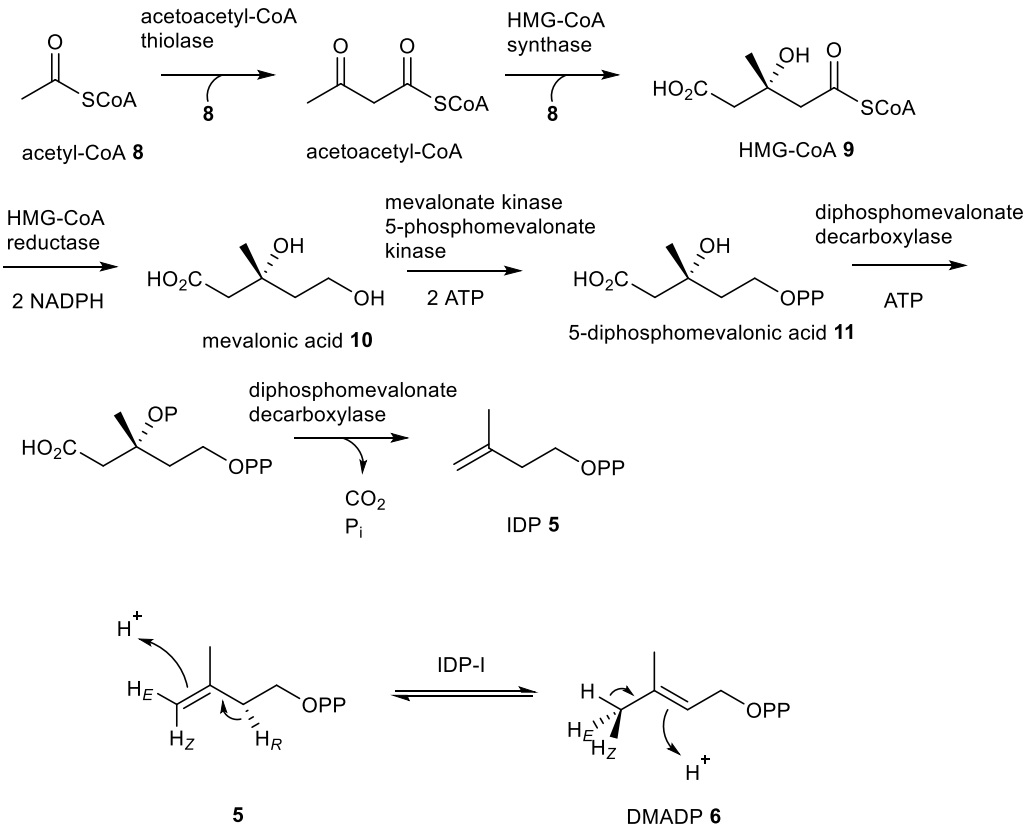


Figure 1.2 Structure of isoprene (A) and regular (B) and irregular (C) arrangements of isoprene units. Examples of mono- sesqui- and diterpene skeletons with the isoprene units highlighted in red, from Ruzicka's landmark paper.[14]

1.2.1 Mevalonate pathway

In eukaryotes such as animals and fungi, the cytosol of plants, and also in some bacteria, IDP and DMADP are biosynthesized through the mevalonate pathway (MVA). Three molecules of acetyl coenzyme A (acetyl-CoA, **8**) undergo sequential Claisen and aldol condensations to yield the key thioester, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA, **9**). HMG-CoA then undergoes NADPH-dependent reduction with two equivalents of NADPH to form mevalonic acid (**10**), the resulting primary alcohol is then diphosphorylated by two molecules of adenosine triphosphate (ATP) to 5-diphosphomevalonic acid (**11**); the subsequent phosphorylation of the 3-hydroxyl drives the decarboxylation and elimination of the tertiary phosphate group to produce IDP (**5**). IDP-isomerase (IDP-I) is responsible for the interconversion of IDP (**5**) and DMADP (**6**) to an equilibrium ratio of approximately 3:1 (DMADP:IDP), through the stereospecific removal of the 2-pro-*R*-proton and the addition of a proton to the re-face of the double-bond of IDP,[15,16] Scheme 1.1.

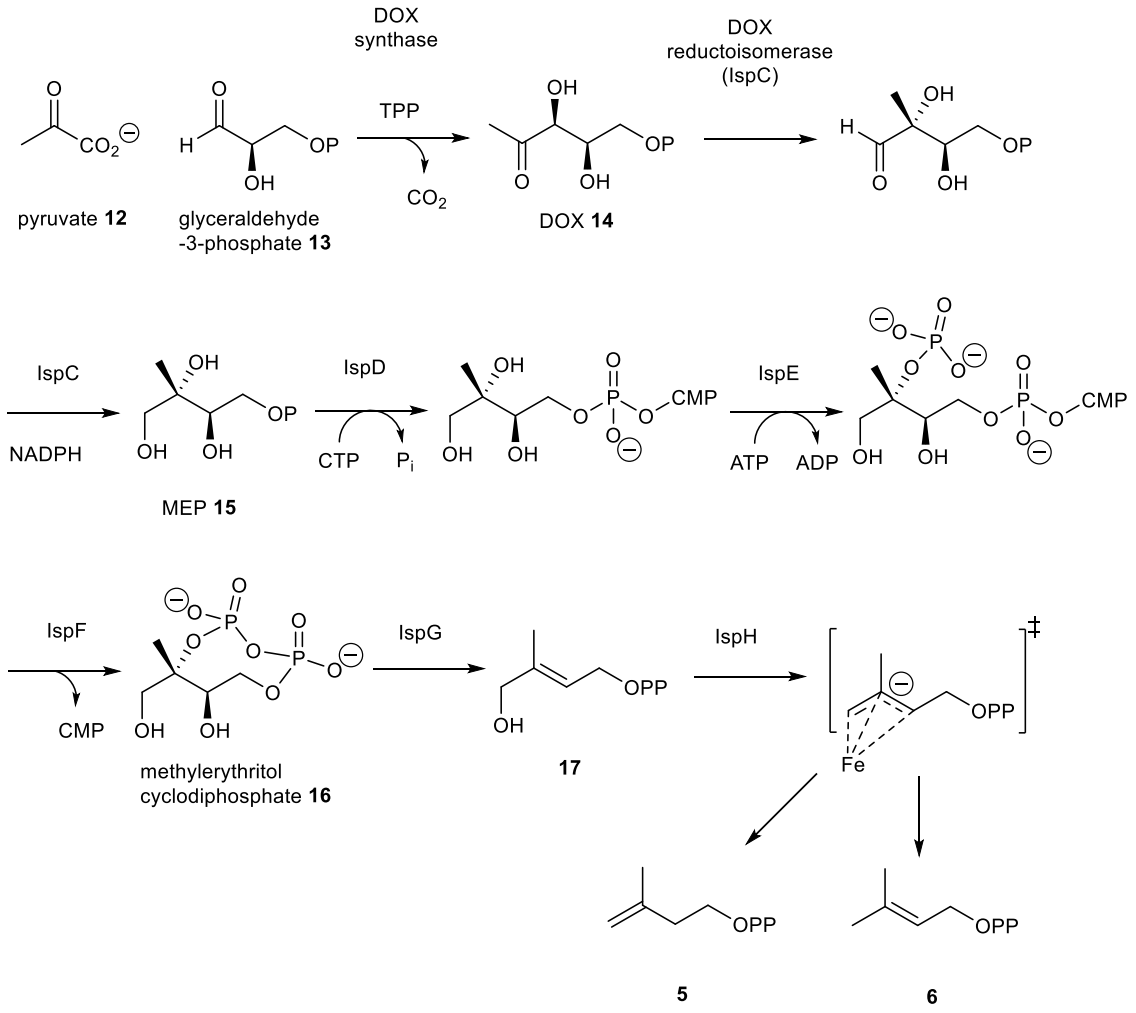


Scheme 1.1 Mevalonate pathway (top) and interconversion of IDP and DMADP (bottom).

1.2.2 Non-mevalonate pathway

In bacteria and the chloroplasts of algae and plants, IDP and DMADP are biosynthesised by the non-mevalonate pathway, also called the deoxyxylulose phosphate (DOX) or methylerythritol phosphate (MEP) pathway after on-path intermediates.[17,18] In this pathway, pyruvate (**12**) and glyceraldehyde-3-phosphate (**13**) are condensed *via* the thiamine diphosphate (TDP) dependent decarboxylation of pyruvate forming 1-deoxy-D-xylulose-5-phosphate (deoxyxylulose phosphate, DOX, **14**). This then undergoes a rearrangement, followed by the NADPH-dependent reduction of the resulting aldehyde to yield 2-*C*-methyl-D-erythritol-4-phosphate (methylerythritol phosphate, MEP, **15**). Methylerythritol phosphate is converted into a cyclic diphosphate, 2-*C*-methyl-D-erythritol-2,4-cyclodiphosphate (**16**), in a series of steps involving the transfer of cytidyl phosphate onto the terminal phosphate, phosphorylation of the 2-hydroxyl and the elimination of cytidyl monophosphate (CMP) upon cyclisation.[19] The final two steps are catalysed by two iron-sulfur proteins, of which the catalytic mechanisms are still in dispute.[20] The first step is a reductive ring opening of the cyclic diphosphate and elimination of water to yield 1-hydroxy-2-methyl-2-(*E*)-butenyl diphosphate (**17**). The following transformation of 1-hydroxy-2-methyl-2-(*E*)-butenyl diphosphate to both IDP

(**5**) and DMADP (**6**) is believed to involve two stepwise single-electron transfers from an iron-sulfur cluster and heterolytic cleavage of the 1-hydroxy group to yield an allylic carbanion which can be deprotonated to yield either IDP or DMADP.[21]

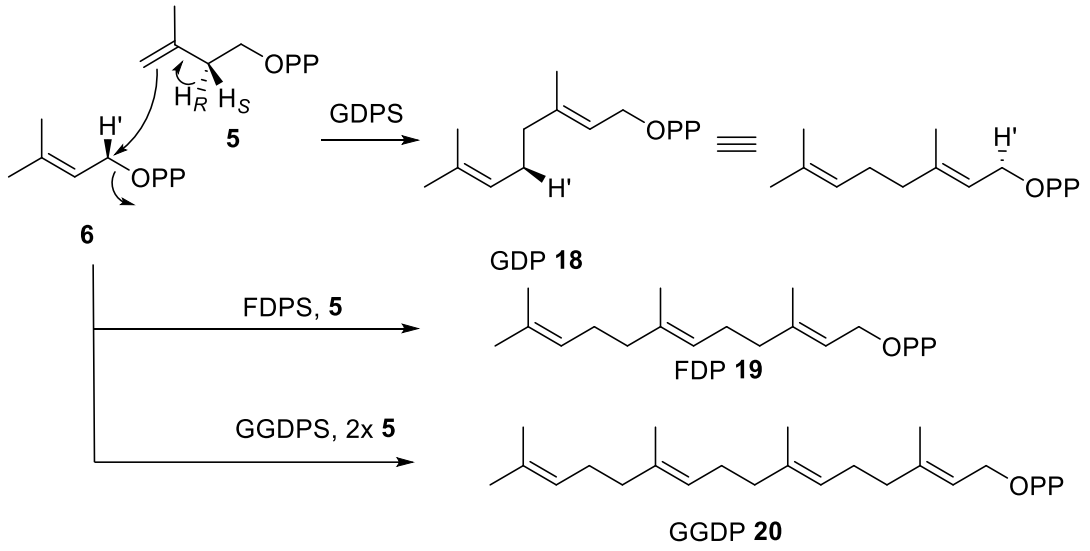


Scheme 1.2 Non-mevalonate pathway to IDP and DMADP

1.2.3 Isoprenyl diphosphate elongation

The linear terpenes, geranyl-, farnesyl- and geranylgeranyl diphosphate (GDP, C_{10}, **18**; FDP, C_{15}, **19** and GGDP, C_{20}, **20** respectively) are formed from the two isoprenoid diphosphate precursors by the subsequent head-to-tail additions of an IDP molecule (**5**) to a starting molecule of DMADP (**6**). These chain elongations proceed by attack of the terminal double bond of IDP at C1 with an inversion of configuration. Stereoselective removal of the 2-*pro-R*-hydrogen of the added IDP quenches the intermediate carbocation and completes formation of the elongated isoprenyl chain.[20] In the case of geranyl diphosphate synthase (GDPS) the newly formed GDP is released from the active site, while for farnesyl diphosphate synthase (FDPS) and geranylgeranyl diphosphate

synthase (GGDPS) subsequent elongations with IDP occur in the same manner to form farnesyl- (**19**, C_{15}) and geranylgeranyl- (**20**, C_{20}) diphosphates respectively, Scheme 1.3.



Scheme 1.3 Isoprenyl diphosphate elongation reactions to form GDP (**18**), FDP (**19**) and GGDP (**20**)

These linear isoprenyl diphosphates undergo carbocationic cyclisations and rearrangements catalysed by terpene synthases to generate a structurally diverse range of cyclic hydrocarbon skeletons, which are then available to undergo P450-dependent oxidations and further transformations to generate complex bioactive terpenoid products. This study focuses on the cyclisations of FDP by sesquiterpene synthases.

1.3 Sesquiterpene synthases

Sesquiterpene synthases are a class of terpene synthase that catalyse the conversion of the 15-carbon isoprenyl diphosphate, FDP, into more than 300 different sesquiterpenes with a diverse range of structures including linear, mono-, bi- and tricyclic examples, with high degrees of stereospecificity.

1.3.1 Structure

While there is typically very low sequence similarity between terpene synthases, they have a highly conserved tertiary structure that was first seen in avian farnesyl diphosphate synthase.[22] This conserved structure is termed the class I terpene synthase fold and is comprised of 10 α-helices joined by short inter-connecting loop regions, with the active site formed by a conical cleft between three main central helices. The loop regions are typically very short, 3 – 5 residues, distal to the active site, while at the end of the helices

proximal to the active site they are longer, 8 – 10 residues in length and more flexible during the catalytic cycle.[23]

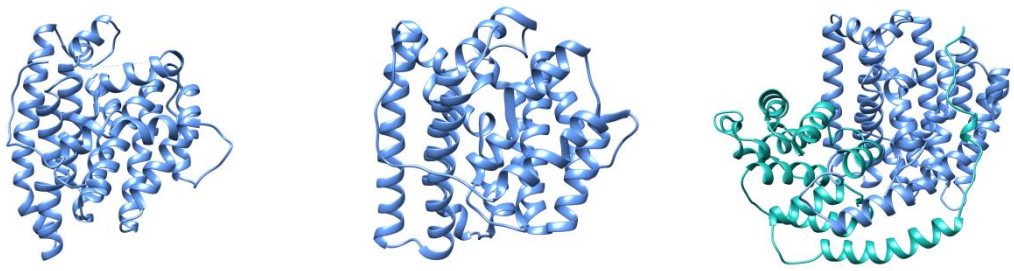


Figure 1.3 Cartoon represtenations of bacterial (pentalenene synthase from S*treptomyces* UC5319, PDB 1PS1) fungal (aristolochene from *Aspergillus terreus*, PDB 2OA6) and plant (5-epi-aristolochene synthase from *Nicotania tabacum*, PDB 5EAT) sesquiterpene synthases, showing the class I fold in blue and, in the plant synthase, the N-terminal domain in turquoise.

Terpene synthases in bacteria and fungi are single domain proteins with this class I fold which is typically 300 amino acids long approximately, while those from plants also have a second N-terminal domain, which is also α-helical in structure.[24] In some plant terpene synthases this second domain appears to be catalytically silent, while it has been shown that in others the N-terminal domain is important, covering a range of roles; domain-changing experiments suggest it may play an important role in ensuring the correct folding of the C-terminal domain,[25,26] and recent work within the Allemann group suggests that the N-terminal segment is important in desolvating and protecting the active site. Shortening of the N-terminal loop segment in δ-cadinene synthase from *Gossypium arboreum* (DCS) by 10, 20 and 30 amino acid residues leads to increasing proportions of a product resulting from premature quenching of a carbocation intermediate.[27]

1.3.2 Metal-binding regions

All terpene synthases share two consensus metal binding regions responsible for the binding of a tri-nuclear magnesium cluster close to the active site entrance. Typically these regions comprise an aspartate-rich region, **DD**xx**D**/**E** found on helix D and the NSE/DTE motif (L,V)(V,L,A)(**N,D**)D(L,I,V)x(**S,T**)xxx**E** on helix H.[24] The DTE-motif is commonly found in plant terpene synthases while the NSE-variant is more common in fungi and bacteria. Altered metal-binding regions have been identified in some cases such as δ-cadinene synthase from *Gossypium arboreum* (DCS) in which the second metal-binding motif is a second DDxxD motif,[28] more commonly seen in isoprenyl diphosphate synthases; and in (*R*)-germacrene D synthase from *Solidago canadensis* the aspartate rich region takes the form of NDxxD.[29] Nonetheless

the two metal-binding regions are responsible for the binding of three magnesium ions, which in turn coordinate to the diphosphate group of farnesyl diphosphate and settle the isoprenoid chain in the active site.

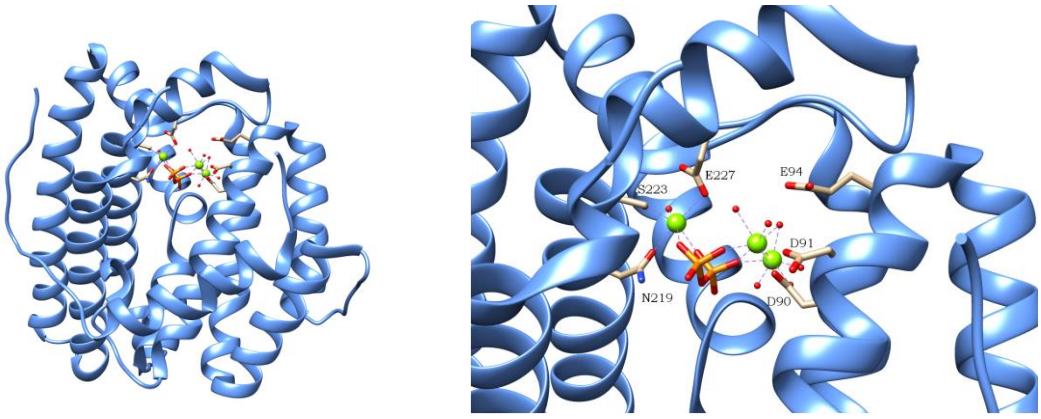


Figure 1.4 Cartoon representation of the crystal structure of AT-AS (PDB 2OA6 chain D) in coordination with $[Mg^{2+}]_3$-PP_i. Magnesium ions are shown as green balls, the PP_i is shown as a stick structure where red denotes oxygen and orange denotes phosphorus. Water in the magnesium coordination sphere is also shown as small red balls. Metal-binding residues are labelled.

The use of X-ray crystal structures is a powerful tool for structural analysis and those of a number of terpene synthases have shown that the aspartate-rich region is responsible for binding two magnesium ions while the NSE/DTE-motif binds to the third magnesium.[22,30] Magnesium ions (or another divalent metal ion) have been shown to been critical for catalysis by terpene synthases, not only are they vital to bind the diphosphate group and orientate the substrate correctly but they also serve to initiate the catalytic cycle by triggering ionisation of the diphosphate group for initial carbocation formation.

1.3.3 Substrate binding cycle

While the X-ray crystal structures of a number of sesquiterpene synthases have been solved,[23,28,30–38] studies with aristolochene synthase from *Aspergillus terreus* (AT-AS) first revealed the closed conformation while complexed with diphosphate and three magnesium ions.[30] Further X-ray crystallographic studies with this enzyme and fluorinated-FDPs presented a series of 12 independent structures displaying an array of binding interactions; while predominantly examples of the substrate binding in the open conformation in the absence of magnesium ions three examples were obtained with increasing numbers of magnesium ions bound and in varying conformations. Using these structures Shishova *et al.* proposed a model for substrate and metal binding, and the associated conformational changes necessary for the catalytic cycle to progress.[37] Initial binding of the substrate is governed by molecular recognition

of the diphosphate by hydrogen bond donors in the active site; the binding of Mg^{2+}_B follows or possibly occurs concurrently, coordinating to the largely pre-organised binding site and causing a reorientation of the diphosphate group. Binding of Mg^{2+}_C to the first aspartate in the DDxxD motif and to the diphosphate group causes greater conformational changes triggering active site closure, and binding of the third magnesium, Mg^{2+}_A, completes the magnesium cluster causing a shift in Mg^{2+}_C such that both Mg^{2+}_A and Mg^{2+}_C are coordinated to the first aspartate in the DDxxD motif and completing the conformational changes required for full active site closure and formation of the Michaelis-complex. Opening of the active site and product release upon completion of the catalytic cycle is proposed to occur in the reverse sequence. Molecular dynamics simulations performed on AT-AS by Van der Kamp *et al.*[39] strongly support the proposed binding cycle and also reveal further details; The apo-enzyme can adopt a range of conformations from open to near-closed, but the binding of FDP (**19**) and Mg^{2+}_B shifts the equilibrium to 'distinctly open conformations', inviting the further binding of the magnesium ions.

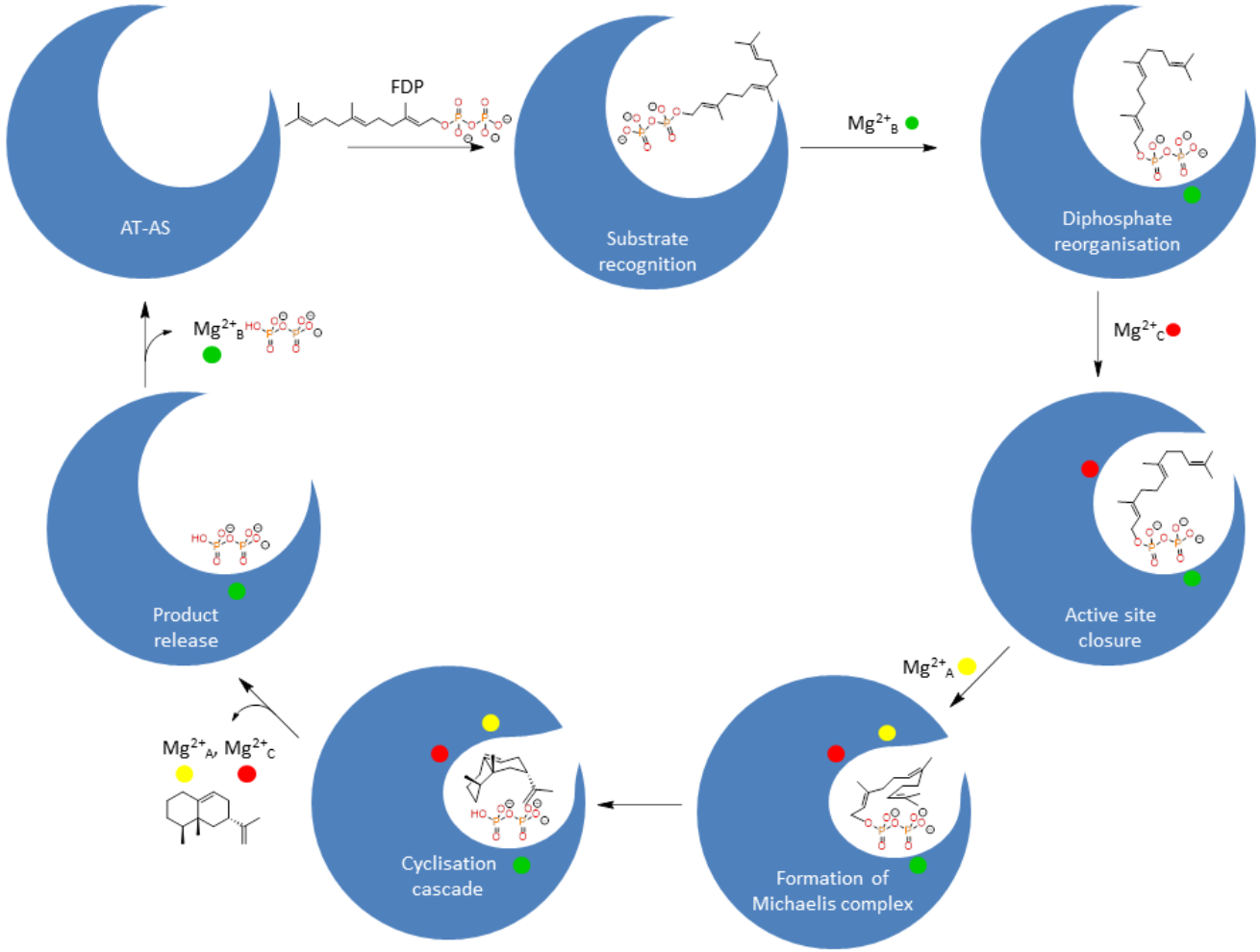


Figure 1.5 Model of metal-binding and conformational changes of AT-AS

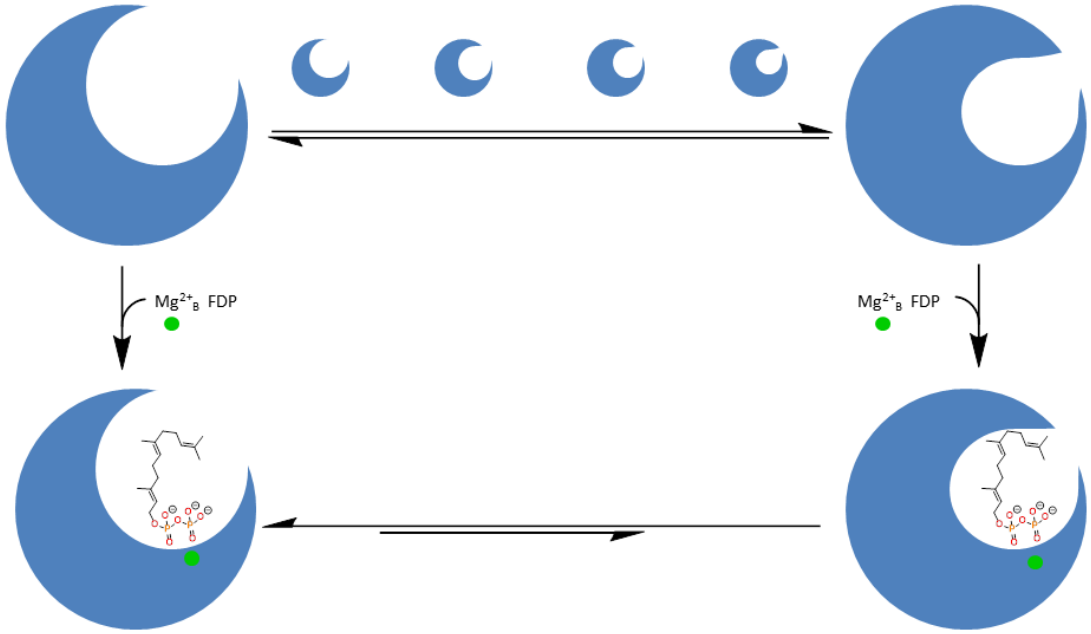


Figure 1.6 Model of the conformational effect of the binding of FDP and Mg^{2+}_B on AT-AS

A recent study by Baer *et al.* has identified significant structural shifts in a bacterial terpene synthase, selina-4(15),7(11)-diene synthase from *Streptomyces pristinaespiralis* (SP-SdS), hinting at an induced-fit mechanism in active-site closure directed by a novel effector triad of a sensor, linker and effector.[36] X-ray crystal structures of SP-SdS in the apo-form as well as in coordination with $(Mg^{2+})_3$-PP_i and in coordination with $(Mg^{2+})_3$-DHFDP (2,3-dihydro-FDP, a nonreactive substrate analogue) were used to identify distinct states of the G1/2-helix break motif between the open and closed conformations. The data also show a significant shift of arginine-178 towards the diphosphate ligand upon substrate binding, while concomitantly reorienting aspartate-181 to form a hydrogen bond with arginine-178. This strictly conserved arginine is part of a hydrogen bonding network[30,31,34] in the active site and to be vital for the catalytic activity in PR-AS.[40] In the study with SP-SdS, the reorientation of arinine-178 and aspartate-181 precede the conformational change of the G1/2-helix break motif which 'moves the catalytic key carbonyl oxygen of glycine-182 inwards into the active site'.[36] The carbonyl oxygen of the residue at the break position in the G helix (glycine-182 in SP-SdS) has been implicated in the stabilisation of the developing carbocation at C1 in the ionisation of FDP by hedycaryol synthase from *Kitasatospora setae* (KS-HcS).[35]

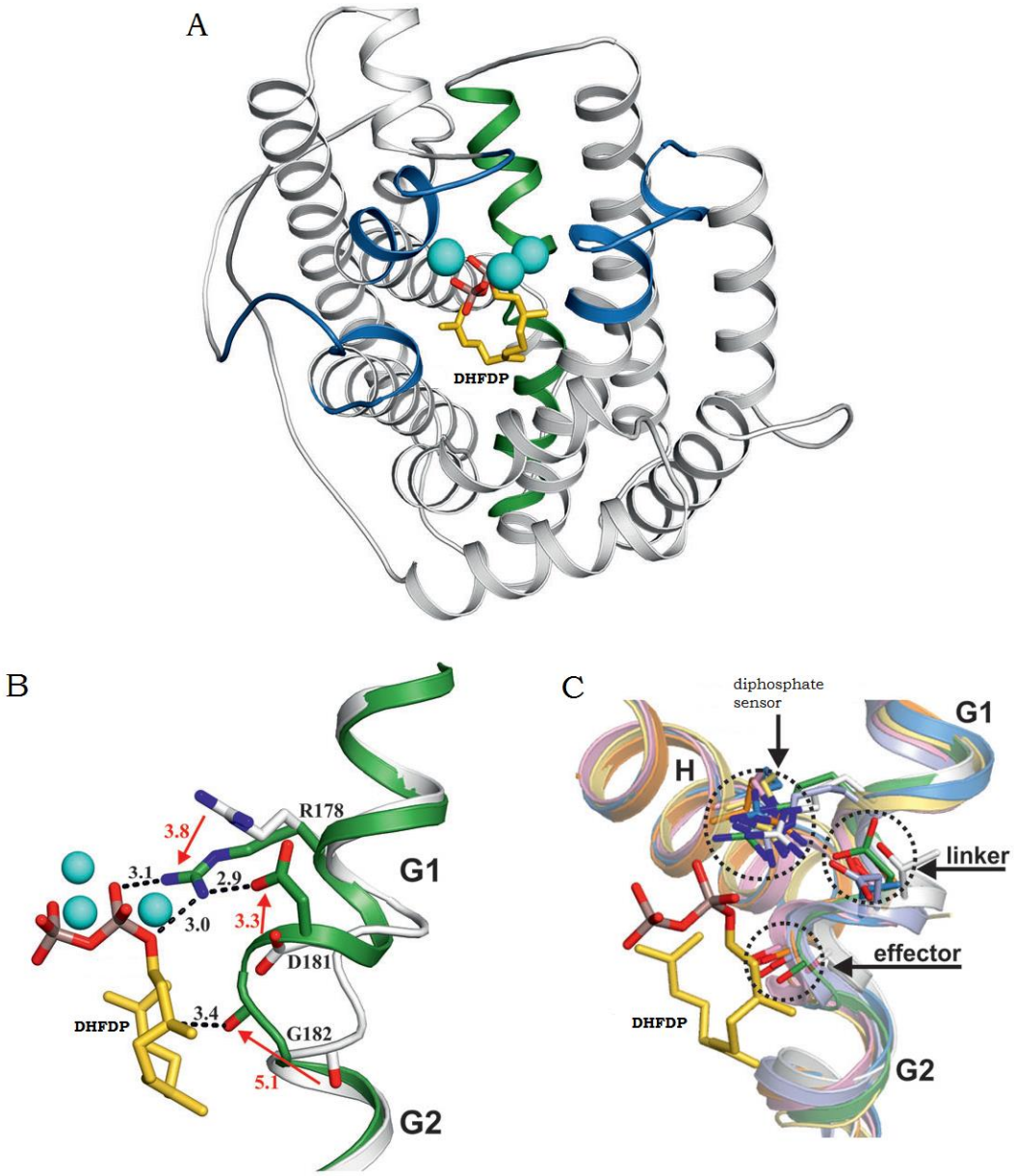
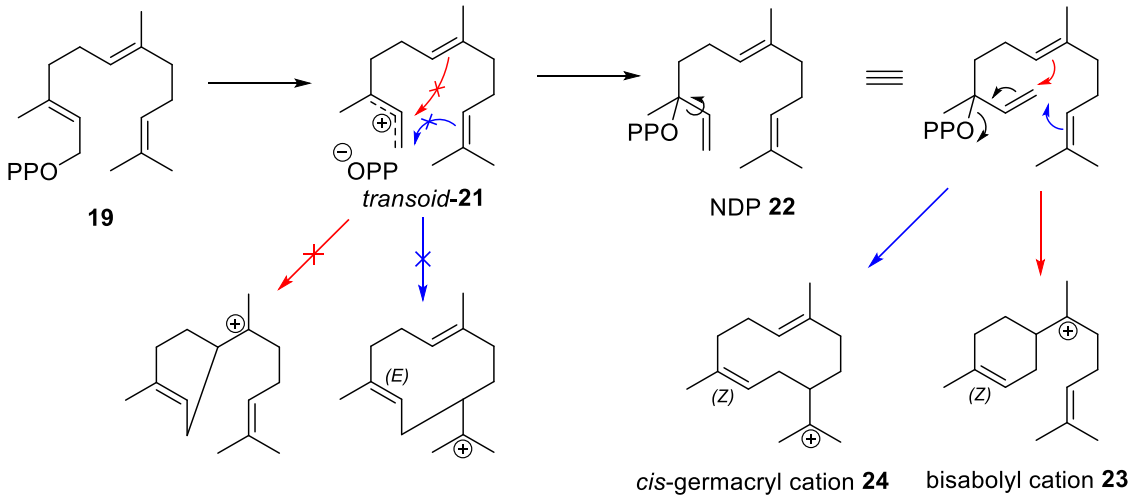


Figure 1.7 Cartoon representations of an X-ray cocrystal structure of SP-SdS in complex with DHFDP, A; Helix G1/2 shown in green, Mg^{2+} in cyan, DHFDP in yellow, oxygen atoms in red, Blue ribbons highlight structural elements formed on DHFDP binding. Superimposition of helix G1/2 in open (grey) and closed (green) conformations, B; dashed lines highlight interactions while the red arrows indicate the induced-fit movements. Superimpositions of the effector triad (sensor, linker, effector as indicated), C; SP-SdS:DHFDP (green), epi-isozizaene synthase (grey, 3KB9), aristolochene synthase (pale blue, 4KUX), bornyl diphosphate synthase (pink, 1N20), limonene synthase (orange, 2ONG), 5-epi-aristolochene synthase (blue, 5EAT) and taxadiene synthase (yellow, 3P5R). Figures taken from reference[36].

1.3.4 Cyclisation through nerolidyl diphosphate (NDP)

For sesquiterpene synthases the first chemical step in the catalytic cycle is the ionisation of FDP (**19**) to the farnesyl allylic cation-diphosphate anion pair followed by nucleophilic attack by a double-bond. Attack by the C10,C11-double bond to form either the 10- or 11-membered rings retaining the (*E*,*E*)-double bond stereochemistry of the FDP is possible through the *transoid*-allylic

farnesyl cation (*transoid*-**21**). However, for those sesquiterpene synthases which undergo a 1,6-cyclisation or form macrocycles with (*Z*,*E*)-double bond stereochemistry, nucleophilic attack by the C6,C7- or C10,C11-double bonds are geometrically forbidden.[41,42] In this case, collapse of the farnesyl allylic cation-diphosphate anion pair has been proposed to form the tertiary diphosphate, nerolidyl diphosphate (NDP, **22**), in the transoid conformation. Rotation around the 2,3-single bond yields the cisoid conformation which is now suitably positioned for 1,6- or 1,10-ring closure either directly with concerted loss of the diphosphate or indirectly through initial formation of the *cisoid*-allylic cation-diphosphate anion pair followed by nucleophilic attack of a double-bond to form the bisabolyl (**23**) or the *cis*-germacryl cation (**24**) which can then participate in further transformations.



Scheme 1.4 Cyclisation of FDP (**19**) to cations containing *Z*-bouble bonds *via* NDP (**22**)

A recent study by Zhang and Tiefenbacher has suggested that a direct isomerisation of the *transoid*-geranyl cation (*transoid*-**25**) to the *cisoid*-isomer (cisoid-**25**) is feasible.[43] The authors investigated the possibility of mimicking the aromatic, hydrophobic active site of terpene synthases using a self-assembling supramolecular structure that forms a large hydrophobic cavity. A resorcinarene unit (**26**) has previously been shown to form a hexamer in apolar solvents, which is known to complex to cationic guests and has been shown to be a mild Brønsted acid, Figure 1.8. This capsule was used to investigate the cyclisation of the monoterpene alcohols geraniol (**27**), nerol (**28**) and linalool (**29**) and was found to successfully catalyse the turnover of all three with the primary products demonstrating interception of carbocationic intermediates by the cleaved leaving group.

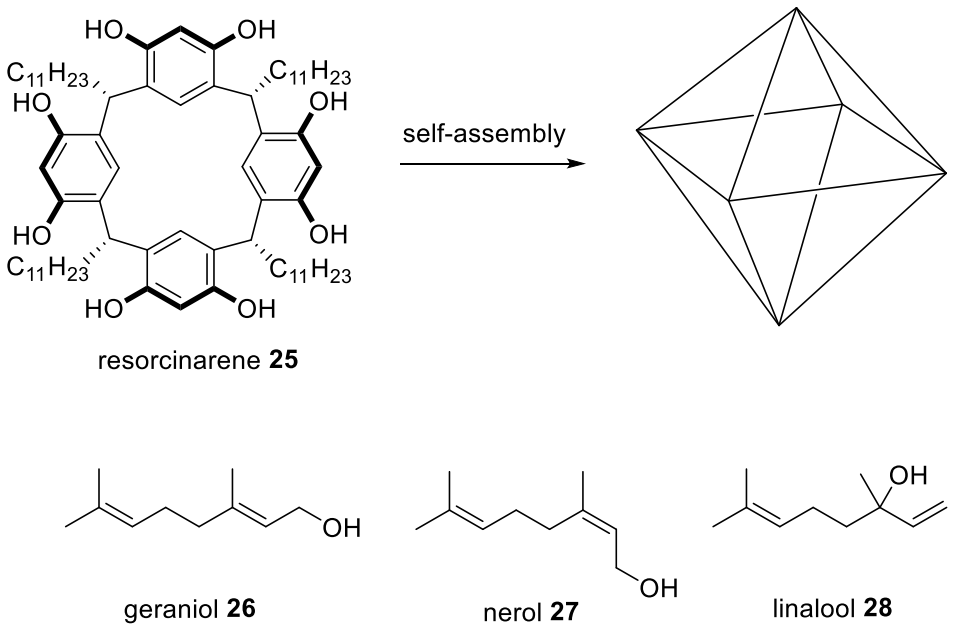
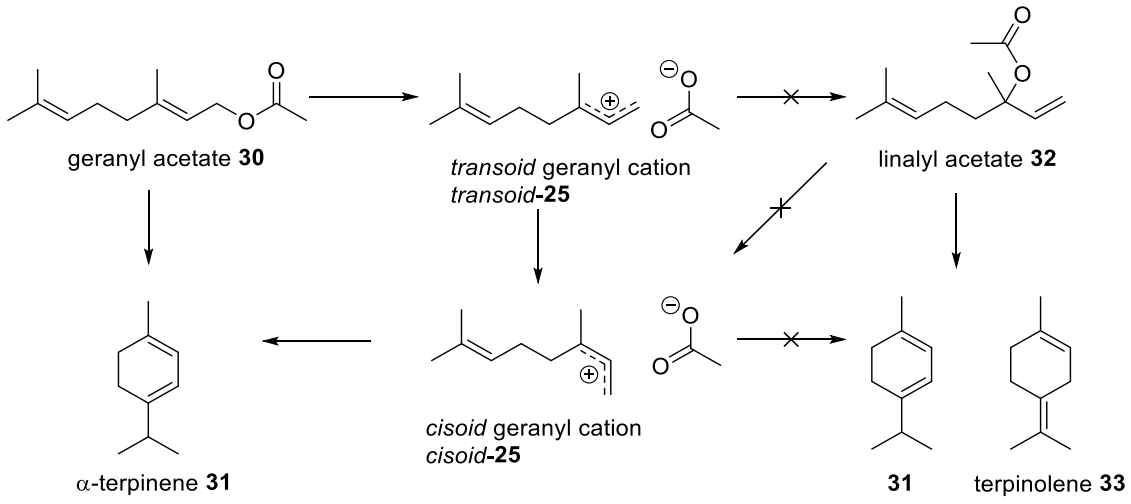


Figure 1.8 Structure of resorcinarene (**25**) and monoterpene alcohols

In order to reduce leaving group interception of intermediates the substrates were switched to the corresponding acetates. When the same reaction conditions were used with geranyl acetate (**30**) it resulted in the selective production of α-terpinene (**31**), while linalyl acetate (**32**) produced both α-terpinene (**31**) and terpinolene (**33**). If cyclisation of geranyl acetate (**30**) forms the *cisoid*-allylic cation (*cisoid*-**25**) *via* linalyl acetate (**32**) intermediate, the product distribution from both these substrates could be expected to be identical; suggesting that there is no conversion of geranyl acetate to linalyl acetate and therefore the energy barrier to isomerisation between the *transoid*- and *cisoid*-cations is easily overcome at ambient conditions, Scheme 1.5.



Scheme 1.5 Products formed by resorcinarene capsule and monoterpene acetates

1.4 Mechanistic investigations

Catalysis by sesquiterpene synthases is initiated by the loss of diphosphate to form the initial farnesyl cation. Sesquiterpene synthases are then able to direct the reaction through cyclisations, hydride- and methyl-shifts, protonations and deprotonations to the final product, shuttling the carbocation through intermediates with an impressive degree of precision and selectivity. While there are examples of 'promiscuous' sesquiterpene synthases (which produce multiple products) such as γ-humulene synthase found in both plants[44,45] and microbes,[46] many sesquiterpenes show exceptionally high-fidelity; producing a single product exclusively or near exclusively. These high-fidelity synthases are able to differentiate between alternate reaction pathways which would normally be of similar energies, achieved through the templating effect of the active site with the metal-binding motives holding the diphosphate group in place and the, predominantly hydrophobic and aromatic, residues lining the active site cavity aligning the isoprenoid-tail into a productive orientation, and stabilising carbocationic intermediates and transition states.

Many of the methods used experimentally to study the mechanisms of terpene synthases focus on probing the carbocationic intermediates; the use of aza-analogues, to mimic carbocation intermediates with ammonium species; fluorinated substrate-analogues, to destabilise intermediate carbocations and intercept the reaction pathway; and single-point mutations of aromatic residues. These methods can also be used to investigate specific steps in the reaction cascade, substrate analogues stereospecifically labelled with deuterium can be used to elucidate regio- and/or stereochemistry of specific steps; and the used of site-directed mutagenesis can be used to target specific residues for single-point mutation to probe their function and effect on catalytic activity or product distribution.

1.4.1 Aza-analogues

Aza-analogues are nitrogen-containing analogues of proposed carbocation intermediates that are either quaternised or protonated in solution to form a stable charged carbocation mimic. Numerous example of these type of products have been shown to effectively mimic intermediates, acting as inhibitors in kinetic studies with sesquiterpene synthases,[47–49] diterpene synthases[50–54] and squalene synthase.[55–57] Some examples have also proved effective as crystallisation aids in mono-[58] and sesquiterpene synthases[38,59] for X-ray crystal structure analysis.

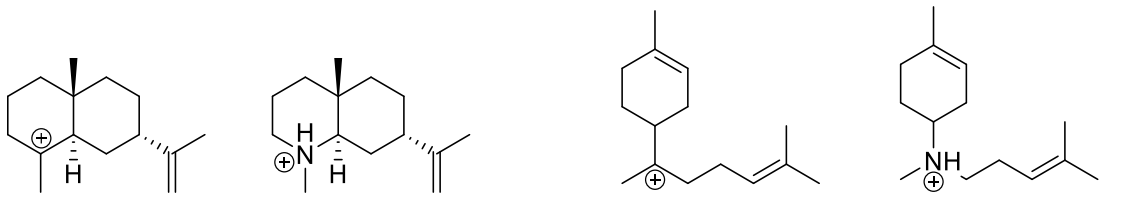


Figure 1.9 Examples of aza-analogues and the proposed carbocations they are designed to investigate

1.4.2 Substrate analogues

Isotopologues

Much of the stereo- and regiochemical detail of the mechanisms of several terpene synthases has been elucidated through the use of isotopically labelled substrates such as deuterated, tritiated and carbon-14 labelled FDPs.[60–66]

Fluorinated analogues

Replacement of hydrogen with fluorine at strategic positions in FDP has been shown to be a valuable tool in the mechanistic investigation of sesquiterpene synthases. Due to the similarity in size to hydrogen, replacement with fluorine would not be expected to have a significant effect on the size and shape of the FDP analogue while exhibiting a strong electronic influence at the site of replacement due to the high electronegativity. This electronegativity has a destabilising effect on carbocations at the β- and γ-positions *via* a through bond inductive effect, while a carbocation in an α-position is stabilised by π-donation of a lone-pair. By strategic positioning of fluorine in FDP-analogues the existence of carbocationic intermediates can be probed through the derailment of the cyclisation cascade; this strategy has been used successfully in the mechanistic elucidation of PR-AS[67–69] as described earlier, Section 1.5.5, and fluorinated-FDP's have also proven successful in the solving of crystal structures.[28,37]

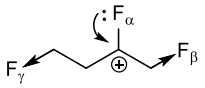


Figure 1.10 Effect of fluorine substitution α-, β- and γ- to a carbocation

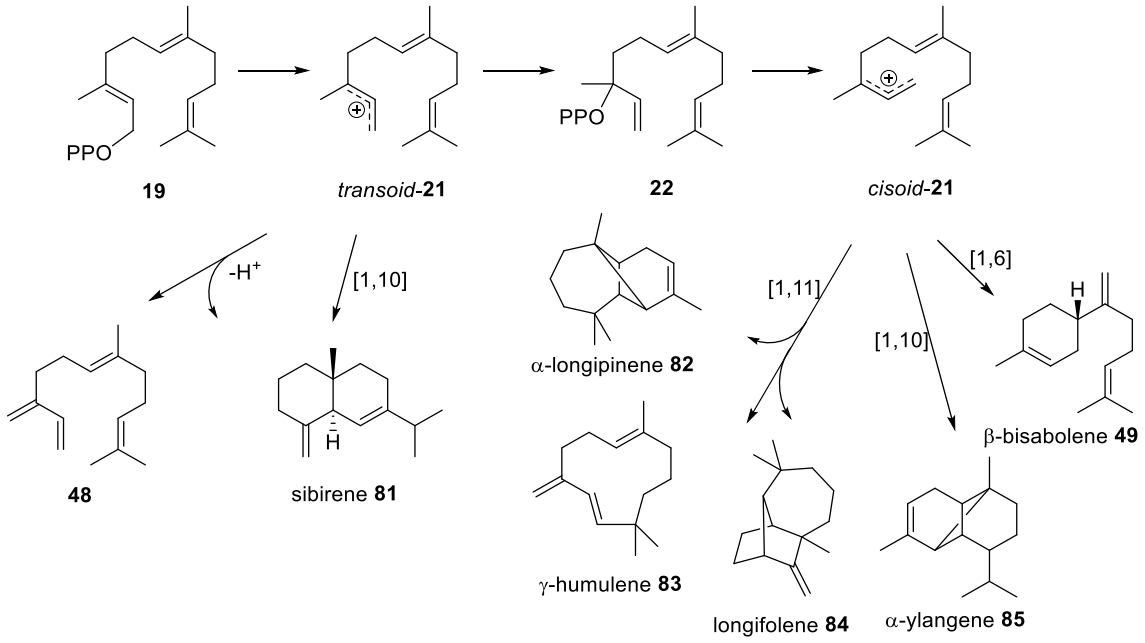
1.4.3 Site-directed mutagenesis

Site-directed mutagenesis (SDM) can be used to specifically alter targeted amino acids in the protein, enabling the importance and function of specific amino acids during catalysis to be assessed through analysis of kinetic

parameters and product distribution. This technique has been used to investigate the metal-binding motifs,[70–72] active-site aromatic residues[73,74] and putative active-site acid/bases,[75] in a number of sesquiterpene synthases; such as PR-AS, Section 1.5.5.

Plasticity residues

Using the promiscuous γ-humulene synthase, a sesquiterpene synthase from *Abies grandis*, grand fir, which produces 52 different sesquiterpenes from incubation with FDP, Yoshikuni *et al.* used saturation mutagenesis of the 19 residues forming the hydrophobic core of the active site to investigate the contribution made by each to different reaction mechanisms; initial 1,6- 1,10- or 1,11-cyclisations, .[76]



Scheme 1.6 Examples of sesquiterpenes produced by γ-humulene synthase via 1,6-, 1,10- and 1,11-cyclisation mechanisms.

The altered product distributions were profiled and normalised to the wildtype enzyme; while many of these residues were 'plastic', causing changes in product profile, four key 'plasticity residues' were identified; W315, M447, S484 and Y565. These plasticity residues significantly affected catalysis, increasing the relative selectivity of the resulting mutants by 100- to 1000-fold (proportions of one product relative to another). Based on the product profiles from saturation mutagenesis, mutations were systematically recombined to produce a desired enzyme function while not severely compromising the efficiencies of those enzymes, typically a maximal 10-fold decrease in catalytic efficiency (k_{cat}/K_M). Using this method the authors were able to create β-

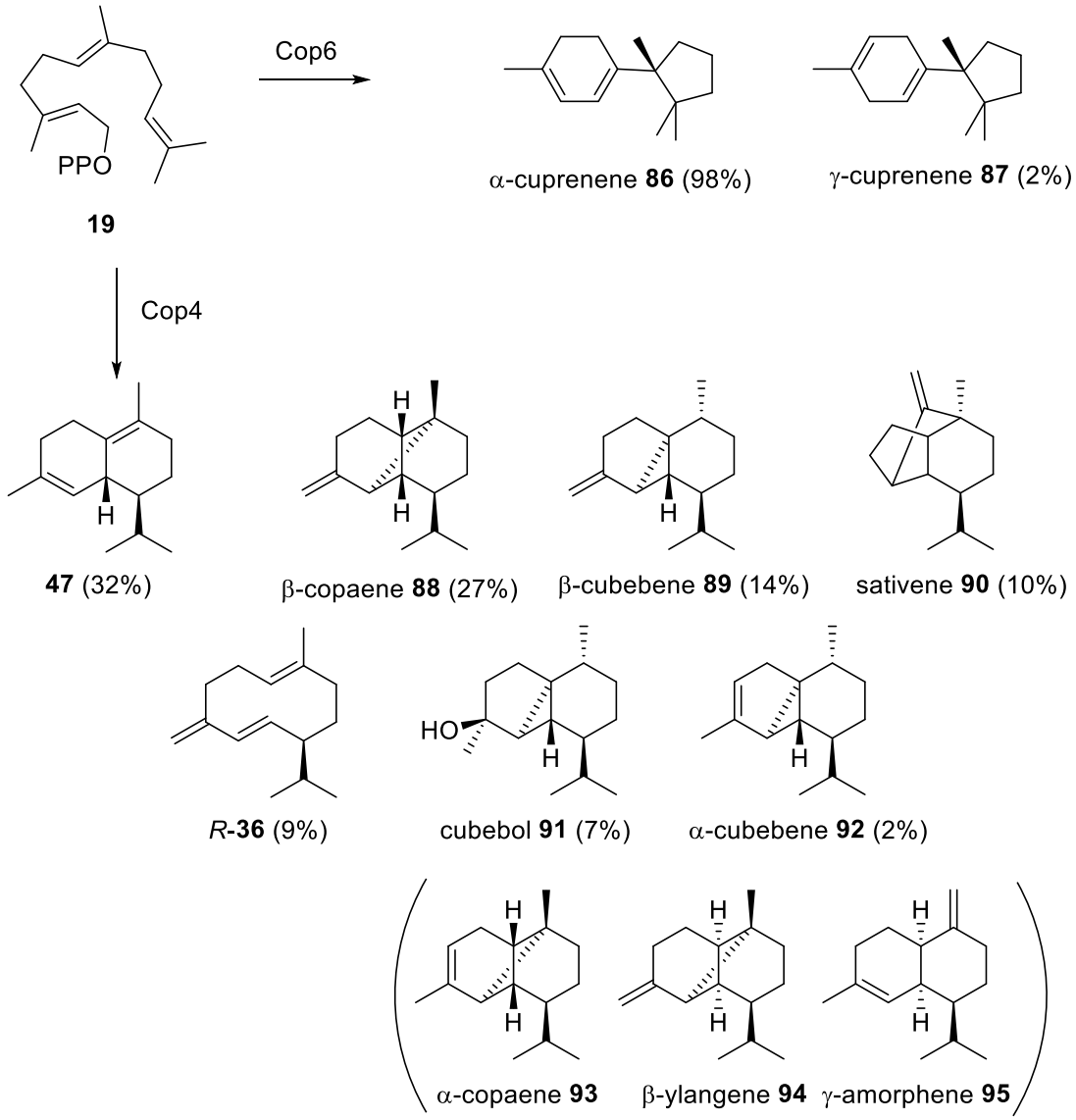
bisabolene (BBA), sibirene (SIB), longifolene (LFN) synthases and a γ-humulene synthase with enhanced selectivity (HUM), by combining up to 5-mutations, . Through this study Yoshikuni *et al.* have demonstrated the innate evolvability of terpene synthases, and an insight into the evolutionary history through a small number of mutations to 'plasticity residues, and provided evidence that a rational design approach to generating novel terpene synthases may be possible in the future.

H-1α loop

A study by Lopez-Gallego *et al.* investigated the importance of the H-1α loop in the cyclisation mechanism of a series of sesquiterpene synthases from *Coprinus cinereus*.[77] Using structural models of Cop4, a non-selective synthase producing 7 products; and Cop6, a highly-selective synthase producing α-cuprenene (**86**), ; the authors postulated interactions between the H-1α loop and other residues lining the active site entrance. Site-directed mutagenesis on residues comprising the H-1α loop identified residues which dramatically affected the product selectivity in Cop4, and severely reduced catalytic efficiency. Interestingly, similar mutations to the highly selective Cop6 had a lesser effect on both catalytic efficiency and product distribution. Mutations to the H-1α loop in Cop6 caused only a small decrease in product specificity, 98% α-cuprenene (**86**) production to 96% in the most adversely affected mutants. A loop switch from Cop4 to Cop6 had only a minor effect on Cop6 (Cop6L4 mutant), with a slight reduction in specificity matching that observed in the single point mutations. The reverse switch, insertion of the Cop6 loop in Cop4, drastically improved the selectivity of the resulting mutant (Cop4L6), producing 71% of germacrene D (**36**) as the major product in addition to two other products; through causing the premature deprotonation during the catalytic cycle to yield primarily the monocyclic product, .

Table 1.2 Product distributions (%) and steady-state kinetic parameters of g-humulene synthase and mutants. ND, not detected. (*E*)-β-farnesene (**48**), sibirene, (**81**), γ-humulene (**83**), longifolene (**84**), α-longipinene (**82**), α-ylangene (**85**), β-bisabolene (**49**).

Enzyme	Mutations	Product distribution							k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)
		48	**81**	**83**	**84**	**82**	**85**	**49**			
WT	-	3.0	23.1	45.1	13.4	4.7	3.8	6.9	0.024	4.7	51
BBA	A336V, M447H, I562T	6.4	0.4	3.8	4.7	1.1	0.1	83.6	0.022	2.9	78
SIB	F312Q, M339A, M447F	6.4	78.1	11.4	2.6	0.1	1.0	0.4	0.00046	3.0	0.15
LFN	A317N, A336S, S484C, I562V	1.7	1.4	12.6	63.0	12.6	3.2	5.5	0.070	3.8	181
HUM	M339N, S484C, M565I	2.7	0.4	85.7	3.4	3.5	4.3	ND	0.0018	2.1	0.87



Scheme 1.7 Products produced from cyclisation of FDP by Cop6 and Cop4 and their proportions. Sesquiterpenes in brackets are those produced by Cop4 mutants.

Table 1.3 Product distributions (%) from Cop6 and Cop4 and their H-1α loop mutants, from the headspace of *E. coli* cultures. U, unidentified sesquiterpenes.

Enzyme	Product distribution												
	86	**87**	**92**	**89**	**90**	**88**	**36**	**47**	**91**	**93**	**94**	**95**	U
Cop6 WT	98	2	-	-	-	-	-	-	-	-	-	-	-
Cop6-C236A	96	3	-	-	-	-	-	-	-	-	-	-	1
Cop6-E237L	96	3	-	-	-	-	-	-	-	-	-	-	1
Cop6-N240L	98	2	-	-	-	-	-	-	-	-	-	-	-
Cop4 WT	-	-	2	14	10	27	9	32	7	-	-	-	-
Cop4-K233I	-	-	10	27	5	18	12	17	3	2	2	-	3
Cop4-H235P	-	-	-	8	-	11	49	4	-	-	15	5	-
Cop4-T236L	-	-	-	12	14	32	-	42	-	-	-	-	-
Cop4-N238L	-	-	-	21	9	32	4	34	-	-	-	-	-
Cop4-N239L	-	-	-	4	-	10	50	3	-	-	15	5	13
Cop6L4	96	3	-	-	-	-	-	-	-	-	-	-	1
Cop4L6	-	-	-	-	-	11	71	-	-	-	18	-	-

Notably, in Cop4 the product fidelity was also found to be sensitive to the pH, with increased production of germacrene D at pH 5 and pH 10; this pH dependence is eliminated in the H235 and N239 mutants, pointing to ionic interactions between these residues and the active site as key in stabilising the H-1a helix in the closed conformation, . This study highlights the role of the H-1a loop in product distribution in low-fidelity terpene synthases, modulated through interactions with the metal-binding motives and the Mg^{2+}_3-PP_i complex during the catalytic cycle. It also demonstrates the complex nature of these enzymes and the subtle differences between them, with the H-1α loop playing an, apparently, less active role in high fidelity terpene synthases.

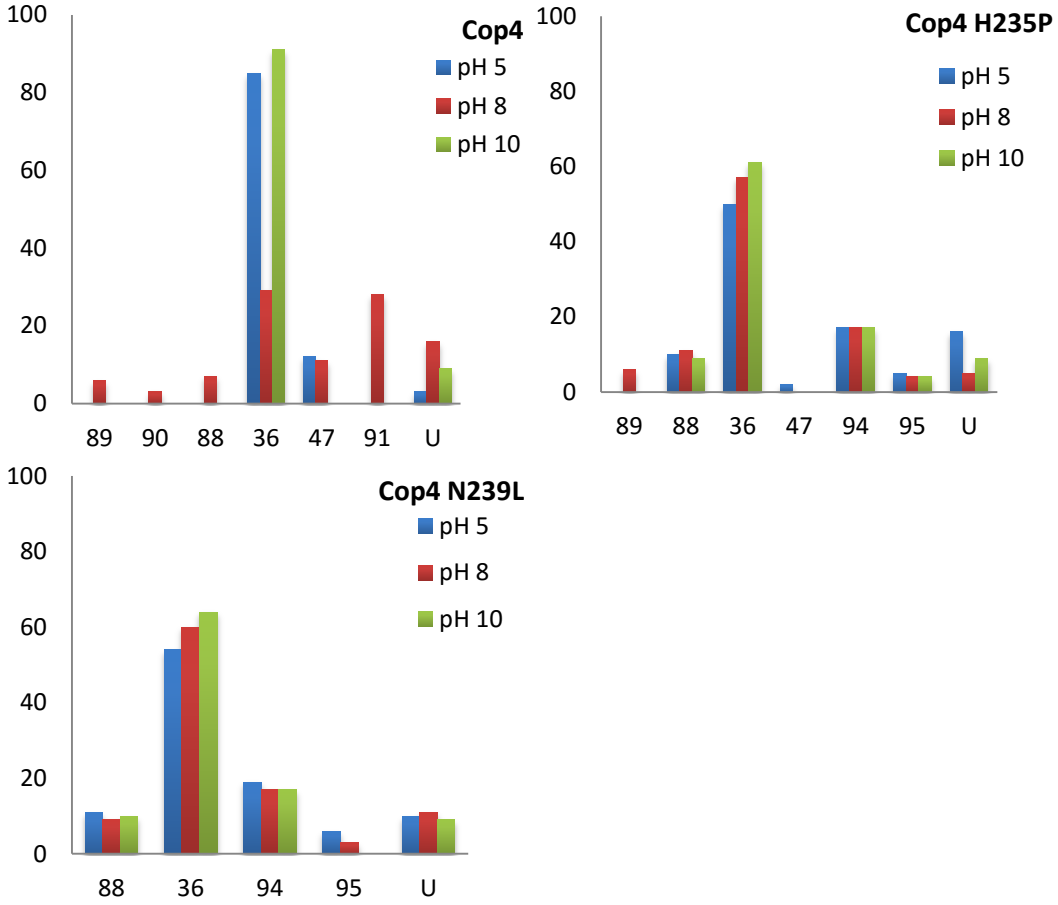


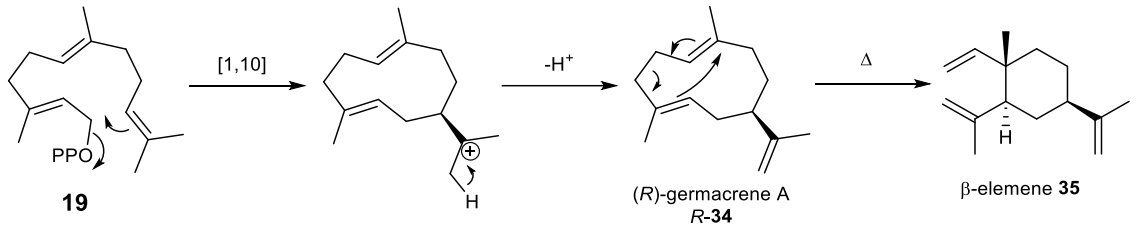
Figure 1.11 Product profiles (%) of Cop4 and mutants at pH 5 (blue), pH 8 (red) and pH 10 (green).

1.5 Terpene synthases used in this study

This study uses a number of previously studied sesquiterpene synthases available within the lab.

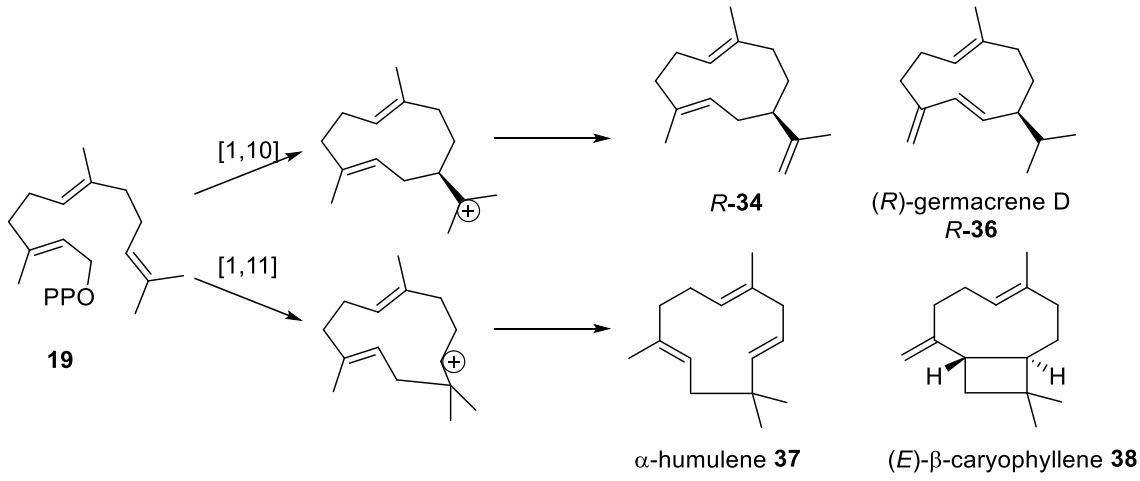
1.5.1 Germacrene A synthase

(+)-Germacrene A synthase from *Solidago canadensis* (GAS) is a 549 amino acid plant sesquiterpene synthase which generates (+)-(10*R*)-germacrene A (*R*-**34**). Germacrene A (**34**) arises from 1,10-ring closure of FDP (**19**) to form the germacryl cation followed by deprotonation at C12 to yield the final product. The absolute stereochemistry of germacrene A (**34**) produced by GAS was determined by chiral GC-MS; exploiting the thermal Cope rearrangement of germacrene A (**34**) to β-elemene (**35**), with retention of configuration, Scheme 1.8.[78]



Scheme 1.8 Formation of (*R*)-germacrene A (*R*-34) from FDP (19), and the thermal Cope rearrangement to β-elemene

In addition to producing germacrene A (**34**) (> 96%), GAS also produces ~2% of germacrene D (**36**) and α-humulene (**37**), along with traces of (*E*)-β-caryophyllene (**38**), Scheme 1.9. The formation of α-humulene (**37**) and (*E*)-β-caryophyllene (**38**), products formed through a 1,11-cyclisation mechanism prompted Gonzalez *et al.* to investigate the origin of this 1,11-cyclase activity.[79]



Scheme 1.9 Products of GAS with FDP (**19**) *via* 1,10- and 1,11-cyclisations

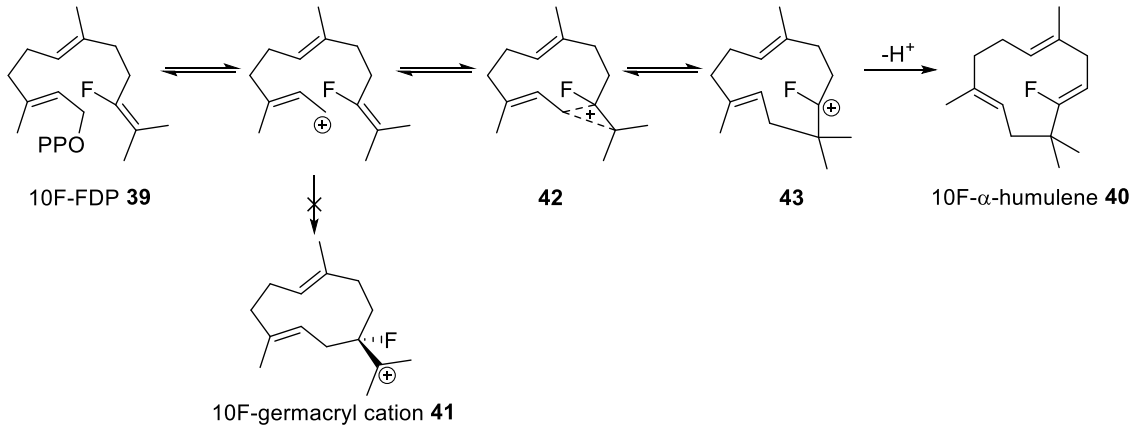
Sequence alignment of multiple plant 1,10- and 1,11-sesquiterpene cyclases identified glycine-402 as a possible determinant in 1,10-cyclisation. Substitution of glycine-402 with residues of increasing volume lead to an increase in the proportions of 1,11-products, up to a maximum of 68% in the

G402C mutant, with many mutants displaying near wild-type catalytic activity, Table 1.4.[80]

Table 1.4 Product distribution (%) of 1,10- and 1,11-products and steady-state kinetic parameters for GAS and mutants

Enzyme	1,10 products	1,11 products	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} mM^{-1})
GAS	98.25	1.75	0.043 ± 0.02	3.4 ± 0.3	12.6 ± 4.8
GAS-G402A	56.7	43.3	0.053 ± 0.01	8.9 ± 1.3	5.9 ± 1.1
GAS-G402S	55.8	44.2	0.021 ± n0.01	3.3 ± 0.9	6.4 ± 1.8
GAS-G402D	50.5	49.5	0.001 ± 0.01	0.5 ± 0.07	2.0 ± 0.2
GAS-G402C	31.6	68.4	0.030 ± 0.01	5.3 ± 1.8	5.7 ± 0.02
GAS-G402T	41.7	58.3	0.054 ± 0.01	4.7 ± 0.5	11.5 ± 0.02
GAS-G402V	42.1	57.9	0.025 ± 001	1.0 ± 0.2	25.0 ± 5.0
GAS-G402F, L, I			Inactive		

Incubation of GAS with 10-fluorofarnesyl diphosphate (10F-FDP, **39**) yielded exclusively 10F-α-humulene (**40**) as the product, albeit 15-fold less efficiently than the production of germacrene A (**34**) with FDP (**19**). This was proposed to be due to the cyclisation bypassing formation of the destabilised 10F-germacryl cation (**41**), by the formation of a bridged carbocation species (**42**). Ring expansion to the stabilised 10F-humulyl-carbocation (**43**) followed by deprotonation yields the 10F-α-humulene (**40**), Scheme 1.10.



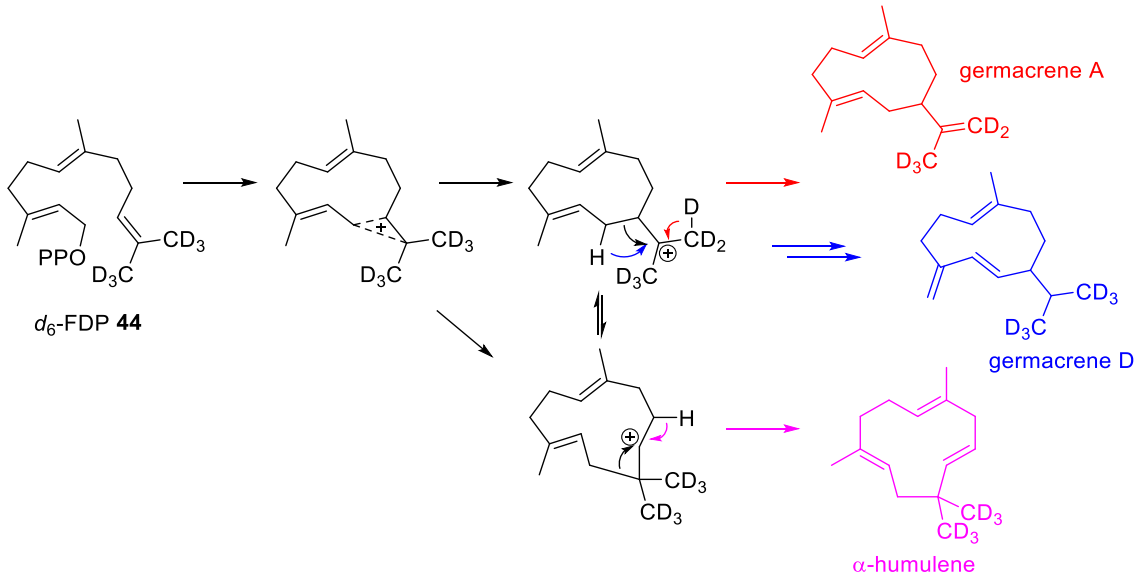
Scheme 1.10 Proposed mechanism for the cyclisation of 10F-FDP (**39**) by GAS

Use of the hexadeuterated [12,12,12,13,13,13-$^2H_6$]-FDP (d_6-FDP, **44**) provided further evidence for the intermediacy of a bridged carbocation. Incubation of GAS and mutants with d_6-FDP (**44**) showed a considerable change in product distribution, with respect to unlabelled FDP (**19**), Table 1.5. The simultaneous perturbation of the germacrene A (**34**): α-humulene (**37**) and the germacrene A (**34**): germacrene D (**36**) product ratios can be attributed to all three products

arising from a common intermediate, subject to a primary deuterium KIE affecting the final deprotonation step, Scheme 1.11.

Table 1.5 Product distributions (%) and ratios from incubations of GAS and mutants with d_6-FDP (**44**) compared to d_0-FDP (**19**). Germacrene A (**34**), germacrene D (**36**) and α-humulene (**37**)

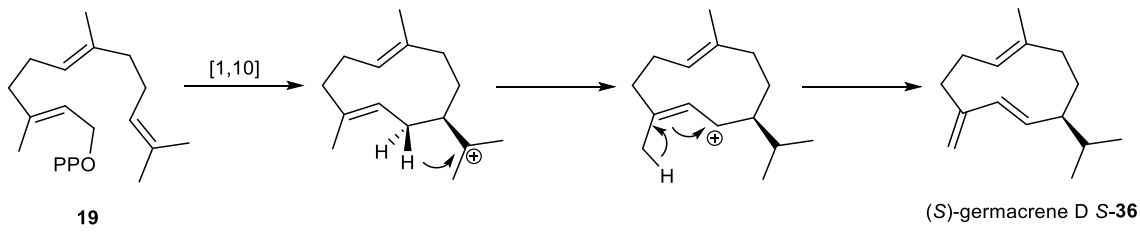
Enzyme	Substrate	**34**	**36**	**37**	**34**:**36**	**34**:**37**	KIE
GAS	d_0-FDP	96.58	1.67	1.75	57.83	55.18	4.98
	d_6-FDP	85	7.3	7.7	11.70	10.98	
GAS-G402A	d_0-FDP	54.1	2.6	41.6	20.80	1.30	4.08
	d_6-FDP	22.4	4.5	71.5	4.97	0.31	
GAS-G402S	d_0-FDP	50.7	5.1	42.5	9.94	1.19	5.09
	d_6-FDP	16.8	8.8	72.5	1.90	0.23	
GAS-G402C	d_0-FDP	13.7	15.9	61.6	0.86	0.22	4.25
	d_6-FDP	3.6	19.9	72.1	0.18	0.05	



Scheme 1.11 Possible mechanisms for the turnover of d_6-FDP (**44**) *via* a bridged carbocation species

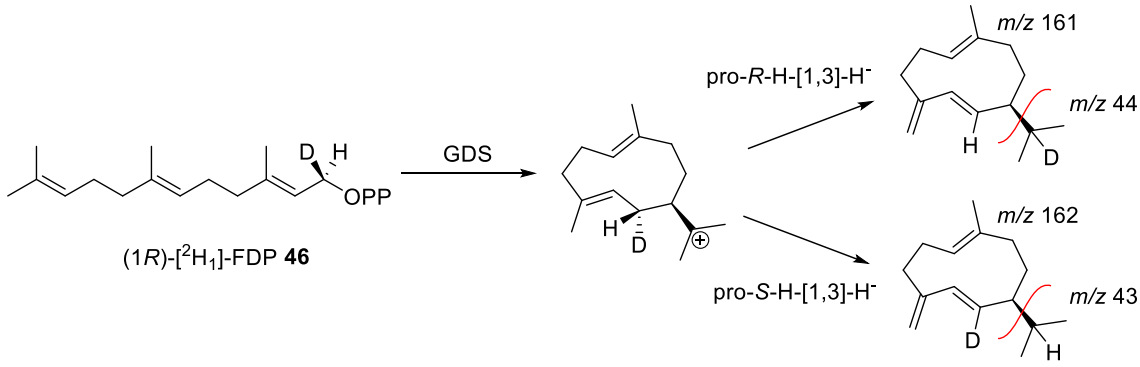
1.5.2 Germacrene D synthase

While *Solidago canadensis* produces both (*R*)- and (*S*)-germacrene D (**36**). This study uses only the (*S*)-germacrene D synthase, which has previously been isolated and characterised.[29,81]



Scheme 1.12 Cyclisation of FDP by (*S*)-GDS.

The mechanism of (*S*)-GDS was proposed to proceed through 1,10-ring closure followed by a 1,3-hydride shift from C6 (C1 of the FDP substrate) followed by deprotonation at C15 to yield (*S*)-germacrene D (*S*-**36**), Scheme 1.12. The stereochemistry of the 1,3-hydride shift was studied by incubation with deuterated-FDPs, (1*S*)-[1-^{2}H]-FDP (**45**) and (1*R*)-[1-^{2}H]-FDP (**46**) and analysis of the products by GC-MS.[81] Incubation with (1*S*)-[1-^{2}H]-FDP (**45**) and (1*R*)-[1-^{2}H]-FDP (**46**) led to products with *m/z* = 205, indicating the presence of one deuterium in place of a proton. The product from incubation with (1*S*)-[1-^{2}H]-FDP (**45**) has the most abundant fragment ion at *m/z* = 162 while that from incubation with (1*R*)-[1-^{2}H]-FDP (**46**) is at *m/z* = 161; both these fragments arise from loss of the isopropyl group from M^{+} indicating that the pro-*R*-H-1 of FDP migrates to the isopropyl group, Scheme 1.13.[64]



Scheme 1.13 Fragmentation of germacrene D from incubation with (1*R*)-[1-^{2}H$_1$]-FDPs

Both (*R*)- and (*S*)-GDS share high sequence identity with each other (85% identity, 89% similarity) however, in the (*R*)-GDS the aspartate rich region is modified to ^{303}NDxxD, due to a single base mutation in the DNA sequence.[29] The 'corrected' mutant N303D was found to be active, and the alteration had little effect on the catalytic efficiency with a slightly improved K_M and no change in k_{cat}. The inverse mutant of (*S*)-GDS, D303N, was however 10-fold less active than the wildtype, due solely to a reduction in k_{cat}, Table 1.6.

Table 1.6 Steady-state kinetic parameters of (*R*)- and (*S*)-GDS and mutants

Enzyme	k_{cat}	K_M
(R)-GDS WT	0.02	3.0 ± 0.3
(R)-GDS-D303N	0.02	2.2 ± 0.2
(*S*)-GDS WT	0.03	2.9 ± 0.3
(*S*)-GDS-N303D	0.002	2.9 ± 0.4

1.5.3 δ-Cadinene synthase

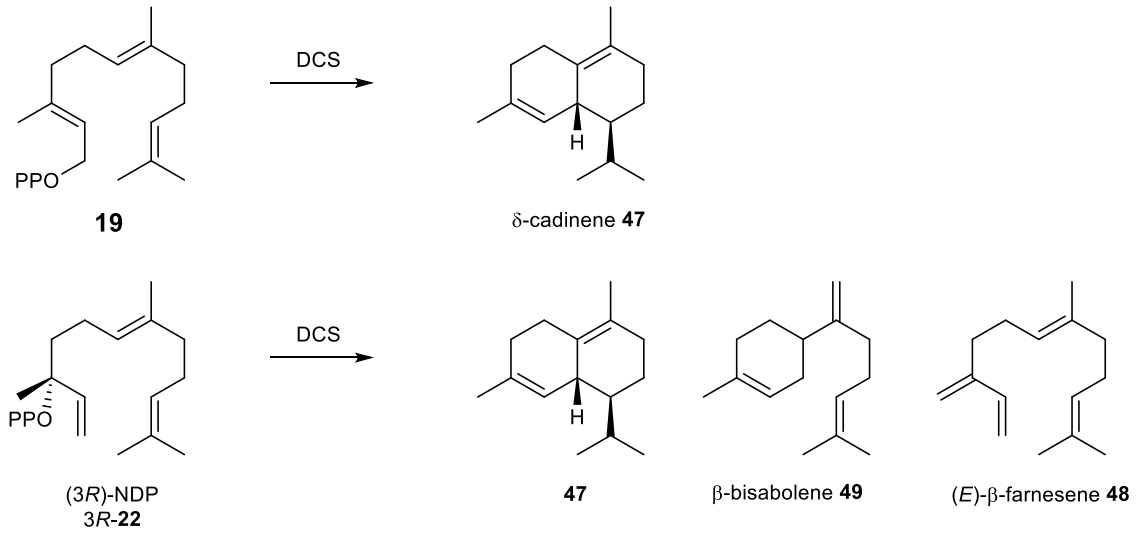
δ-Cadinene synthase from *Gossypium arboreum* (DCS) is a plant sesquiterpene synthase which catalyses the conversion of FDP to the bicyclic sesquiterpene (+)-δ-cadinene (**47**).[82] In place of the NSE/DTE motif common to terpene synthases DCS has a second aspartate rich region on helix H, ^{451}DDVAE; the roles of both aspartate rich regions were probed by site directed mutagenesis. The ^{307}DDTYD motif on helix D interacts with Mg^{2+}_A and Mg^{2+}_C, Section 1.3.2, and while the D307A, D308A and D311A mutants all produced δ-cadinene (**47**) only the D307A mutant was active enough to measure steady-state kinetic parameters. While the turnover number for this mutant is comparable to wildtype, replacement of the aspartate with an alanine has a dramatic effect on the K_M, Table 1.7.[28] Mutants of the second aspartate rich region similarly all produce δ-cadinene, but it is clear that the glutamate-455 has a significant role in coordinating with the Mg^{2+}_B, with the catalytic activity severely compromised in the E455A mutant. While the DCS-D451A and DCS-D452A display essentially unchanged steady-state kinetic parameters with respect to wildtype.

Table 1.7 Steady-state kinetic parameters of DCS and mutants. n/m denotes not measurable

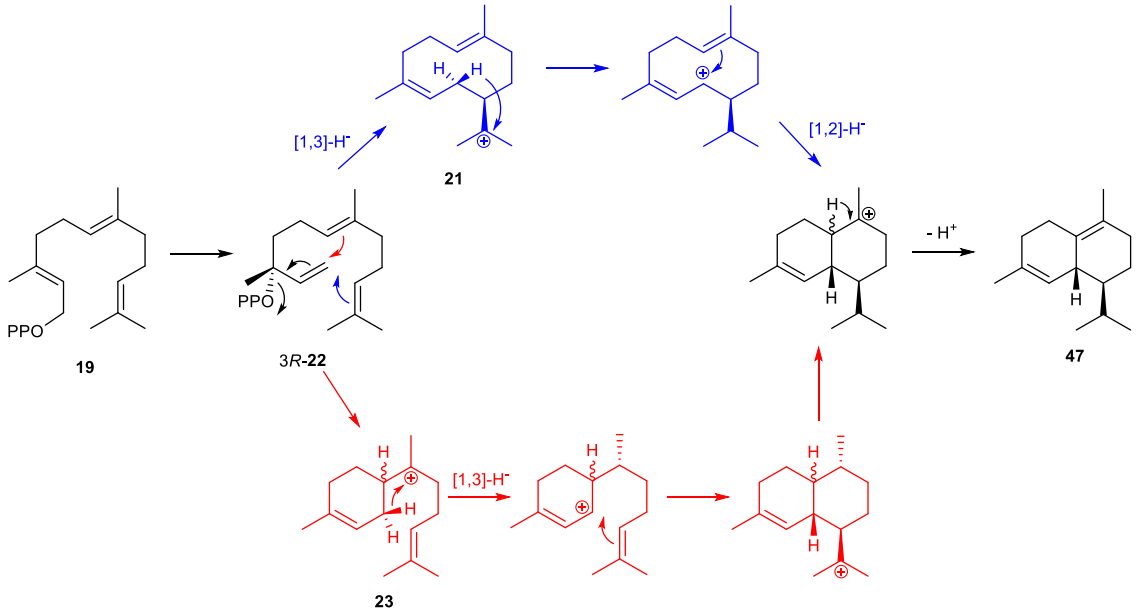
Enzyme	k_{cat} (s^{-1})	K_M (μM)
DCS WT	0.010 ± 0.001	3.2 ± 0.5
DCS-D307A	0.012 ± 0.001	43 ±1 6
DCS-D308A	n/m	n/m
DCS-D311A	n/m	n/m
DCS-D451A	0.043 ± 0.001	2.4 ± 0.3
DCS-D452A	0.014 ± 0.001	3.1 ± 1.2
DCS-E455A	n/m	n/m

Benedict *et al.* have shown that DCS is able to turnover nerolidyl diphosphate (NDP, **22**) into δ-cadinene (**47**), albeit with lower fidelity than with FDP, also producing (*E*)-β-farnesene (**48**) and β-bisabolene (**49**).[82] Competitive studies with (3*R*)-NDP (3*R*-**22**) and (3*RS*)-NDP (3*RS*-**22**) demonstrate that (3*R*)-NDP

(3*R*-**22**) is the enantiomer turned over by the enzyme, but the formation of β-bisabolene (**49**) in the turnover of NDP (**22**) highlights the possibility of an initial 1,6-ring closure followed by hydride shift and 1,10-ring closure.



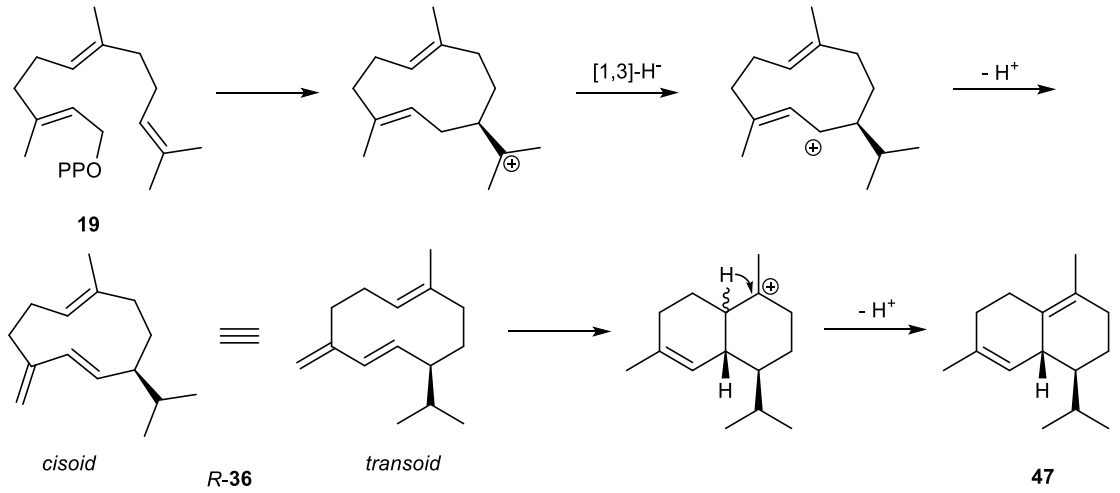
Scheme 1.14 Products of the turnover of **19** and (3*R*)-NDP (3*R*-**22**) by DCS



Scheme 1.15 Possible 1,10- (blue) and 1,6-cyclisation (red) mechanisms of DCS.

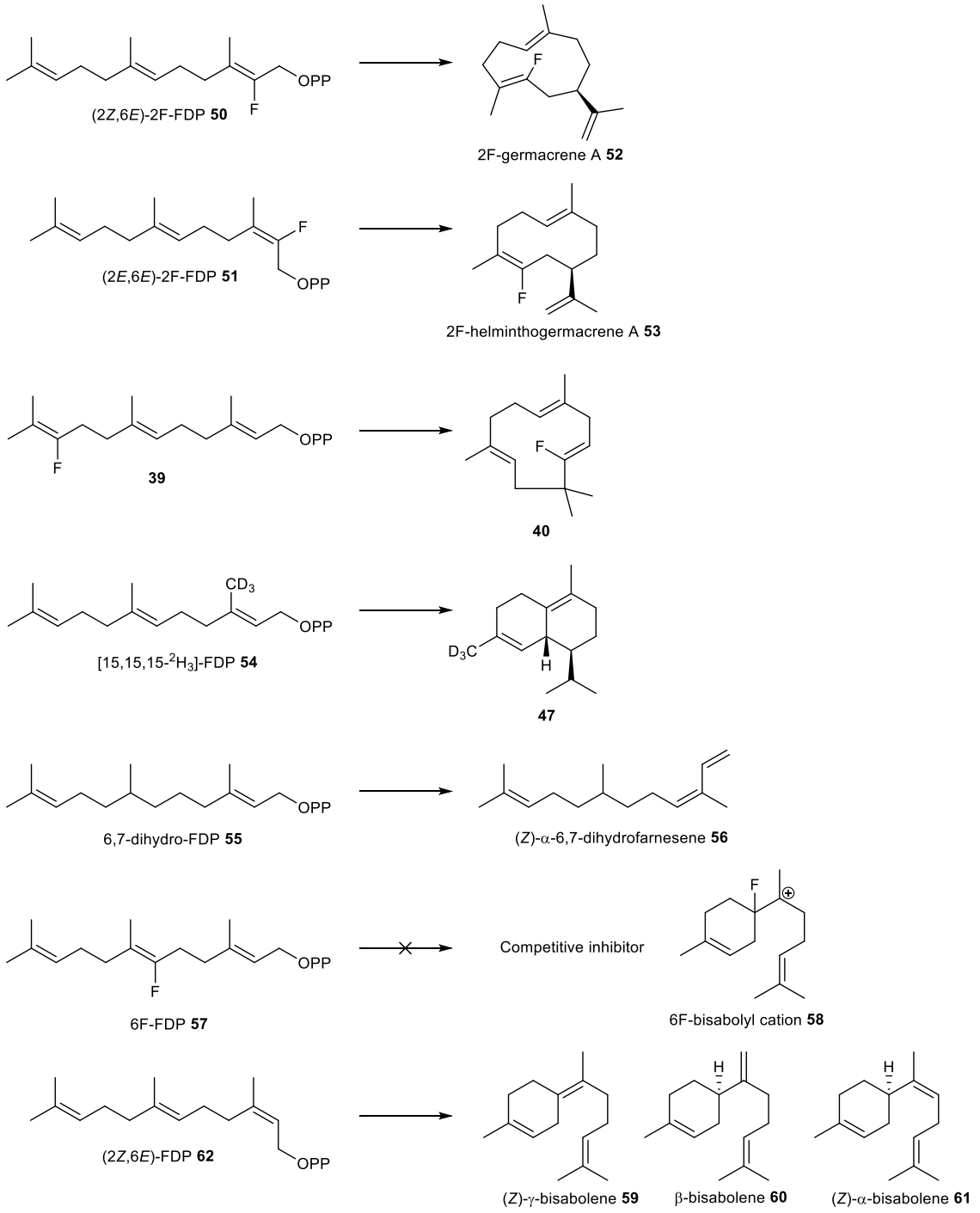
Faraldos *et al.* have performed a detailed study into the mechanism of DCS using deuterated and fluorinated FDP-analogues to distinguish between 1,6- and 1,10-cyclisation pathways.[65] In addition to 1,6- and 1,10-ring closures subsequent to isomerisation to NDP (**22**), a mechanism *via* direct 1,10-ring closure to form (*R*)-germacrene D (*R*-**36**) as a neutral enzyme-bound on-path intermediate, followed by 1,6-ring closure driven by protonation of the distal double bond has also been proposed, Scheme 1.16. This mechanism would avoid the necessity for the formation of NDP (**22**) and has been shown to

operate in a promiscuous, multiproduct sesquiterpene synthase from *Medicago truncatula*.[83]



Scheme 1.16 Mechanism for conversion of **19** to δ-cadinene *via* enzyme bound (*R*)-germacrene D (*R*-**36**) intermediate

Incubations with (2*Z*,6*E*)-2F-FDP (**50**), (2*E*,6*E*)-2F-FDP (**51**) and 10F-FDP (**39**) led to the formation of macrocyclisation products 2F-germacrene A (**52**), 2F-helminthogermacrene A (**53**) and 10F-α-humulene (**40**) respectively, supporting a 1,10-ring closure mechanism. However, the intermediacy of enzyme bound (*R*)-germacrene D (*R*-**36**) was ruled out through the use of [15,15,15-$^{2}H_3$]-FDP (**54**), which resulted in δ-cadinene (**47**) with full incorporation of the deuterium atoms. When DCS was incubated with 6,7-dihydro-FDP (**55**), however, no 1,10-cyclisation products were produced, instead the main product was (*Z*)-α-6,7-dihydrofarnesene (**56**). Along with the strong competitive inhibition of DCS by 6F-FDP (**57**), implicating the inhibition of the formation of the 6F-bisabolyl cation (**58**), and the exclusive production of bisabolene (**59**,**60**,**61**) products upon incubation with (2*Z*,6*E*)-FDP (**62**) these results are in better agreement with the 1,6-ring closure mechanism.



Scheme 1.17 Products formed by incubations of DCS with substrate analogues

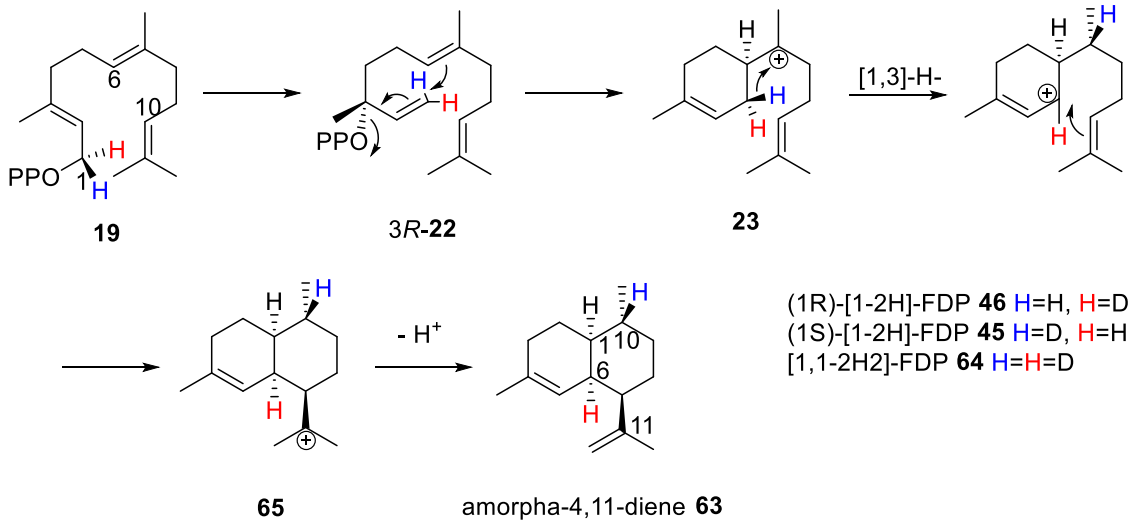
Overall the results did not unambiguously favour either 1,6- or 1,10-cyclisation mechanisms instead demonstrating the possibility of multiple mechanistic pathways and the versatility of DCS in its ability to distinguish between them in a substrate dependent manner.

Use of (1*R*)-[1-$^{2}H_{1}$]-FDP (**46**) as described with (*S*)-GDS (Section 1.5.2) resulted in the deuterated δ-cadinene (**47**) product and through the analysis of the mass-spectral fragmentation pattern of the product, it was demonstrated that

the pro-*R*-H-1 remains at the C1 position, identifying the stereochemistry of the 1,3-hydride shift as migration of the pro-*S*-H-1.

1.5.4 Amorphadiene synthase

Amorphadiene (amorpha-4,11-diene) synthase (ADS) is a plant sesquiterpene synthase from *Artemisia annua,* that catalyses the formation of amorpha-4,11-diene (**63**) from FDP (**19**).[84,85] Amorpha-4,11-diene (**63**) is a key intermediate in the biosynthesis of the naturally occurring malarial treatment artemisinin (**2**). Picaud *et al.* proposed a mechanism through formation of NDP (**22**) and 1,6-ring closure followed by a 1,3-hydride shift, based on the observation of bisabolene byproducts.[85] Investigations using (1*R*)-[1-^{2}H]-FDP (**46**), (1*S*)-[1-^{2}H]-FDP (**45**), and [1,1-$^2H_2$]-FDP (**64**) with analysis of the formed amorphadiene by GC-MS and NMR showed that the pro-*S*-H-1 is transferred to C10, while the pro-*R*-H-1 remains at the C1 position (C6 in amorphadiene numbering), Scheme 1.18. Two subsequent 1,2-hydride shifts after formation of the bisabolyl cation (**23**) would results in deuterium incorporation at C1 and C6 (amorphadiene numbering), while an initial 1,10-ring closure followed by a 1,3-hydrixde shift would result in deuterium at C11.



Scheme 1.18 Incorporation of deuterium into amorphadiene from deuterated-FDPs

Only the proposed mechanism accounts for the incorporation of deuterium into both C6 and C10 upon incubation of ADS with [1,1-$^2H_2$]-FDP (**64**).[66,86] After 1,6-cyclisation and 1,3-hydride shift, the second ring closure forms the bicyclic amorphane skeleton (**65**), with deprotonation forming the final product (**63**).

Recently, O'Maille and coworkers have identified an epistatic network of residues which controls cyclisation of FDP.[87] Using ADS and (*E*)-β-Farnesene (EBFS), also from *Artemisia annua*, a large library of enzymes was created (>27,000). ADS and BFS share 49% amino acid identity, and structure-based combinatorial protein engineering was used to breed variation form ADS into BFS. The generated mutant enzymes were sampled and screened by GC-MS to identify cyclic and linear products. Subsequent sequence analysis and further mutation enabled the authors to identify a single substitution, Y402L, which was sufficient to activate cyclisation in EBFS, causing approximately 75% of cyclic products. Further mutations and screening lead to the identification of the converse cyclisation suppressing replacement, V467G. This substitution reverted cyclase mutants to farnesene synthases and similarly, removal of this substitution in farnesene-producing mutants resulted in 6-fold increase in cyclic products. A further network of epistatic residues was also identified, some in distant positions in the protein structure highlighting the role of long-range communication across the protein. This epistatic network was also shown to control cyclisation in an EBFS homologue from *Citrus junos*.

1.5.5 Aristolochene synthase

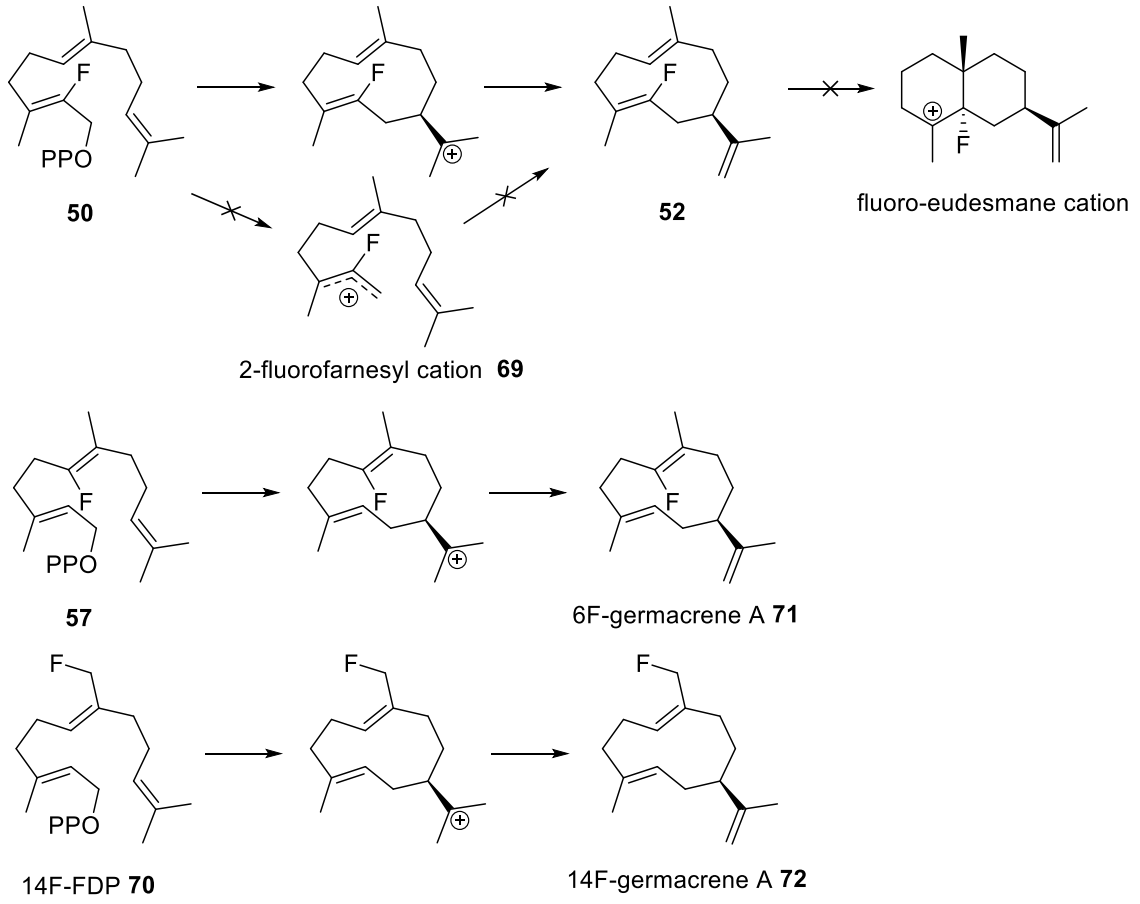
Aristolochene synthases producing (+)-aristolochene (**66**) have been identified in *Aspergillus terreus* (AT-AS)[60,88] and *Penicillin roquefortii* (PR-AS),[89] and both

have been extensively studied. In this study the enzyme used is PR-AS, producing 92 % (+)-aristolochene (**66**), 7.5 % (*S*)-germacrene A (*S*-**34**) and < 0.5 % valencene (**67**).

The relatively high proportion of (*S*)-germacrene A (*S*-**34**) produced by PR-AS led Calvert *et al.* to postulate germacrene A (**34**) as an on path intermediate, and upon replacement of an active site tyrosine with a phenylalanine the proportion of germacrene A (**34**) increased in the resulting mutant, Y92F.[90] Further alterations to this tyrosine residue also lead to increased proportions of (*E*)-β-farnesene (**48**) produced by the Y92V mutant, demonstrating the importance of tyrosine-92 in shaping FDP (**19**) into a suitable conformation for cyclisation.[73] After 1,10-cyclisation and deprotonation to form the enzyme bound germacrene A (**34**), protonation of the C6-7 double bond is proposed to drive the second ring closure to form the eudesmane cation (**68**). A subsequent series of hydride and methyl shifts followed by deprotonation at C8 yields

aristolochene (**66**). A small proportion of deprotonation occurs at C6, yielding 0.5% valencene (**67**).

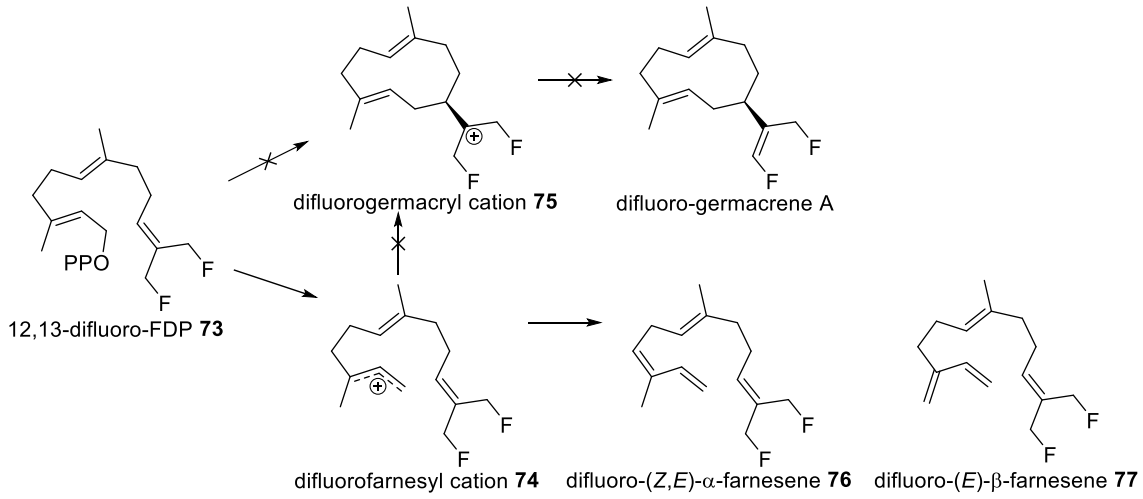
Incubation of PR-AS with (2Z,6E)-2F-FDP (2F-FDP, **50**) produced a single product which was identified as 2-fluorogermacrene A (**52**), providing compelling evidence that germacrene A (**34**) is an on-pathway product of PR-AS while simultaneously indicating that ionisation and cyclisation in PR-AS is a concerted process. The productive cyclisation to form 2-fluorogermacrene A (**52**) indicates that there is no development of positive charge in a β-position to the fluorine; If ionisation and cyclisation were to occur in a stepwise manner, formation of the allylic farnesyl cation (**21**) might be destabilised by the 2-fluoro-substituent leading to an abortive process.[68] Incubations with 6-fluoro- (**57**) and 14-fluoro-FDPs (**70**) both produced the respective fluorinated germacrene A analogues (**71**, **72**), further supporting the role of germacrene A (**34**) as an on path intermediate to aristolochene (**66**).[69]



Scheme 1.19 Cyclisation of fluorinated-FDPs (**50**, **57**, **70**) by PR-AS

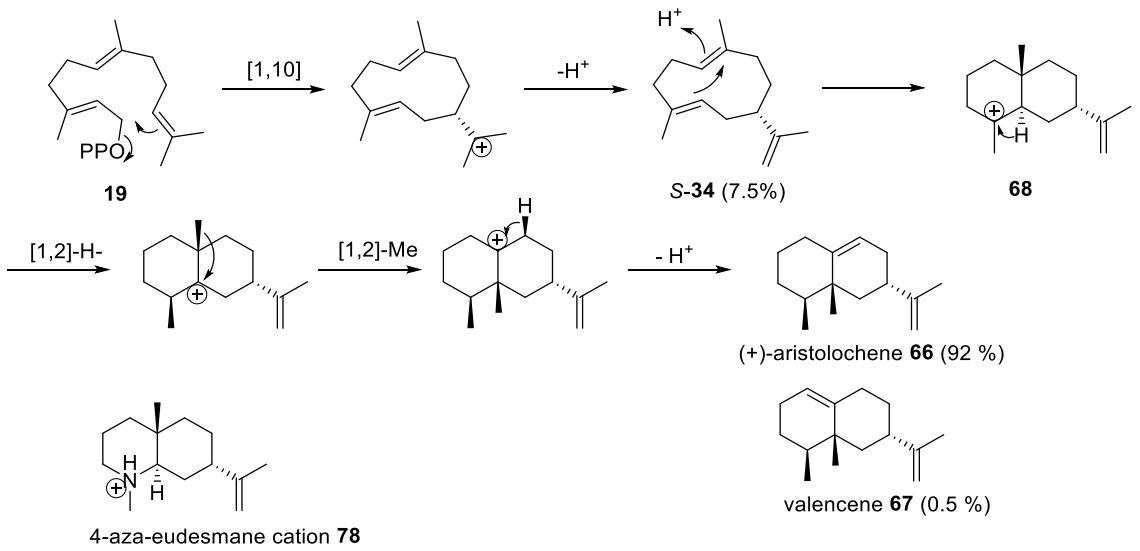
A concerted mechanism for the cyclisation by PR-AS was further supported by the use of the [12,13]-F_2-FDP (12,13-difluoro-FDP, **73**) analogue, incubation with this analogue yielded no products and it was demonstrated to act as a

competitive inhibitor with a measured K_i comparable with the reported K_M (K_i, 0.8 μM ≈ K_M, 2.3 μM).[67] In a stepwise mechanism where the farnesyl cation (**21**) is generated prior to cyclisation, accumulation of the 12,13-difluorofarnesyl cation (**74**) would be expected due to the destabilising effect of the two fluorine substituents on the 12,13-difluorogermacryl cation (**75**) at C11. Deprotonation of the 12,13-difluorofarnesyl cation (**74**) at C4 or C15 would then lead to the 12,13-difluoro-(*E*,*E*)-α-farnesene (**76**) or 12,13-difluoro-(*E*)-β-farnesene (**77**), respectively. In a situation where loss of the diphosphate is in concert with cyclisation the first carbocation to develop is on C11 in the formation of the 12,13-difluorogermacryl cation (**75**); in the destabilising presence of the two fluorine substituents this cation would not develop, leading to competitive inhibition.



Scheme 1.20 Projected turnover of 12,13-difluoro-FDP by PR-AS

The cyclisation of FDP (**19**) to aristolochene by PR-AS has been determined to go through a concerted 1,10-cyclisation and loss of diphosphate to form the germacryl cation which is deprotonated at C12 to yield the enzyme bound intermediate (*S*)-germacrene A (*S*-**34**), a small proportion of which 'leaks' from the active site. Protonation of the 6,7-double bond drives the second ring closure to generate the bicyclic eudesmane cation (**68**), subsequent 1,2-hydride and methyl shifts are followed by abstraction of the H_{si} on C8 to produce aristolochene (**66**).



Scheme 1.21 Cyclisation of FDP (**19**) by PR-AS

The 4-aza-eudesmane cation (**78**) has been synthesised and found to be a potent competitive inhibitor for PR-AS, these results suggest that this compound is an effective transition state mimic for the eudesmane carbocation (**68**) supporting the intermediacy of this carbocation in the reaction mechanism catalysed by PR-AS.[47,48]

PR-AS has also been extensively studied using site-directed mutagenesis to probe active site residues, in particular the metal binding motives[71] and residues identified as part of the $[Mg^{2+}]_3$-PP_i hydrogen bonding network.[40] Replacement of the residues in the aspartate rich region (^{115}DDxxE) showed aspartate-115 and -116 to play a greater role in magnesium binding than that played by glutamate-119; the D115E and D116E mutants displayed 12-fold and 60-fold reduction in catalytic activity with respect to wildtype, while replacement of glutamate-119 to generate the E119D mutant did not considerably affect steady-state kinetic parameters. Further alteration of D115 to a glutamine produced the inactive D115N mutant, while a similar alteration to D116 produced the D116N mutant, which was near identical in activity to the D116E mutant. In addition to an effect on catalytic efficiency alterations to the aspartate rich region also caused a decrease in fidelity in the resulting mutants, with greater proportions of germacrene A (**34**) detected in all mutants, with the exception of E119D. These altered product distribution highlights the necessity of correct and precise substrate binding and positioning for correct functioning of the enzyme.[71]

Table 1.8 Steady-state kinetic parameters and product distribution (%) of PR-AS and mutants. ND, not detected. Aristolochene (**66**), valencene (**67**) and germacrene A (**34**).

Enzyme	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)	Product distribution		
				66	**67**	**34**
PR-AS	0.043 ± 0.002	0.6 ± 0.1	72	94	2	4
AS-D115E	0.016 ± 0.0004	2.7 ± 0.2	6.0	75	6	19
AS-D115N	Inactive					
AS-D116E	0.0013 ± 0.00005	1.1 ± 0.13	1.2	62	3	35
AS-D116N	0.002 ± 0.0001	1.5 ± 0.1	1.5	63	2	35
AS-E119D	0.0048 ± 0.0001	0.15 ± 0.02	30	94	2	4
AS-N244L	Inactive					
AS-N244D	0.00012 ± 0.000003	3.3 ± 0.2	0.036	19	ND	81
AS-S248A	0.0013 ± 0.00004	5.5 ± 0.4	0.24	21	ND	79
AS-E252D	0.0022 ± 0.00009	3.4 ± 0.4	0.65	19	ND	81
AS-E252Q	0.00022 ± 0.00001	10.1 ± 0.7	0.022	ND	ND	100

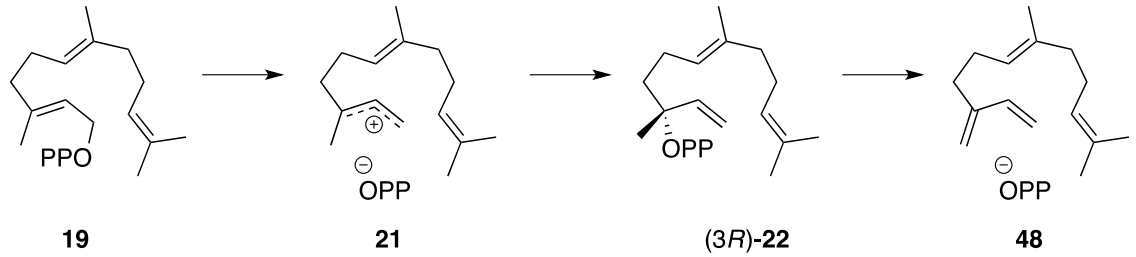
This study also probed the magnesium binding residues in NSE-motif (^{244}NDxxSxxxE). Replacement of these residues had a greater effect on both catalytic efficiency and product distribution than that reported to the above mutants of the aspartate rich region; the replacement of asparagine-244 with the isosteric leucine caused a complete loss of function, while the N224D mutant displayed a 2000-fold decrease in catalytic efficiency with a dramatic change in product distribution, producing >80 % germacrene A (**34**). This dramatic shift in product distribution was also mirrored by alterations to other residues in the NSE-motif; the S248A and E252D mutants both produced ~80 % germacrene A (**34**) coupled with a 200-fold decrease in activity, while mutation of the glutamate-252 to a glutamine yields a high-fidelity germacrene A synthase, albeit with 3500-fold lower catalytic efficiency than PR-AS.

In addition to the metal-binding motifs, other conserved residues identified as part of the hydrogen bonding network to the $[Mg^{2+}]_3$-PP_i complex in crystal structures of AT-AS[30,37] were also investigated by SDM.[40] All replacements of these residues, reducing the level of hydrogen-bonding capacity, caused a decrease in catalytic efficiency. Alterations of arginine-200 and arginine-340 to uncharged or negatively charged residues yielded inactive mutants, highlighting the importance of salt bridges between these arginine residues and the aspartate rich region or the diphosphate anion. The vital role of R200 and R340 in catalysis is cooperatively aided by tyrosine-341 and lysine-251, and replacing these residues with amino acids of reduced hydrogen bonding capability produced mutants, K251Q and Y341F, with 20- and 10-fold reductions in catalytic activity respectively but relatively unchanged product profiles. This indicates that while they are important in the recognition and

binding of the diphosphate group, they have a lesser role for active site closure and contouring.

1.5.6 (*E*)-β-farnesene synthase

(*E*)-β-Farnesene (**48**) is an acyclic sesquiterpene produced by both plants and animals, that acts as a semiochemical in roles such as alarm and defensive pheromones, and feeding and pollination stimulants.[91,92] (*E*)-β-Farnesene synthase (EBFS) from *Mentha* X *piperita*, peppermint, is a 550 residue enzyme made up of the catalytic domain and the noncatalytic N-terminal domain common to plant sesquiterpene synthases. EBFS catalyses the deprotonation and conjugate elimination of FDP (**19**) to form (*E*)-β-farnesene (**48**); Faraldos *et al.* investigated the mechanism of this transformation using FDP analogues. (2Z,2E)-2F-FDP was found to be a potent competitive inhibitor of EBFS, suggesting a stepwise mechanism for the formation of (*E*)-β-farnesene *via* the farnesyl cation. While the possible intermediacy of NDP was examined using the (2*Z*,6*E*)-FDP (**62**) and (3*R*)-NDP, ((3*R*)-**22**) both of which were converted to (*E*)-β-farnesene almost exclusively (>93%). From these data it was proposed that EBFS catalyses the conversion of FDP to (3R)-NDP via the farnesyl cation followed by deprotonation and elimination of the diphosphate to yield (*E*)-β-farnesene (**48**) Scheme 1.22.



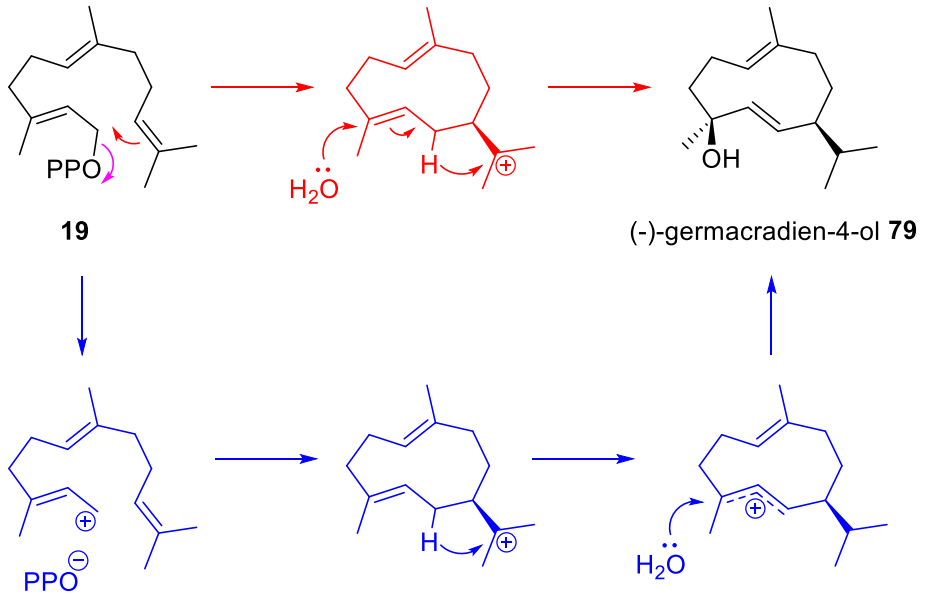
Scheme 1.22 Cyclisation of FDP (**19**) by EBFS

1.5.7 Germacradien-4-ol synthase

Germacradien-4-ol synthase (GdolS) is a 316 amino acid bacterial sesquiterpene synthase discovered through genome mining for putative sesquiterpene synthases in *Streptomyces citricolor*.[93] This enzyme catalyses the conversion of FDP (**19**) to germacradien-4-ol (**79**), a 10-membered macrocyclic alcohol, which is referred to by different names in the literature, 1,6-germacradien-5-ol and germacrene D-4-ol, in addition to germcradien-4-ol. Throughout this report it shall be referred to as germacradien-4-ol. This product has previously been identified in plant species,[94–97] and Yoshikuni *et al.* have previously engineered a bacterial δ-cadinene synthase to produce

primarily germacradien-4-ol (**79**) through the use of error-prone PCR,[98] but this is the first bacterial terpene synthase which has been identified for its biosynthesis and the first synthase which produces germacradien-4-ol as the sole product.[99]

GdolS yields germacradien-4-ol (**79**) as the single product from incubation with FDP (**19**), demonstrating a remarkable ability to effectively protect the carbocationic intermediates from water while directing the incorporation of water at the end of the cyclisation cascade with exceptional stereo- and regio-control. While Nakano *et al.* proposed a mechanism for the formation of germacradien-4-ol (**79**), thus far, there has been no mechanistic investigation into how the enzyme is able to direct this transformation and control the incorporation of water, Scheme 1.23.



Scheme 1.23 Proposed concerted (red) or a stepwise (blue) mechanisms for the conversion of FDP (**19**) to (-)-germacradien-4-ol (**79**) by GdolS.

1.6 Aims of this Project

1.6.1 Context

While terpenoids such as paclitaxel (**1**) and artemisinin (**2**) are valuable therapeutic products, the vast majority of the supply of these drugs is still extracted from natural sources due to the challenges in the synthesis of these compounds with a high degree of stereocomplexity. A growing proportion of the supply of artemisinin (**2**) is produced in a semisynthetic manner from synthetic modification of the hydrocarbon precursor amorpha-4,11-diene (**63**) which is able to be produced in relatively high amounts, >25 g/L, *via* a fermentation process.[100] More recently a group from the Max-Plank-Institute have developed

a continuous flow process to generate artemisinin (**2**) and a number of its derivatives also used in the treatment of malaria, from dihydroartemisinic acid.[101]

Despite these advances in utilising terpene synthases for the preparation of terpenoid precursors, exploiting their ability to generate multiple stereo-centres highly effectively, difficulties in selectively derivatising the hydrocarbon skeleton remains the major barrier to the conversion of the hydrocarbon terpenes to therapeutically viable terpenoids. If simple terpenoids such as terpene alcohols or epoxides could be generated using terpene synthases, the oxygen moiety could provide a 'handle' for further reactivity to be added in a regio- and stereospecific manner. The predictable and reliable design and generation of simple terpenoids could open the door for the cheap and efficient synthesis of complex terpenoids as well as providing means for the preparation of novel terpenoids as lead compounds in the development of new therapeutics.

1.6.2 Aims

The project was divided into two parts both designed to investigate methods of generating simple terpenoids. The first part focussed on a recently discovered bacterial sesquiterpene synthase, germacradien-4-ol synthase (GdolS),[93] which incorporates water in the cyclisation cascade to generate a single terpene alcohol, with a high degree of stereo- and regio-specificity. This enzyme was characterised and the mechanism probed using substrate analogues and single-point mutations to elucidate how GdolS is able to incorporate water into its product while still protecting the highly reactive carbocation intermediates.

The second part of the project involved the synthesis of a range of FDP-analogues designed to offer means of trapping carbocation intermediates in the catalysis by a range of sesquiterpene synthases. These analogues were primarily epoxy-FDP's, designed to produce simple terpenoids with hydroxy-, epoxy- or ether functionalities; providing further information about the versatility of terpene synthases as well as producing novel terpenoids which may have further uses.

2 Characterisation and mechanistic investigation of germacradien-4-ol synthase (GdolS)

2.1 Preface

As described previously (Section 1.5.7) germacradien-4-ol synthase is a 316 amino acid bacterial sesquiterpene synthase from *Streptomyces citricolor*. The enzyme converts FDP (**19**) to germacradien-4-ol (**79**) as the single product, demonstrating an ability to effectively protect the carbocationic intermediates from water while directing the incorporation of water at the end of the cyclisation cascade with exceptional stereo- and regio-control. Thus far, there has been no mechanistic investigation into how the enzyme is able to direct this transformation and control the incorporation of water.

Here, we aimed to fully characterise germacradie-4-ol synthase to determine its structure and stability, and the germacradien-4-ol (**79**) product. We also aimed to investigate and elucidate the mechanism of germacradien-4-ol synthase using fluorinated- and deuterated-FDP analogues, and to use isotopologues of water to determine the source of the water incorporation.

2.2 Enzyme characterisation

2.2.1 Expression

The expression vector containing the gene for germacradien-4-ol synthase (pET16b-SC1) was a generous gift from Professor Ohnishi, University of Tokyo. The pET16 expression vector contains a T7 promoter for controlling gene expression. The gene for GdolS had been cloned into the polycloning site with an in-frame sequence encoding an N-terminal deca-histidine tag enabling purification by nickel-affinity chromatography and the plasmid was used as it was received from Professor Ohnishi.[93] *E. coli* BL21-(DE3) strains are an ideal host for expressing these genes, because the chromosome contains the gene for T7 RNA polymerase, whose expression is regulated by the lac operon and thus allows gene expression to be controlled by IPTG (isopropyl-β-D-1-thiogalactopyranside). Two cells lines, BL21-(DE3) and BL21-Star™-(DE3), were transformed with the pET16b-SC1 expression vector and the expressions tested as described in methods, Section 6.1.10. Both cell lines have high transformation efficiencies ($1 - 5 \times 10^7$ and 1×10^8 for BL21-(DE3) and BL21-Star™-(DE3), respectively) and are deficient in *lon* and OmpT proteases; this deficiency reduces the degradation of heterologous proteins. In addition to these characteristics, BL21-Star™-(DE3) carries a mutated version of the *rne131* gene, encoding a truncated RNAse E enzyme lacking the ability to

degrade mRNA. This mutation results in an increase in the stability of mRNA and generally increases the yield of recombinant proteins from T7 expression vectors. Also available was the BL21-CodonPlus(DE3)-RP strain. CodonPlus strains contain extra copies of the genes encoding the tRNAs which most commonly limit translation of recombinant proteins. The BL21-CodonPlus-RP strain contains copies of the genes encoding the tRNAs for arginine and proline, which are rare in *E. coli* but more frequently used in GC-rich organisms, such as *Streptomyces*.

Samples taken from the expressions at regular intervals were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE, Section 6.1.12), Figure 2.1. Expression in *E. coli* BL21-(DE3) showed increasing levels of a protein after IPTG addition. This band is at a molecular weight of approximately 35 to 40,000, which corresponds to the molecular weight of GdolS with the N-terminal His-tag (molecular weight 38,687). In the test expression in *E. coli* BL21-Star™-(DE3) there is a small increase in the protein band at 35,000 up to 3 hours post-IPTG addition, but not the level of over-expression seen in BL21-(DE3). The harvested cell pellets were resuspended in buffer and cells lysed by sonication (Section 6.1.11). The resulting suspension was clarified by centrifugation and the supernatant solution and pellet were analysed by SDS-PAGE, Figure 2.1, showing the presence of recombinant protein in both the supernatant solution and as inclusion bodies in the pellet.

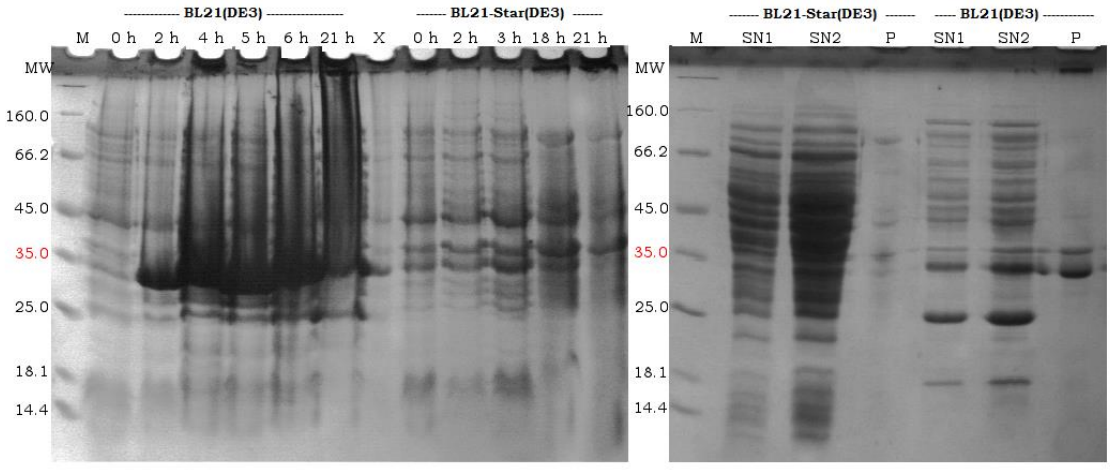


Figure 2.1 SDS-polyacrylamide gels of the test expression of pET16b-GdolS in BL21(DE3) and BL21-Star(DE3). Left, analysis of samples taken at hourly periods after induction, where the lanes are labelled by the number of hours after induction (X is an empty lane). Right, analysis of the cell pellet (P) and supernatant solution (SN) taken after 21 h, after cell lysis, SN2 is twice the concentration of SN1 in both cases. MW, molecular weight x10^3, approximate expected mass of GdolS highlighted in red.

Due to the effective expression in *E. coli* BL21-(DE3) it was decided not to investigate the expression in *E. coli* BL21-CodonPlus(DE3)-RP. In order to increase the proportion of protein in the supernatant solution further expressions were performed in *E. coli* BL21-(DE3); with reduction of the temperature to 16 °C after addition of IPTG, and inducing with a lower concentration of IPTG and maintaining the temperature at 37 °C.

When the temperature of the cell culture was reduced to 16 °C after addition of IPTG, expression was allowed to continue for 21 hours. While expression was effective and soluble protein was obtained from the supernatant solution after cell lysis, there was still considerable protein formed as inclusion bodies. When reducing the concentration of IPTG used to induce the expression, from 0.5 mM to 0.2 mM, and allowing the expression to continue at 37 °C for 3 hours, the proportion of soluble protein appeared to be greater than that formed as inclusion bodies, Figure 2.2. Expression in this manner produced sufficient protein for further studies, and no further optimisation was carried out.

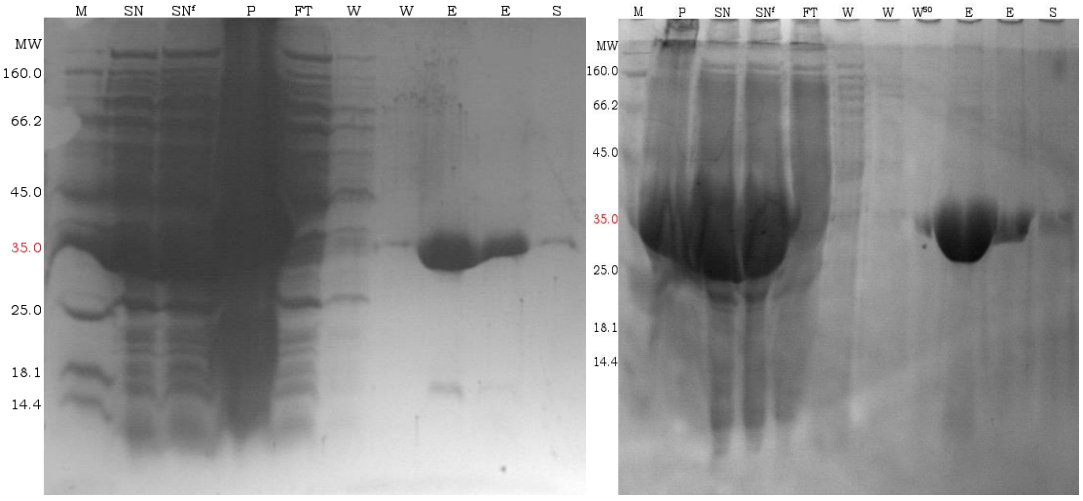


Figure 2.2 SDS-polyacrylamide gels of the expression of pET16b-GdolS in BL21(DE3). Left, at 37 °C for 3 h after addition of 0.2 mM IPTG. Right, at 16 °C for 21 h after addition of 0.5 mM IPTG. Also showning purification fractions; pellet (P), supernatant solution before and after filtering (SN and SN^{f}), flowthrough (FT), washes (W), elution (E) and stripping with 500 mM imidazle (S).MW, molecular weight x10^{3}, approximate expected mass of GdolS highlighted in red.

Expressed protein was isolated and purified to greater than 95 % purity by nickel-affinity chromatography (Figure 2.3, Section 6.1.11). The supernatant solution was applied to the Ni-NTA resin pre-equilibrated in the same buffer (50 mM Tris, 100 mM NaCl, 10 mM imidazole, pH 8.0), and then washed sequentially with increasing concentrations of imidazole (10 mM, 250 mM and 500 mM). The 10 mM imidazole washes remove any proteins bound through non-specific interactions and proteins with low histidine content, and the desired protein is eluted in the 250 mM imidazole washes. While a small

amount of protein is also removed in the 500 mM imidazole wash, only the 250 mM imidazole elutions were combined and carried forward. Nickel-affinity chromatography yields a single band between molecular weights 35 and 40,000 by SDS-PAGE, Figure 2.3. The purified protein was dialysed to remove imidazole, and concentrated by ultrafiltration before it was characterised in terms of structure, stability and steady-state kinetic parameters.

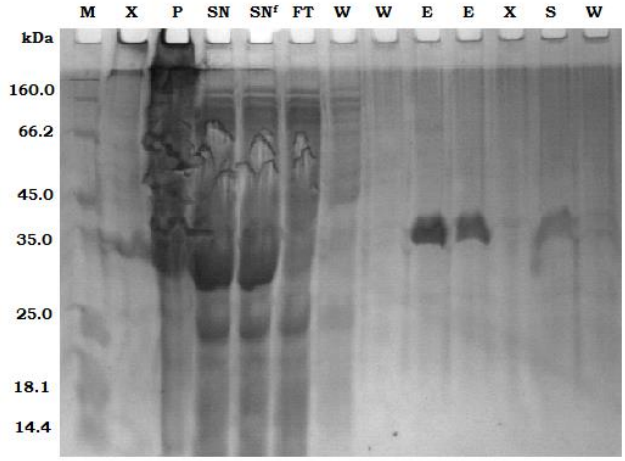


Figure 2.3 Purification of GdolS from expression of pET16b-GdolS in BL21(DE3), showing the pellet (P), supernatant solution before and after filtering (SN and SNf), flowthrough (FT), washes (W), elution (E) and stripping with 500 mM imidazole (S).

2.2.2 Kinetics

The steady-state kinetic parameters for the reaction catalysed by GdolS were measured using a radiolabelled-substrate assay commonly used with terpene synthases, involving incubation of a range of concentrations of [1-^{3}H]-FDP ([1-^{3}H]-**19**) with a known concentration of enzyme for a defined length of time.[102] Measurements of the radioactivity of the hexane extractable products through scintillation counting are then fitted against the Michaelis-Menten equation to give V_{max} and K_M. (Section 6.1.19)

Prior to the measurement of steady-state kinetic parameters optimal conditions for pH, concentration of magnesium ions and concentration of enzyme were determined. While optimising parameters the buffer conditions reported by Nakano *et al.* were used as a starting reference, 50 mM HEPES (2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid) buffer at pH 8, 5 mM 2-mercaptoethanol was used with 5 nM enzyme and 4 μM FDP (**19**) for the optimisation of magnesium ion concentration in the range of 0.5 to 10 mM, and incubation of the reaction for 10 minutes at 30 °C. Within the range [Mg^{2+}] = 0.5 – 10 mM there was only a small increase in rate and so 2.5 mM was chosen, in consistency with Nakano *et al.*, Figure 2.4.

When optimising pH, 50 mM buffers (MES, pH 6 – 6.5; HEPES, pH 7 – 8.5; CHES, pH 9 – 10) with 2.5 mM $MgCl_2$, 5 mM 2-mercaptoethanol were used with 14 nM enzyme and 4 μM FDP (**19**), and incubation of the reaction for 10 minutes at 30 °C. In the pH range 7.5 to 9.5 GdolS proved stable with pH having very little effect on the rate of catalysis, while beginning to drop off either side of this plateau, at pH 7 and pH 10. A pH of 8 was therefore chosen for the measurement of kinetic parameters, a value on the stable plateau while not too close to the point at which the rate reduces.

The enzyme activity was measured in the range, [GdolS] = 10 nM – 1 μM using 50 mM HEPES buffer at pH 8 with 2.5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 4 μM FDP and incubation of the reaction for 10 minutes at 30 °C. Activity versus enzyme concentration was found to be linear in the range below 50 nM, at this concentration the rate plateaued, suggesting that the enzyme was beginning to aggregate, a behaviour which has previously been reported with other terpene synthases.[71] Due to this, a concentration of enzyme of 20 nM was used for the measurement of kinetic parameters, a concentration in the linear range while with sufficient activity for effective detection and measurement.

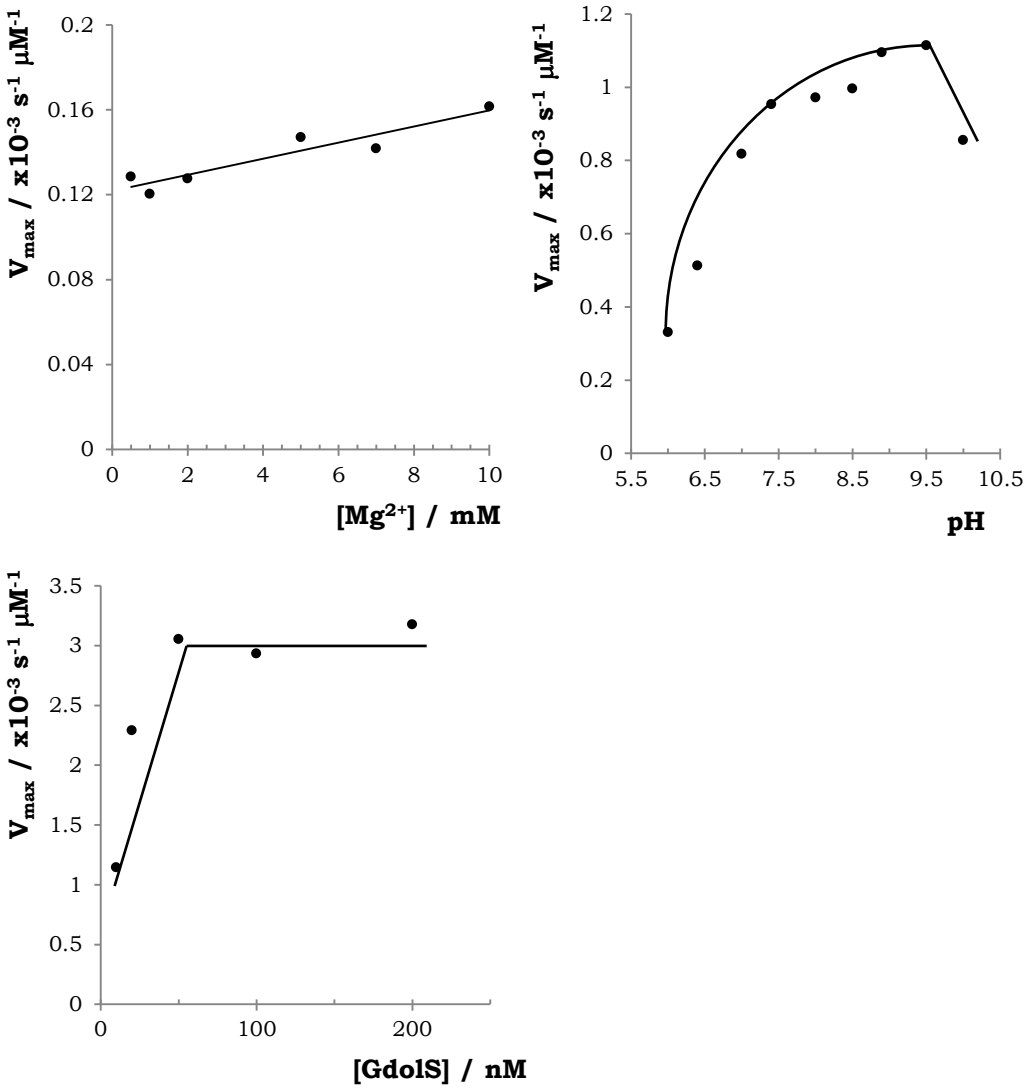


Figure 2.4 Graphs showing maximal rate versus (top left) [Mg^{2+}], (top right) pH and (bottom) enzyme concentration for turnover of [1-^{3}H]-FDP with GdolS

Using the optimised conditions of [Mg^{2+}] = 5 mM, pH 8.0 and [E] = 30 nM, measurement of the steady state kinetic parameters was carried out varying [1-^{3}H]-FDP ([1-^{3}H]-**19**) concentration from 0.02 to 7 µM, Figure 2.5. The k_{cat} and K_M were found to be 0.079 ± 0.003 s^{-1} and 1.07 ± 0.13 µM respectively. Ohnishi *et al.* measured steady state kinetic parameters using a GC-MS based method and the values they reported differ to those reported here. The k_{cat}, 0.198 ± 0.004 s^{-1}, is approximately three-fold higher than we find while there is a 10-fold decrease in K_M, 0.115 ± 0.0035 µM. Differences between steady state kinetic parameters based on the method used have been reported for other terpene synthases.[103,104] Both sets of values reported here are comparable with those reported for other bacterial terpene synthases, although the majority of those are also measured using an assay with radiolabelled FDP (**19**).

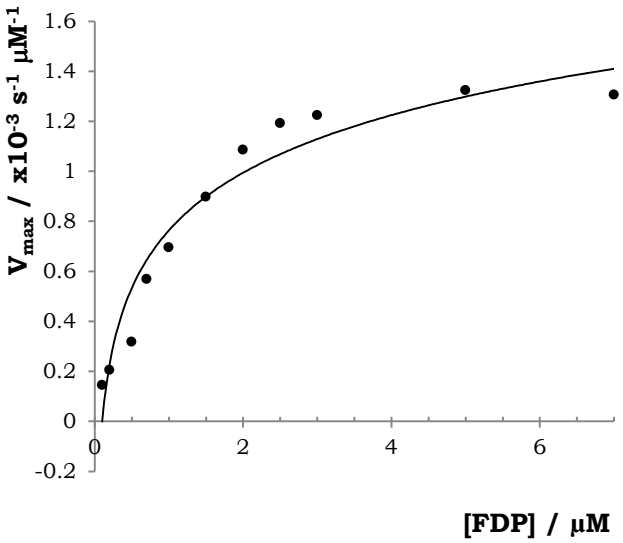


Figure 2.5 Representative graph for the calculation of steady-state kinetic parameters of GdolS

2.2.3 Product Analysis

Analysis of the pentane extractable products from the incubation of GdolS with FDP (Section 6.1.17) by GC-MS showed a single product peak with a molecular ion of *m/z* = 222 (major fragments *m/z* = 207 ($[M - CH_3]^+$), 204 ($[M - H_2O]^+$), 161 ($[M - (H_2O + C_3H_7)]^+$)). The mass spectrum did not compare well with that reported in the literature for this compound.[93,96,98] In the mass spectra reported in the literature the most abundant peak is that at *m/z* 81, with the molecular ion, $[M]^+$ (*m/z* 222), approximately 5% and the peaks for $[M - H_2O]^+$ (*m/z* 204) and $[M - CH_3]^+$ (*m/z* 207) approximately 10 and 15% respectively. In comparison, the spectra reported here has the most abundant peak at *m/z* 161, with the molecular ion at approximately 1%. The peak for $[M - H_2O]^+$ is much more abundant, 60%, while the peak for $[M - CH_3]^+$ is also less intense, 2%. These differences have been attributed to variations in the tuning of the mass spectrometers.

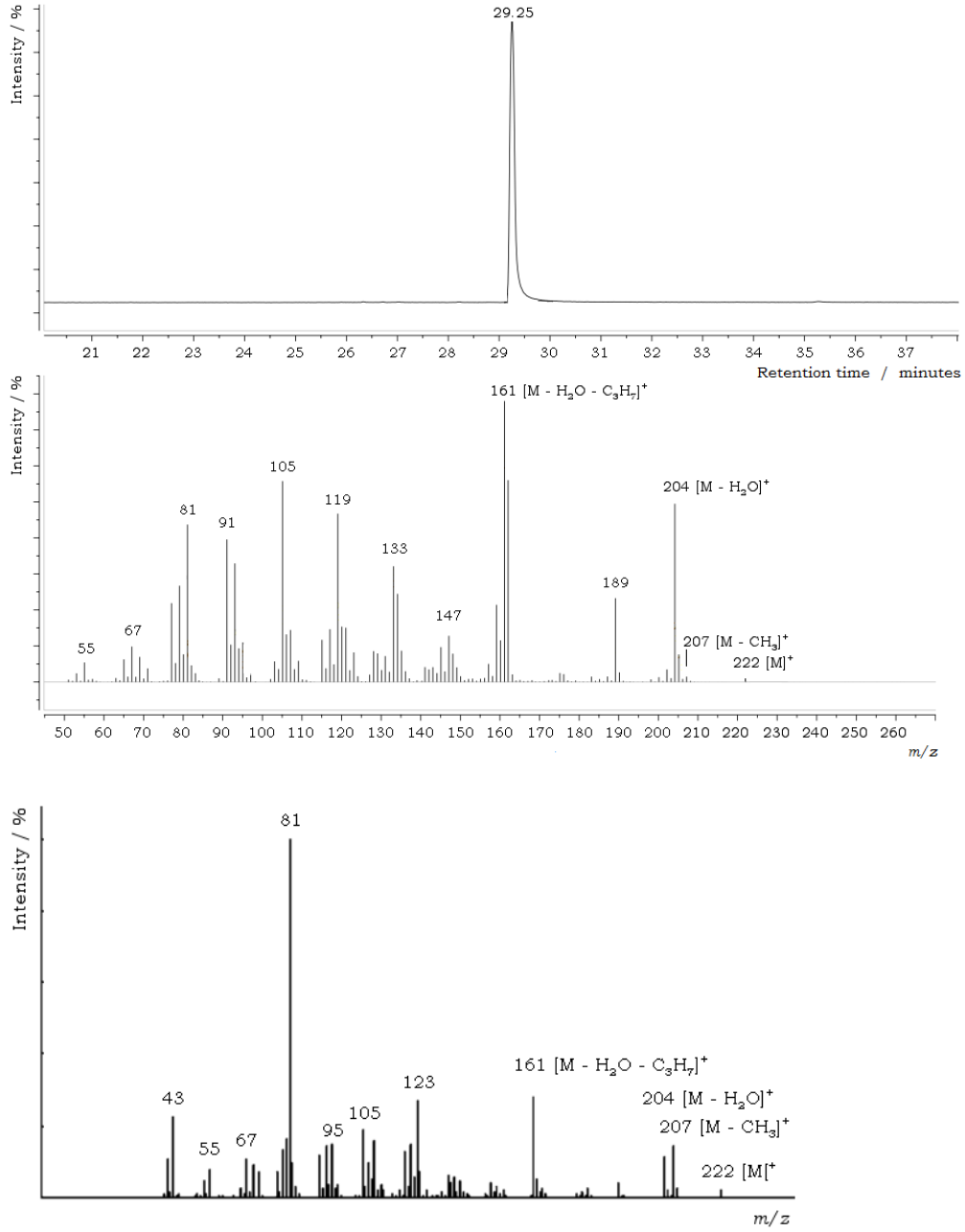


Figure 2.6 GC-MS analysis of the pentane extracted products of an overnight incubation of GdolS with FDP. Top, gas chromatogram. Middle, background subtracted mass spectrum of the peak at 29.25 minutes. Bottom, reported mass spectrum of germacradien-4-ol.[93]

NMR spectroscopic analysis of germacradiene-4-ol.

While analytical scale incubations (total volume 250 µL) are usually sufficient to analyse the pentane extractable products of the enzymatic incubations by GC-MS, it was necessary to prepare a larger sample to unambiguously identify the product of GdolS and so a preparative-scale incubation (total volume 500 mL) was used to produce milligram quantities of the terpene alcohol. As described in Methods (Section 6.1.21) 90 mg of FDP was incubated over 3 days with GdolS (0.4 mM) and the product extracted into pentane. After gentle removal of the pentane at room temperature under reduced pressure, a single

product was recovered in 80.5% yield (36.8 mg) as a colourless oil. The sample was characterised using both ^{1}H and ^{13}C NMR spectroscopy using a series of 1D and 2D experiments and was identified as the reported germacradien-4-ol (**79**), showing the characteristic doublets for the isopropyl group at 0.78 and 0.82 ppm, in addition to the singlets for the allylic-C14, 1.54 ppm, and the C15, 1.19 ppm, methyl groups. The vinylic-proton, H1, shows as a broad doublet with a coupling constant of 11 Hz consistent with a large dihedral angle of coupling, while the vinylic H5 and H6 protons are a doublet and a doublet of doublets respectively, coupling to each other with a coupling constant of 16 Hz, Figure 2.7, Table 2.1.

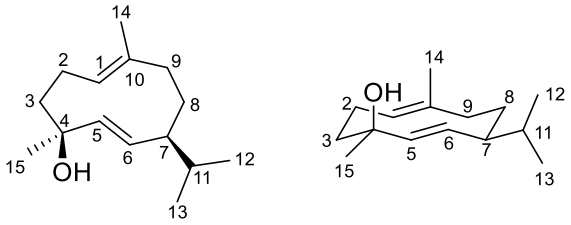


Figure 2.7 Structures of germacradien-4-ol (**79**) with numbering system used in assignment

Table 2.1 Complete assignment of germacradien-4-ol (**79**). Chemical shifts are reported relative to TMS in ppm

	^{1}H	^{13}C
1	4.95 (bd, *J* = 11 Hz, 1H)	129.0
2	2.50 (dddd, *J* = 14, 12.5, 11, 3.5 Hz, 1H) 1.99 – 1.88 (m, 1H)	23.8
3	1.68 – 1.56 (m, 2H)	39.7
4	-	73.3
5	5.25 (d, *J* = 16 Hz, 1H)	140.2
6	5.16 (dd, *J* = 16, 9.0 Hz, 1H)	125.8
7	2.07 – 1.95 (m, 1H)	52.9
8	1.40 – 1.33 (m, 2H)	26.1
9	2.32 – 2.16 (m, 2H)	41.4
10	-	132.7
11	1.46 – 1.37 (m, 1H)	33.1
12	0.78 (d, *J* = 6.5 Hz, 3H)	20.8
13	0.82 (d, *J* = 6.5 Hz, 3H)	19.1
14	1.54 (s, 3H)	16.9
15	1.19 (s, 3H)	30.8

This data is consistent with the (4*S*, 7*S*)-stereoisomer as reported by Nordin *et al.*,[105] who performed the most in depth assignment of this compound to-date. The optical rotation of the enzymatically produced sample was measured as -163.4° ($[\alpha]^{20}_{D}$ *c*.0.991, $CHCl_3$), while the maximum value reported by Nordin *et al.*[105] is – 184.6° ($[\alpha]^{25}_{D}$ *c*.1.02, $CHCl_3$), differing by approximately 20° from the sample generated herein.

GC-FID analysis of enzymatically produced germacradien-4-ol (79)

To examine the optical purity of the sample produced by GdolS, it was analysed by gas chromatography using a column containing a cyclodextrin based chiral stationary phase (SUPELCO Aztec CHRIRALDEX™ B-DM). In the absence of a racemic sample to ensure separation of enantiomers, the sample was eluted using a variety of methods including those that have been successfully used for the separation of the tertiary terpene alcohols, bisabolol (**96**) and nerolidol (**97**) (Programs 1 & 2, respectively).[106] Care was taken not to overload the column by injecting 5 μL and using a 50:1 split of the injected sample, the peak shape was carefully monitored throughout for any indication of shoulder peaks.

Table 2.2 GC-FID Programs

Program	1	2	3	4
Initial temperature (°C)	70	90	155	100
Initial time (minutes)	3	2	60	2
Temperature ramp (°C/minute)	5	5	-	2
Final temperature (°C)	160	170	-	150
Hold time (minutes)	30	40	-	30
Program length (minutes)	51	58	60	57
Injection	Split 50:1	Split 50:1	Split 50:1	Split 50:1

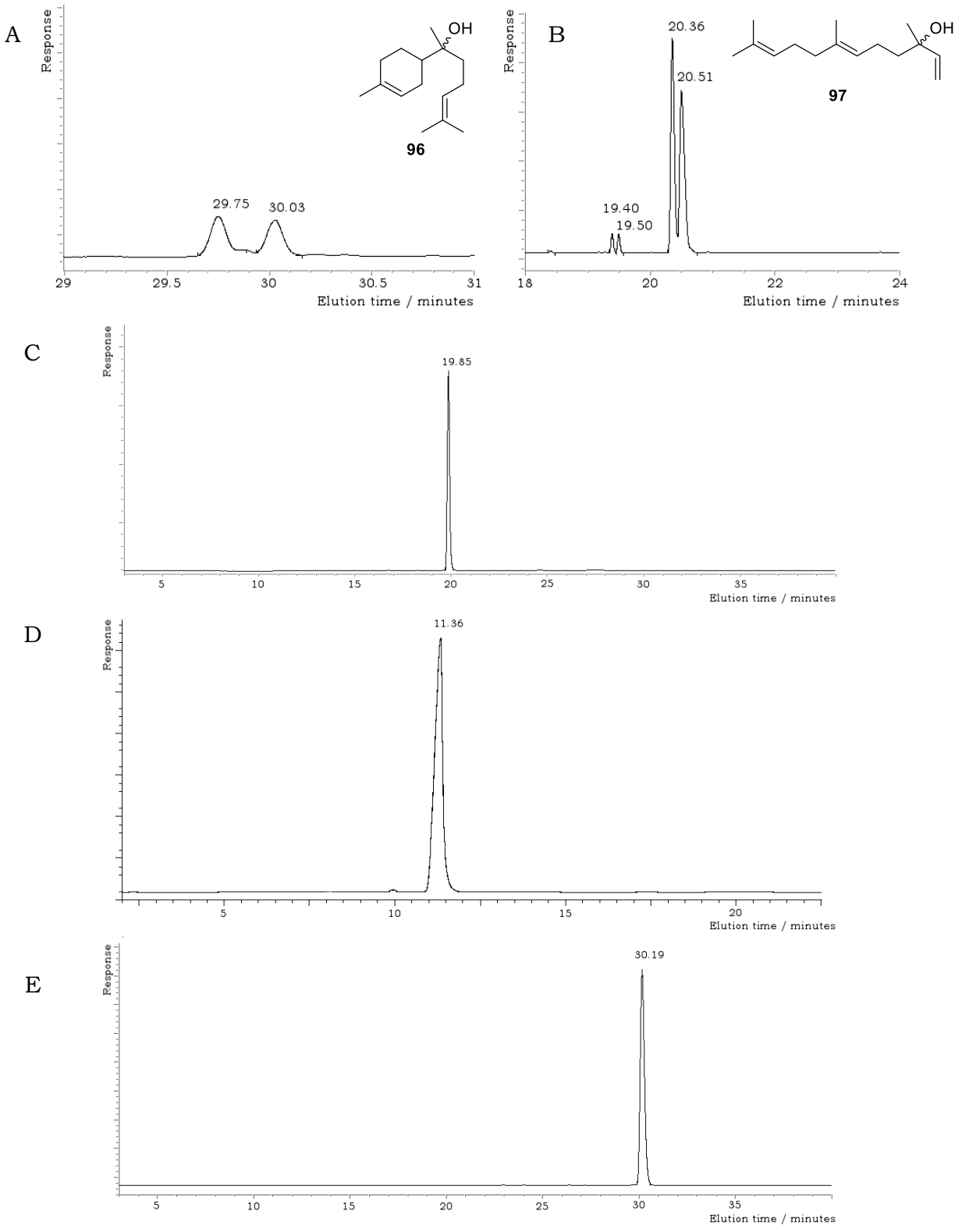


Figure 2.8 GC-FID Chromatograms showing (A) separation of racemic bisabolol (**96**) using method 1, (B) separation of racemic *trans*-nerolidol (**97**) using method 2,and analysis of germacradien-4-ol (**79**) using methods 2, 3 and 4 (C, D and E respectively).

In all methods, only a single compound eluted and the peak shape was regular and symmetrical, particularly when a temperature gradient was used. The observed $[\alpha]_D$ (- 163.4°) compared to that reported (- 184.6°) gives a theoretical

enantiomeric excess (*e.e*) of 88.5%, indicating that any secondary peak would be approximately 6 % of the major peak and should be easily distinguishable from the baseline. While not conclusive, this is strongly suggestive that only a single stereoisomer was present.

While the majority of the time only a single product was seen by GC-MS, which was consistent with the NMR spectra, occasionally two other products were seen in the chromatogram by GC-MS analysis. One peak was identified as δ-cadinene by comparison with a genuine enzymatically produced sample, while the second peak was tentatively assigned as γ-cadinene by comparison with the NIST mass spectral library.[107] These peaks were not consistent in proportion with the germacradien-4-ol and it was proposed that they were being generated by rearrangement of the germacradien-4-ol to eliminate water. Indeed, further review of the literature found reports that after treatment with dilute formic acid germacradien-4-ol yields a mixture of hydrocarbon and alcohol products.[94,105] In the absence of water in the gas chromatograph no alcohol products were observed, Figure 2.9. The addition of triethylamine to the pentane extracted products prior to injection in the GC-MS reduced the amount of δ- and γ-cadinene to nearly zero, while replacement of the GC inlet liner and syringe eliminated any rearrangement. In future chapters the presence of δ- and γ-cadinene in gas chromatograms will not be discussed, but will be taken into account in integrations for the calculation of product distributions.

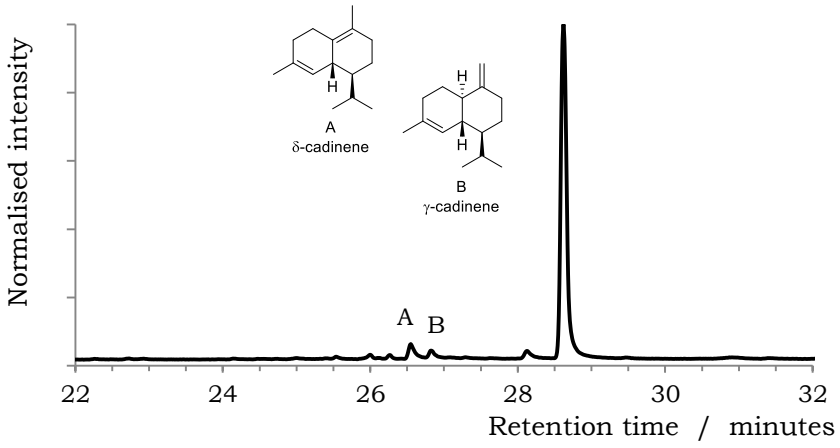


Figure 2.9 Gas chromatogram of the pentane extracted products from incubation of GdolS with FDP, showing traces of δ-cadinene (A) and γ-cadinene (B)

2.2.4 Structure of GdolS

The protein, like other terpene synthases, was expected to have a high content of α-helices. GdolS proved to be unstable in 50 mM phosphate buffer, clearly

precipitating in the cuvette instantly; due to the strong absorbance of Tris and HEPES buffers below 200 nm, measurements were taken with enzyme diluted to approximately 10 μM in distilled water. Circular dichroism of GdolS was recorded between 190 and 270 nm and showed a typical α-helical spectrum with minima at 208 and 222 nm, and a maximum at 190 nm. Measuring the mean residual ellipticity at 222 nm while increasing the temperature gave a melting temperature of 55 °C, Figure 2.10. Full spectra were also recorded at different temperatures to give more information about the melting temperature of the protein, while recording the spectrum of the sample after returning to 20 °C indicated that the structural changes upon heating irreversibly denatured the protein.

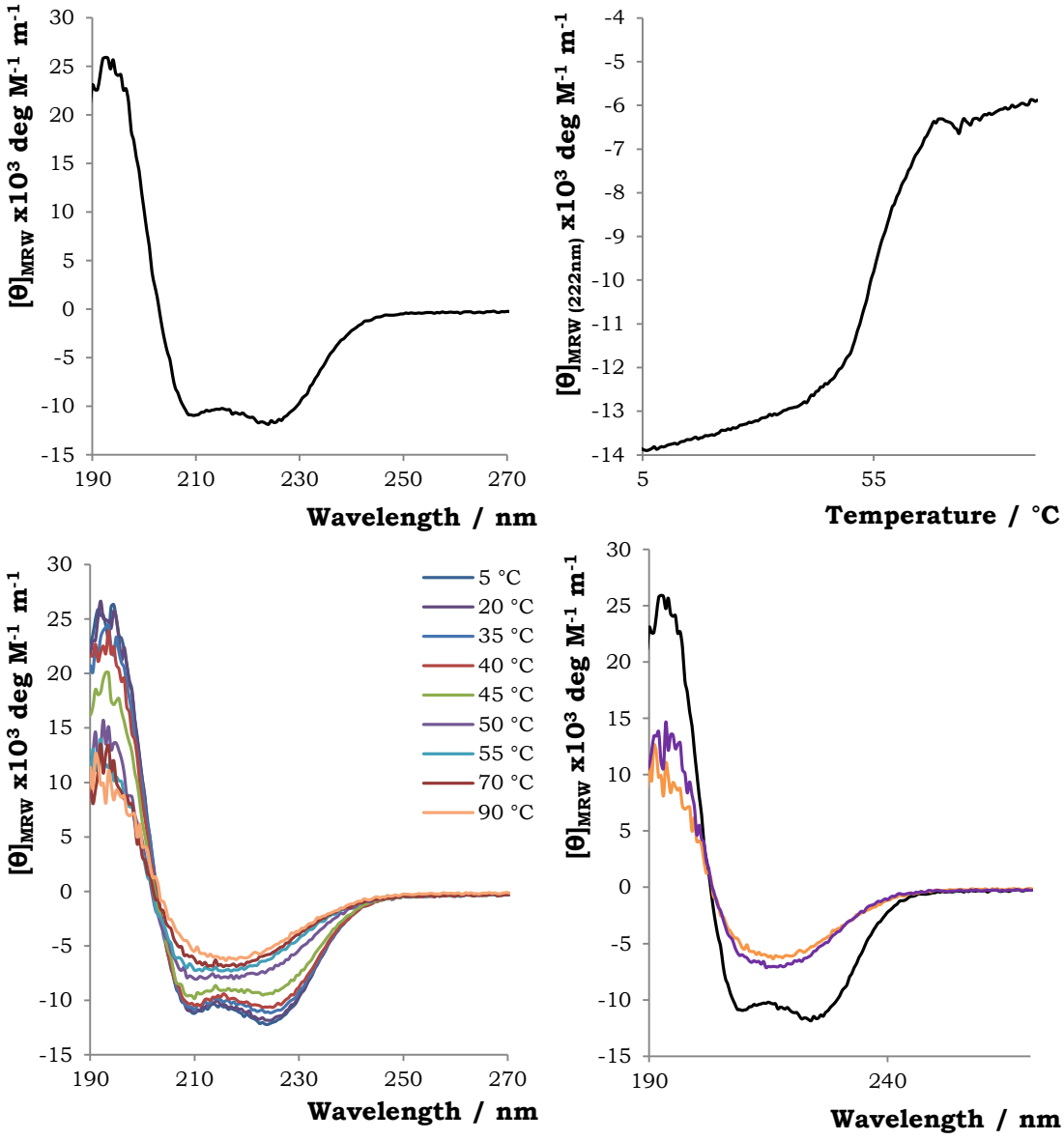


Figure 2.10 Top left, Circular dichroism spectra of GdolS. Spectrum of GdolS at 20 °C. Top right, Mean residue ellipticity at 222 nm across a temperature gradient from 5 – 90 °C. Bottom left, Full CD spectrum of GdolS at increasing temperatures. Bottom right, Comparison of GdolS at 20 °C, Black; 90 °C, orange; and at 20 °C after cooled from 90 °C, purple.

CD spectra were also measured across a range of pH values to compare with the findings of the pH optimisation. There was no change to the spectrum of GdolS at pH 8 over 60 minutes. The protein was also stable over 10 minutes at pH 7 to 10, while at pH 6 the spectrum was visibly less α-helical upon the first measurement, and showed further reduction in helicity after 10 minutes, Figure 2.11. Notably, at pH 10 there is very little (if any) decrease in helicity even after 10 minutes, suggesting that while at pH 6 the reduction in rate of catalysis is due to a loss of structure this is not the case at pH 10 and may be due to some other effect on the rate of the reaction.

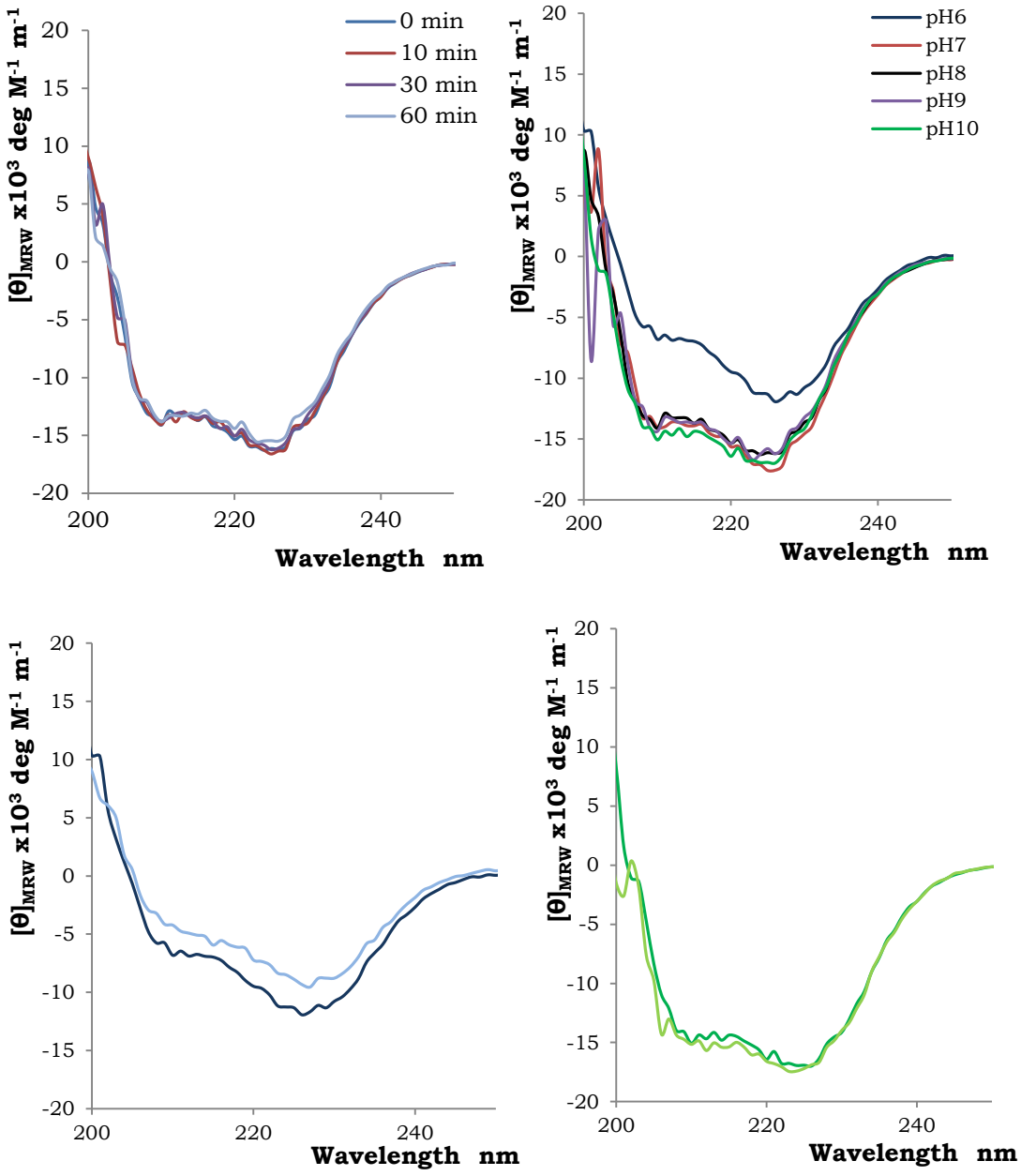
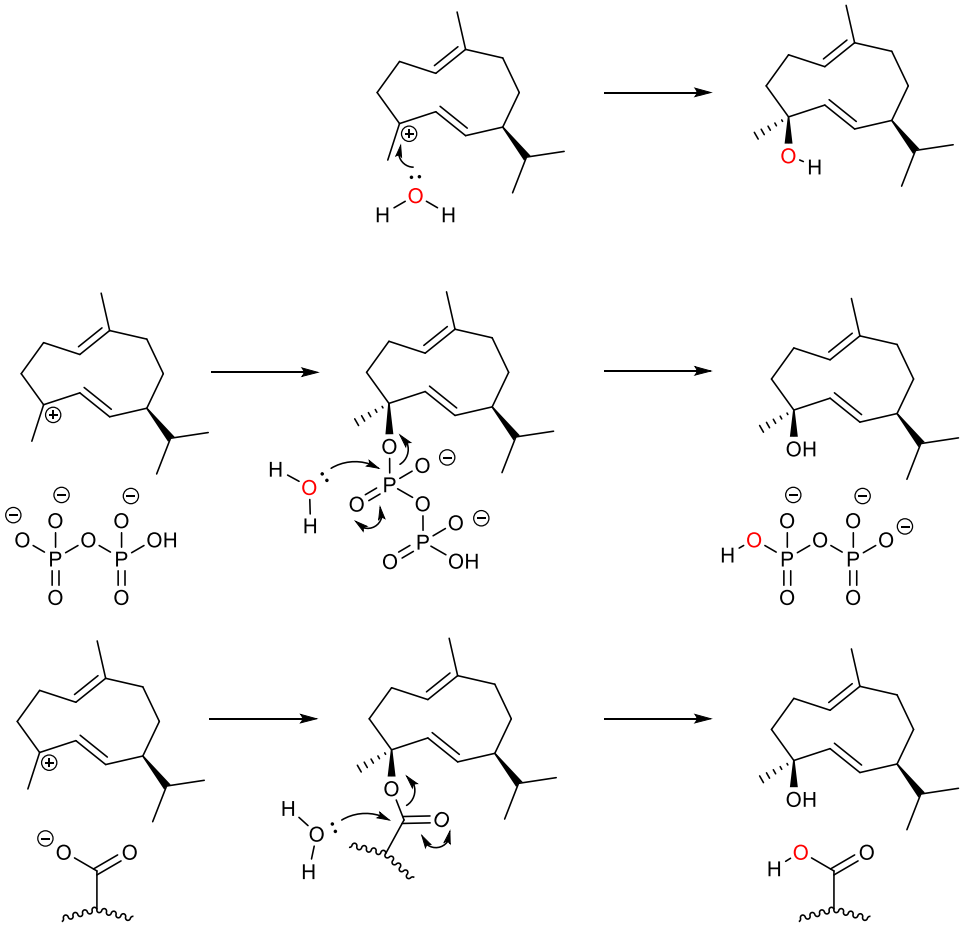


Figure 2.11 CD spectra of GdolS at different pH levels. Top left, Spectra at pH 8 at 10 minute intervals. Top right, Spectra at different pH values. Bottom, Spectra at 0 (dark) and 10 minute (light) intervals in pH 6 (Bottom left) and pH 10 (Bottom right).

2.3 Investigation of the mechanism of GdolS

2.3.1 Investigation into the origin of the C4-hydroxyl in germacradien-4-ol (79)

The hydroxyl group at C4 was proposed to originate directly from the bulk solvent, through quenching of a carbocation by an active site bound water molecule or through a channel into the active site. Other possible mechanisms for the incorporation of the hydroxyl could include involvement of an active site amino acid or the diphosphate anion, to form either an enzyme bound intermediate or a tertiary diphosphate, followed by hydrolysis to yield germacradien-4-ol, Scheme 2.1. These postulated mechanisms were investigated using incubations performed in buffers prepared using isotopologues of water.



Scheme 2.1 Possible pathways for the GdolS catalysed incorporation of water in germacradien-4-ol. Top, direct attack of a free or enzyme-bound water molecule; Middle, formation of a germacradienyl diphosphate followed by hydrolysis; Bottom, formation of an enzyme-bound germacradienyl ester followed by hydrolysis. The fate of the oxygen atom derived from the bulk solvent is highlighted in red.

Incubation in buffer prepared with D_2O (^2H_2O) showed an increase of 1 amu in the molecular ion to *m/z* 223, and an increase of 1 amu to *m/z* 208 for the $[M - CH_3]^+$ fragment, while no change was observed in any of the fragments of lower molecular weight after dehydration (*m/z* 204, $[M - H_2O]^+$).

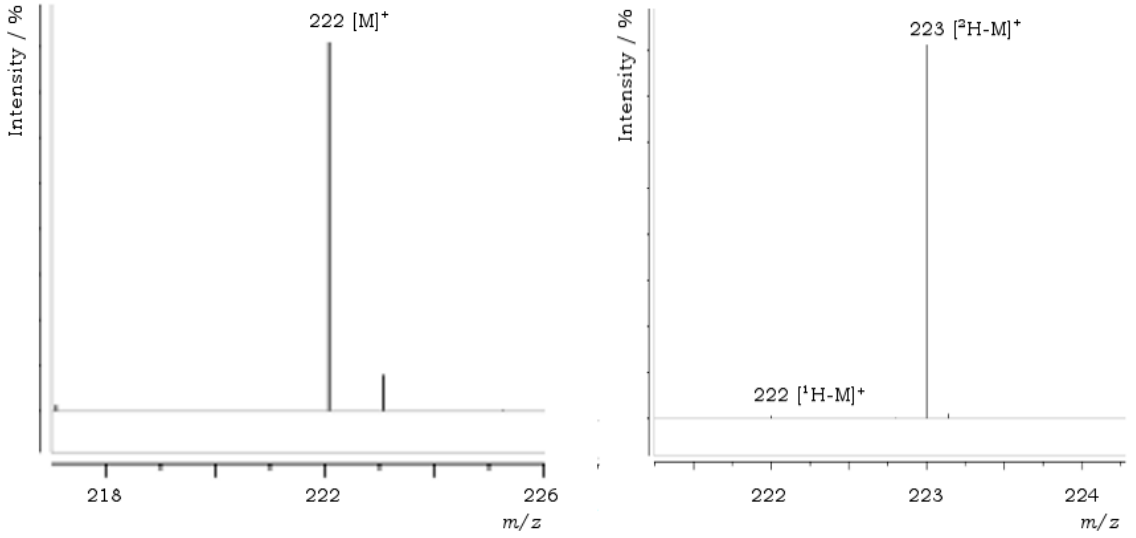


Figure 2.12 Mass spectrum of germacradien-4-ol arising from incubation of GdolS and FDP in, left, H_2O and right, D_2O buffer. Both show only expansion of the molecular ion.

This indicates that the proton on the hydroxyl group may originate from the bulk solvent; however, this could also be contributed to exchange of the proton with the solvent. To trace the origin of the oxygen atom, proposed to provide a truer indication of the origin of the hydroxyl, incubations were carried out in buffer containing 50% $H_2{}^{18}O$.

These incubations show incorporation of an ^{18}O atom in to 4-hydroxyl group of germacradien-4-ol. Analysis of the mass spectrum shows 0.75:1 incorporation of ^{16}O:^{18}O, a 57% incorporation from $H_2{}^{18}O$. indicating that the hydroxyl arises from direct quenching of the carbocation with water from the bulk solvent.

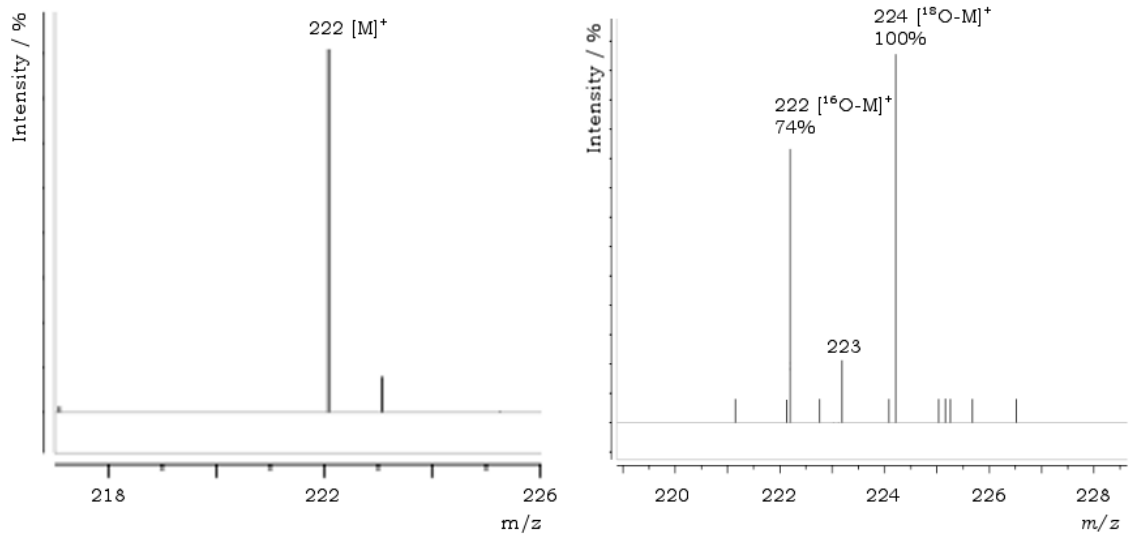


Figure 2.13 Mass spectrum of germacradien-4-ol arising from incubation of GdolS and FDP in, left, H_2O and, right, 50 % $H_2{}^{18}O$ buffer, showing expansion of the molecular ion.

2.3.2 Incubation of substrate analogues with GdolS

Incubations of GdolS with FDP (**19**) and FDP-analogues were carried out as described in methods, Section 6.1.17. 12,13-F_2-FDP (**73**) was prepared as described in Chapter 4, Section 4.2.1, while the other FDP-analogues were samples available in the lab. These analogues were primarily prepared by Drs. Juan Faraldos, David Miller and Sabrina Touchet, who have my thanks for allowing me use of their samples.

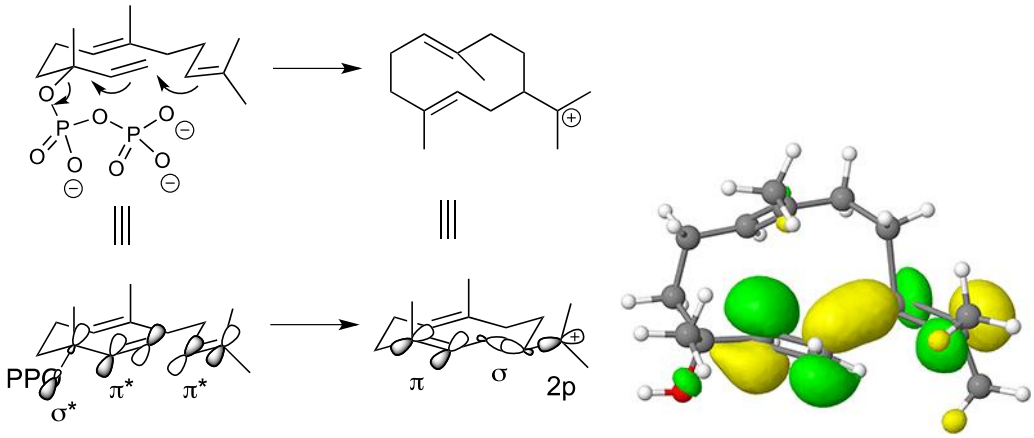
Nerolidyl diphosphate (NDP, 22)

Incubation of GdolS with the tertiary diphosphate nerolidyl diphosphate (NDP, **22**), produces germacradien-4-ol (**79**) as a single product. Measurement of the steady-state kinetic parameters with racemic-^{3}H-NDP (prepared by Dr Juan Faraldos) gave k_{cat} and K_M of 0.076 ± 0.003 s^{-1} and 1.3 ± 0.03 μM respectively, values remarkably similar to those measured for FDP (**19**). Use of the enantiopure (^{3}H)-(3*R*)-NDP (3*R*-**22**, prepared by Dr Juan Faraldos) for the measurement of kinetic parameters showed a 3-fold increase in k_{cat} paired with a 2-fold in increase in K_M (0.24 ± 0.007 s^{-1} and 2.6 ± 0.18 μM respectively) leading to a modest improvement in catalytic efficiency over the substrate FDP (**19**), Table 2.3.

Table 2.3 Comparison of steady-state kinetic measurements of turnover of [1-^{3}H]-substrates by GdolS

[1-^{3}H]-substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} μM^{-1})
FDP	0.079 ± 0.003	1.07 ± 0.13	73.8 x 10^{-3}
Rac-NDP	0.076 ± 0.003	1.30 ± 0.03	58.5 x 10^{-3}
(3*R*)-NDP	0.240 ± 0.007	2.60 ± 0.18	92.3 x 10^{-3}

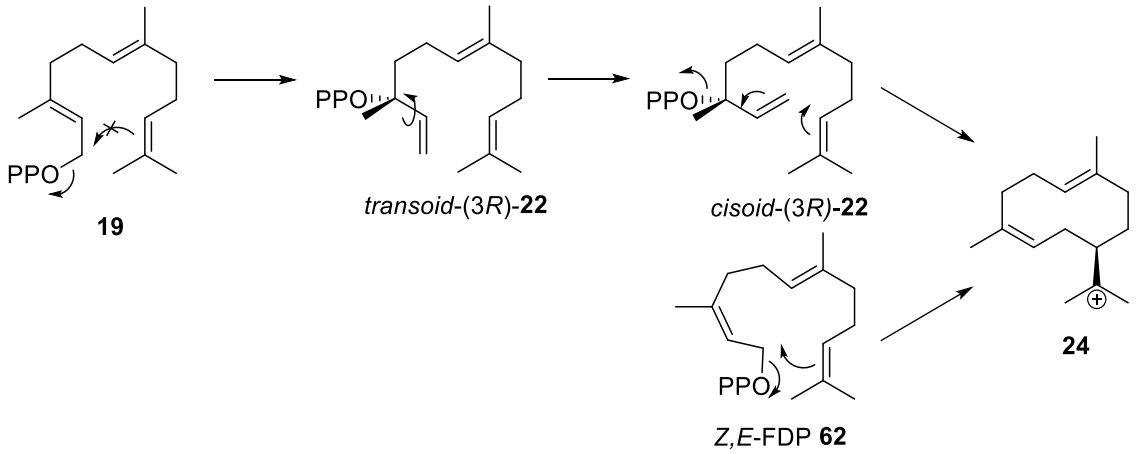
While these results may suggest that the catalytic mechanism of GdolS proceeds through isomerisation of FDP (**19**) to NDP (3*R*-**22**) prior to 1,10-ring closure this proposal seems counterintuitive due to the apparent lack of necessity for this isomerisation. The proposed active site conformation of (3*R*)-NDP (3*R*-**22**) potentially shows a favourable alignment between the 1,2-π*-anti-bonding orbital and the 10,11-π*- and 3C-O-σ* orbitals which would allow the concerted ring-closure and allylic elimination of the diphosphate in a single step, Scheme 2.2. Molecular modelling of the LUMO of (3*R*)-nerolidol (kindly performed by Dr Stefano Leoni) constrained to the active site volume shows that in this conformation the orbitals are aligned to productively drive ring closure and the allylic elimination of the diphosphate upon excitation of electrons from the HOMO. While product release is rate determining for the overall conversion of FDP by terpene synthases, the initial ionisation of FDP (**19**) is widely accepted as the rate limiting chemical step.[108,109] This configuration of NDP (**22**) is optimal to directly generate the *transoid*-germacryl (**98**) cation in a concerted manner, hence accelerating the reaction.



Scheme 2.2 Left, Active site conformation of (3*R*)-NDP (3*R*-**22**) and ring closure to the *transoid*-germacryl cation (**98**). Right, Calculated LUMO of (3R)-nerolidol confined to active site volume.

(2*Z*,6*E*)-Farnesyl diphosphate ((2*Z*,6*E*)-FDP, 62)

Incubation of GdolS with (2*Z*,6*E*)-FDP (**62**) did not solely yield germacradien-4-ol (**79**) as would be expected if isomerisation of the 2,3-double bond were necessary to allow cyclisation. Instead germacradien-4-ol (**79**) is produced as only a minor product (<3%) among a large number (>20) of other products, which are currently unidentified.



Scheme 2.3 Conversion of (*E*,*E*)-FDP (**19**) to NDP (**22**) to allow cyclisation to the *cis*-germacryl cation (**24**) and direct cyclisation of (*Z*,*E*)-FDP (**62**)

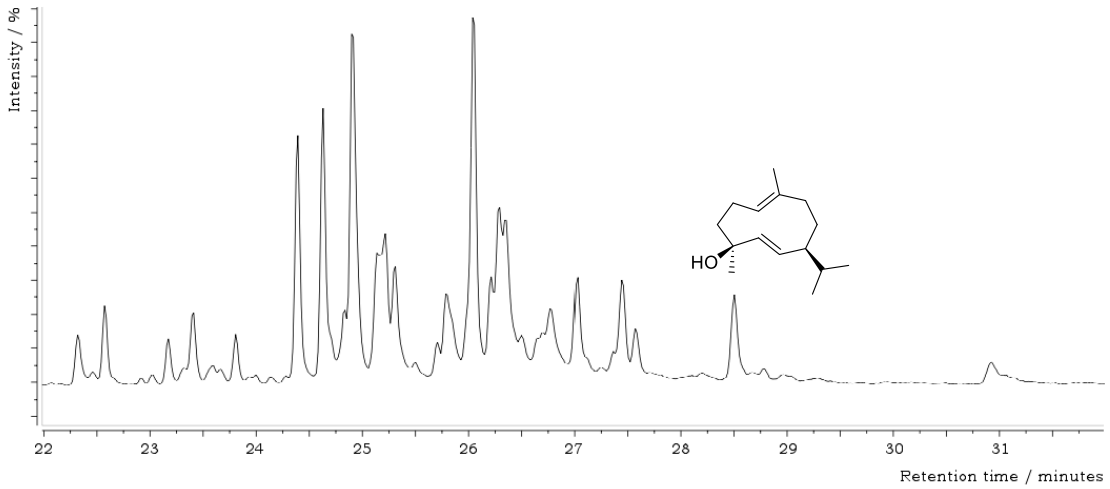


Figure 2.14 Gas chromatogram of the pentane extracted products from the incubation of GdolS with (*Z*,*E*)-FDP (**62**). The only peak positively identified is that of germacradien-4-ol (**79**, highlighted).

Geranyl diphosphate (GDP, 18)

Many sesquiterpene synthases are able to turn over the ten-carbon isoprenyl diphosphate, geranyl diphosphate (**18**), into linear and cyclic monoterpenes[44,65,66,110,111] and so geranyl diphosphate (**18**) was incubated with GdolS and the pentane extractable products analysed by GC-MS. GdolS was able to turn over GDP (**18**), albeit much less effectively than the natural substrate as determined by comparison of the total ion count (3 x 10^4, GDP; 2 x 10^6, FDP). The products were identified by comparison with genuine samples of known monoterpenoids showing primarily linear (>88%), but also small proportions of cyclic (<12%), monoterpenes; the presence of cyclic products suggests that GdolS might be catalysing the conversion of GDP (**18**) to linalyl diphosphate (LDP, **99**). However, as discussed earlier, Section 1.3.4, a recent

study by Zhang and Tiefenbacher has found allylic cations to be rotatable under ambient conditions.[43]

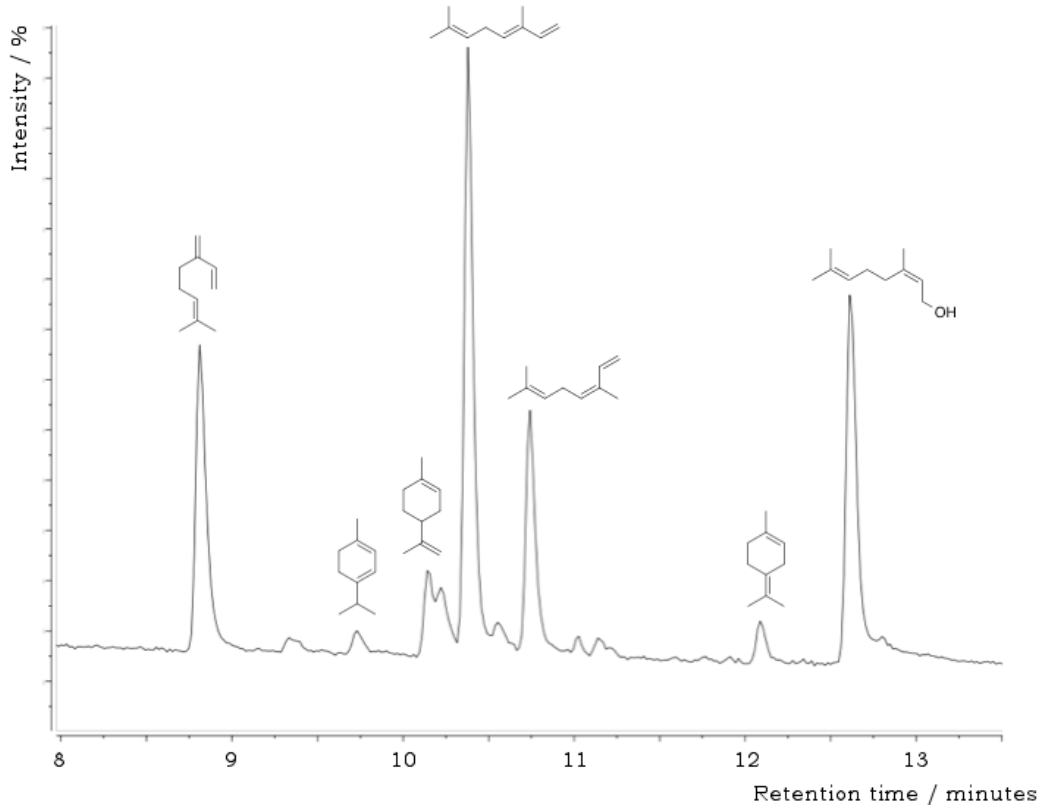
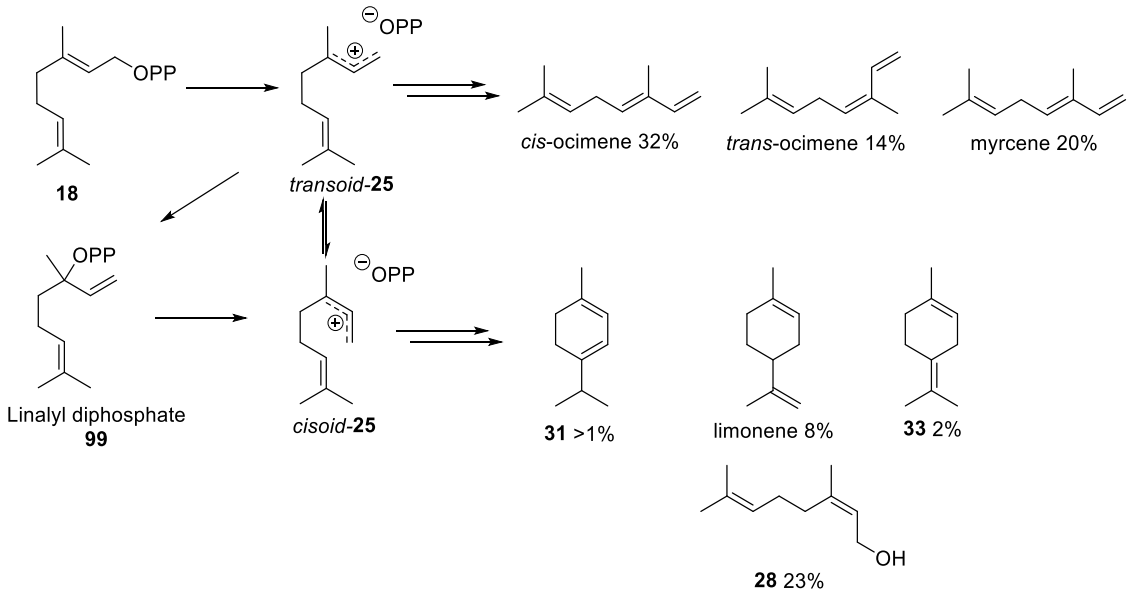


Figure 2.15 Gas chromatogram of the pentane extractable products arising from incubation of GdolS with GDP (**18**). Compounds identified by comparison with authentic samples.



Scheme 2.4 Formation of the products identified from incubation of GDP (**18**) with GdolS. Percentages are approximate to the closest significant figure.

Inhibition of GdolS by 2F-FDP (50) and [15,15,15]-F_3-FDP (100)

Incubation of GdolS with 2F-FDP (**50**) and [15,15,15]-F_3-FDP (15F_3-FDP, **100**) generated no pentane extractable products that could be detected by GC-MS

analysis, even after extended incubation times. The effect of these two FDP-analogues as inhibitors was investigated by measuring steady-state kinetic parameters in the presence of varying amounts of the fluorinated-FDP.

15F_3-FDP (**100**) was found to be a good inhibitor with a K_i of 3.9 ± 0.2 μM, comparable to the K_M for FDP (**19**); a Lineweaver-Burke double reciprocal plot of the data showed inhibition to be competitive, as is typically found with terpene synthases due to having a single active site into which binding of the diphosphate group is highly favourable in the presence of Mg^{2+}. The 2F-FDP (**50**) was found to be a potent inhibitor with a K_i of 0.3 μM, however the type of inhibition could not be clearly determined despite numerous attempts. The Lineweaver-Burk plot shows the lines intercepting before the *y*-axis and above the *x*-axis, indicating a difference in V_{max} and an effect on K_M, indicative of mixed inhibition. Mixed inhibition is rare in systems such as terpene synthases where there is a single binding site. In terpene synthases substrate binding is primarily driven by the [Mg^{2+}]-PP_i coordination and the absence of any secondary regions for binding means deriving an explanation for mixed inhibition is difficult.

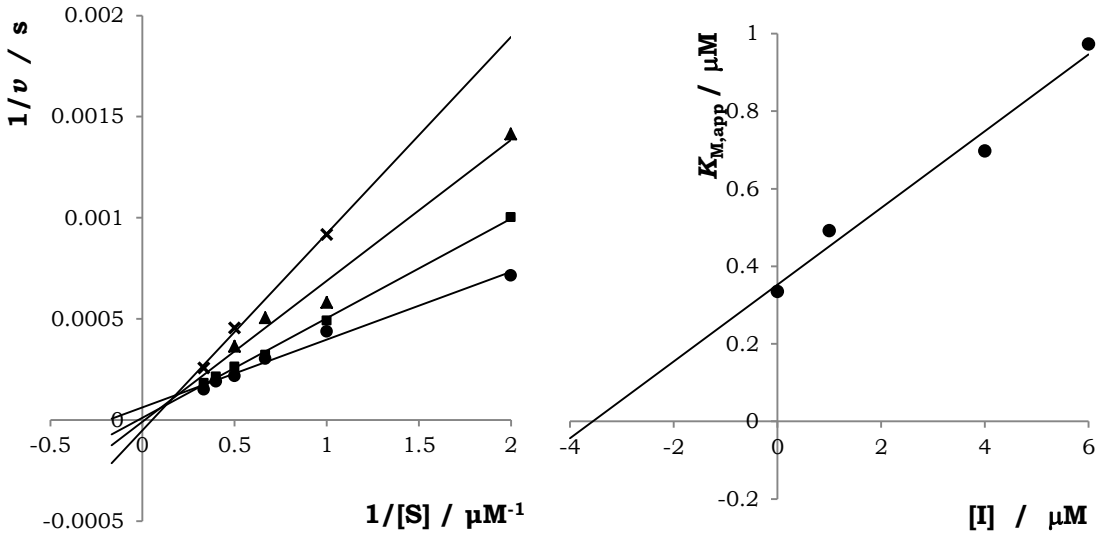


Figure 2.16 Lineweaver-Burk (left) and plot of $K_{M,app}$ against [I] for the calculation of K_i (right) for [15,15,15]-F_3-FDP (**100**), legend shows the concentration of inhibitor.

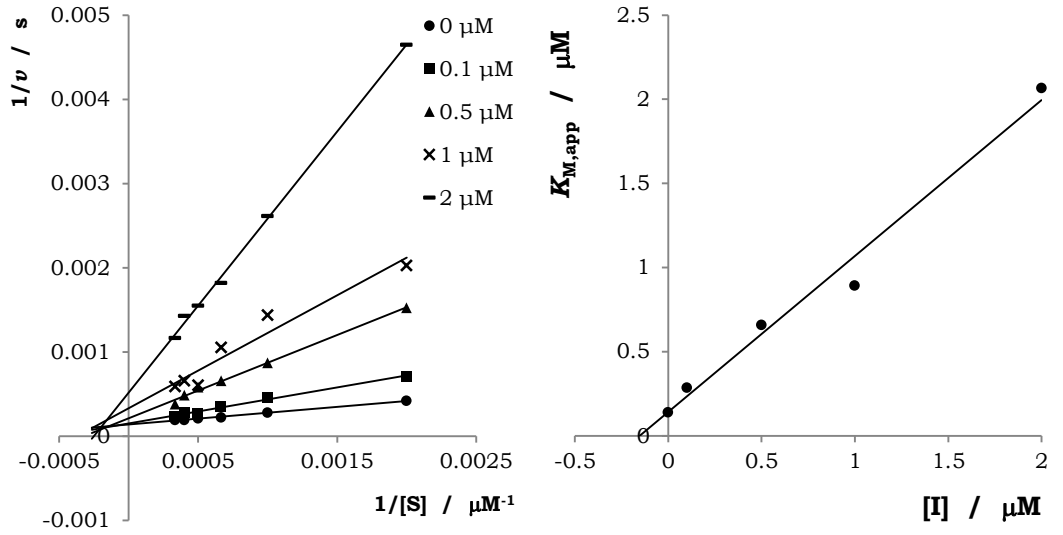
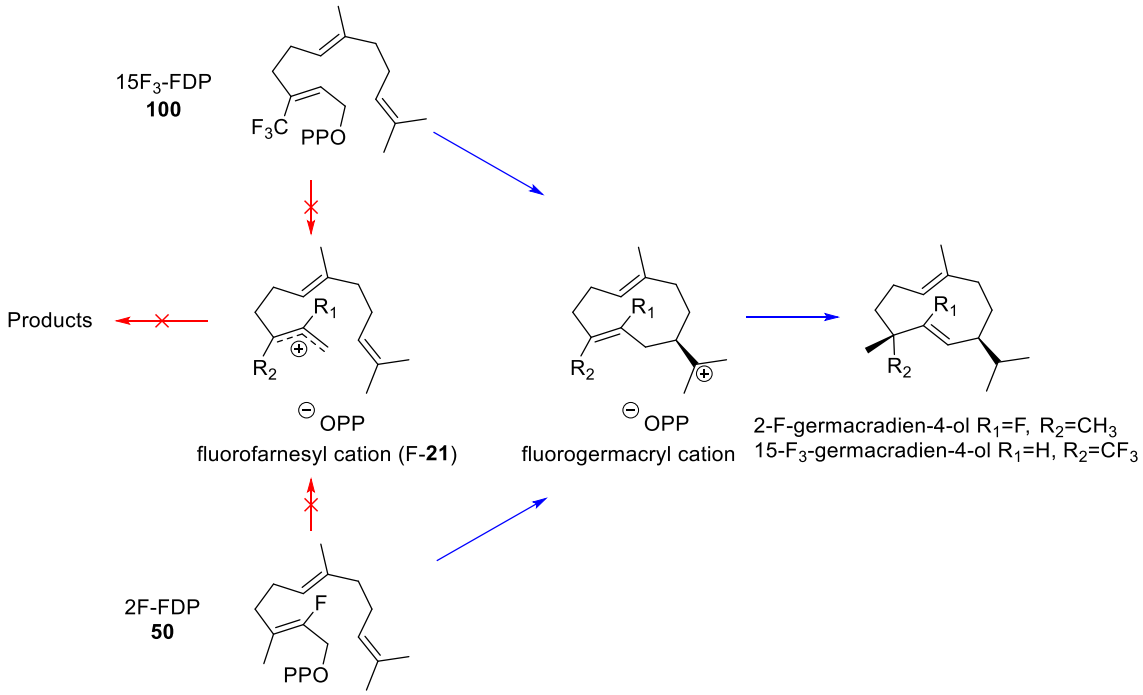


Figure 2.17 Lineweaver-Burk plot (left) and plot of $K_{M,app}$ against [I] for the calculation of K_i (right) for 2F-FDP (**50**), legend shows the concentration of inhibitor.

The inhibition of GdolS by both the 2F-FDP (**50**) and the 15F$_3$-FDP (**100**) is strong evidence for a stepwise cyclisation mechanism, whereby dissociation of the diphosphate occurs prior to the ring-closure. This dissociation generates the allylic farnesyl cation (**21**); it is this cation which is destabilised by the electron-withdrawing nature of the fluorine-substituents in the cases of the 2F-FDP (**50**) and 15F$_3$-FDP (**100**) and formation of this cation is prevented, and the cyclisation is unable to proceed.



Scheme 2.5 Abortive stepwise (red) and productive concerted (blue) cyclisation mechanisms with 2-F-FDP (**50**) and 15F$_3$-FDP (**100**). Intermediates where R$_1$=F, R$_2$=CH$_3$ arise from 2-F-FDP (**50**), where R$_1$=H, R$_2$=CF$_3$ arise from 15F$_3$-FDP (**100**).

GdolS catalysed turnover of 12,13F_2-FDP (73)

GC-MS analysis of the pentane extractable products from incubation of 12,13F_2-FDP (**73**) with GdolS shows a single peak. Analysis of the mass-spectrum is consistent with this peak being the 12,13-*E*-β-farnesene (**77**), arising from inhibition of the 1,10-ring closure. This is supported by the production of the same product from incubation of 12,13-F_2-FDP with EBFS, Figure 2.18.

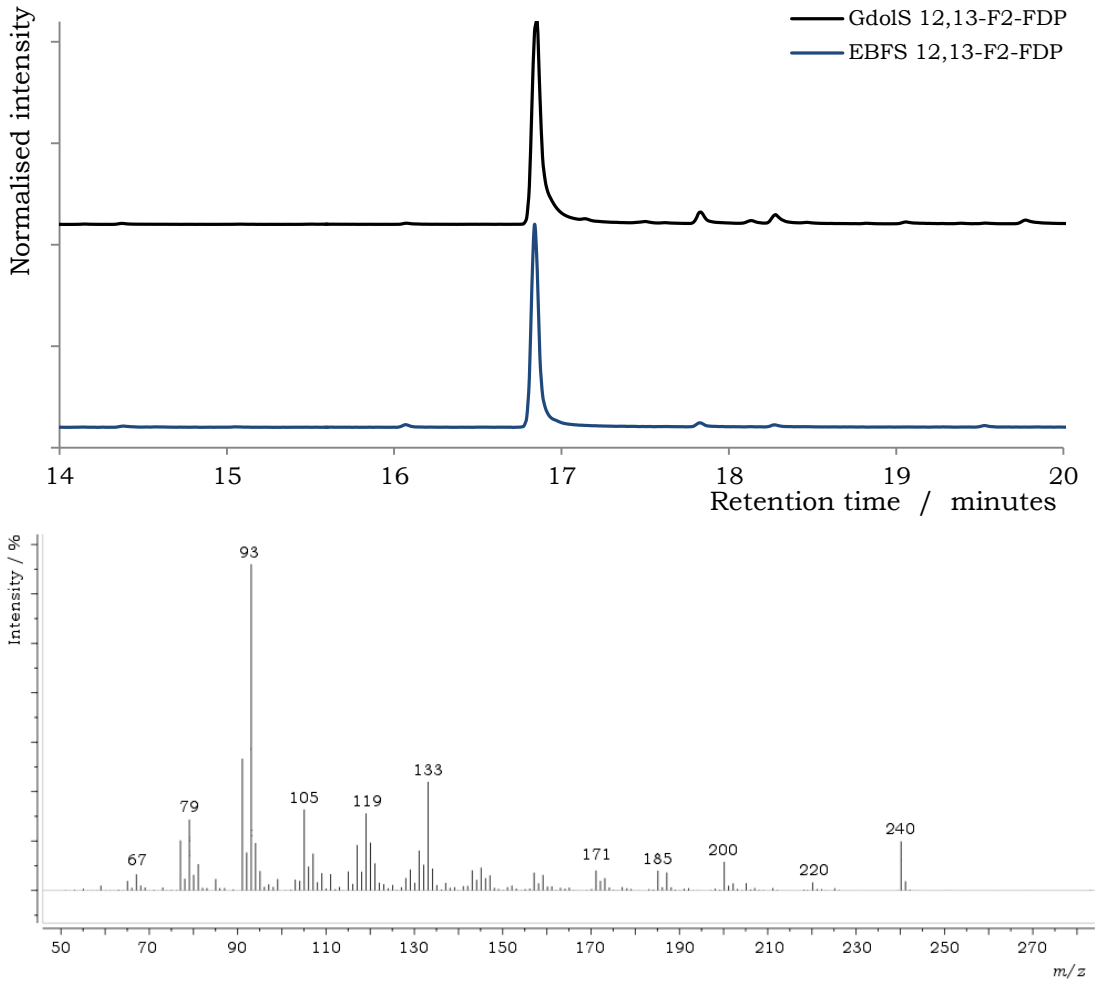
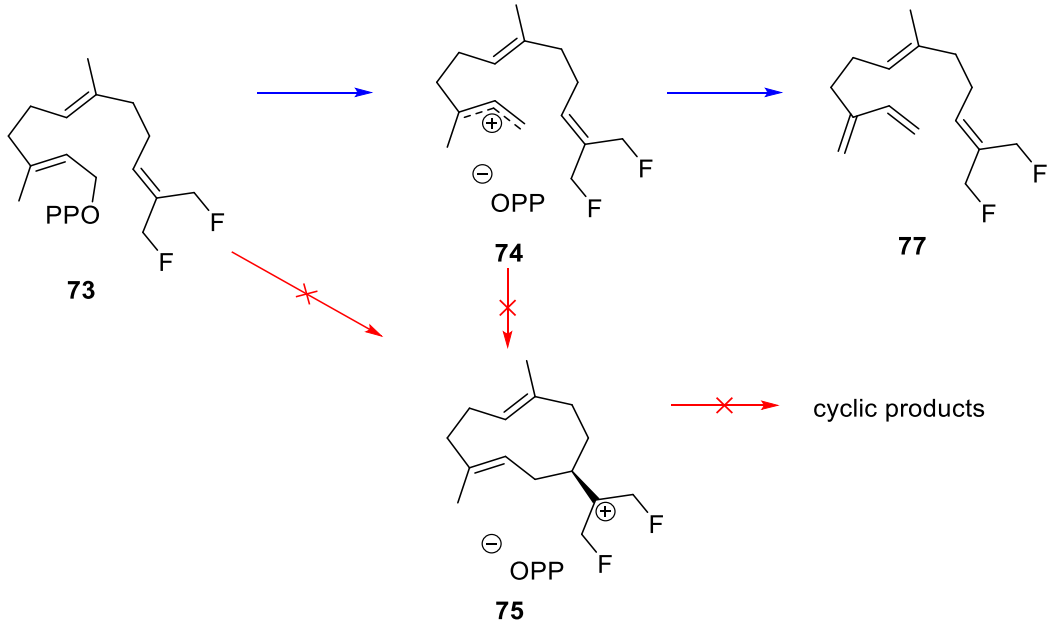


Figure 2.18 GC-MS analysis of the pentane extracted products of the incubation of GdolS with 12,13F_2-FDP (**73**).

Formation of the 12,13-difluorofarnesyl cation (**74**) is not destabilised and so dissociation of the diphosphate is able to proceed; the subsequent 1,10-ring closure would generate the 12,13-difluorogermacryl cation (**75**) and this would be strongly destabilised by the presence of the fluorines on C12 and C13 and so this ring closure does not occur; instead direct deprotonation leads to the farnesene derivative (**77**). The formation of linear instead of cyclic products could also be explained by a reduction in the nucleophilicity of the 10,11-double bond due to the electron withdrawing nature of the fluorine

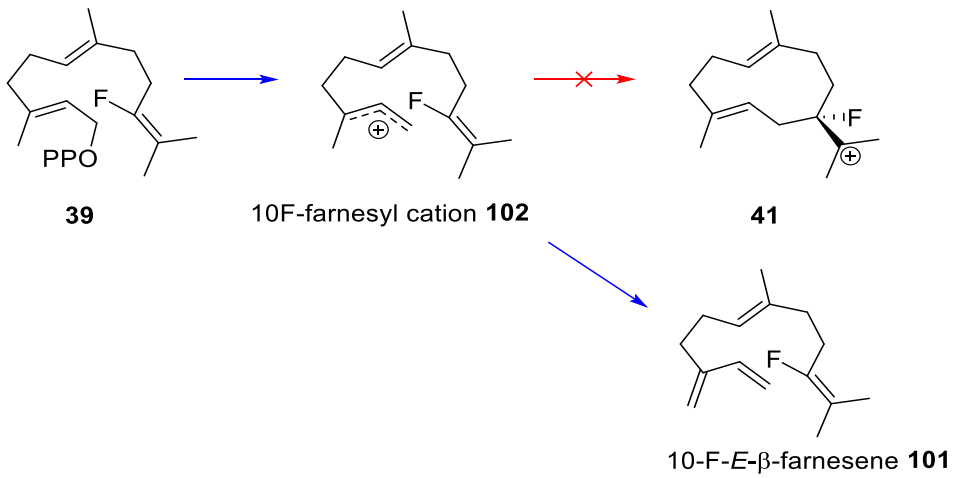
substituents, but this would still be in support of a stepwise mechanism with the intermediacy of the 12,13-difluorofarnesyl cation.



Scheme 2.6 Productive stepwise (blue) and abortive concerted GdolS catalysed cyclisation mechanisms with 12,13F_2-FDP (**73**).

GdolS Catalysed Cyclisation of 10F-FDP (39)

Incubation with 10F-FDP (**39**) was expected to yield a similar result as that with 12,13F_2-FDP (**73**), producing the farnesene derivative due to destabilisation of the 10F-germacryl cation (**41**) by the fluorine in the 10-position, Scheme 2.7. However, analysis of the pentane extractable products by GC-MS showed three main products. Two of these peaks were identified as 10F-*E*-β-farnesene (**101**) and 10-F-α-humulene (**40**); a product of a 1,11-cyclisation. These peaks were identified by coelution with enzymatically produced authentic samples using GDS(Section 1.5.2), Figure 2.19.[65,79] The third peak was not identified.



Scheme 2.7 Anticipated results from incubation of GdolS with 10F-FDP.

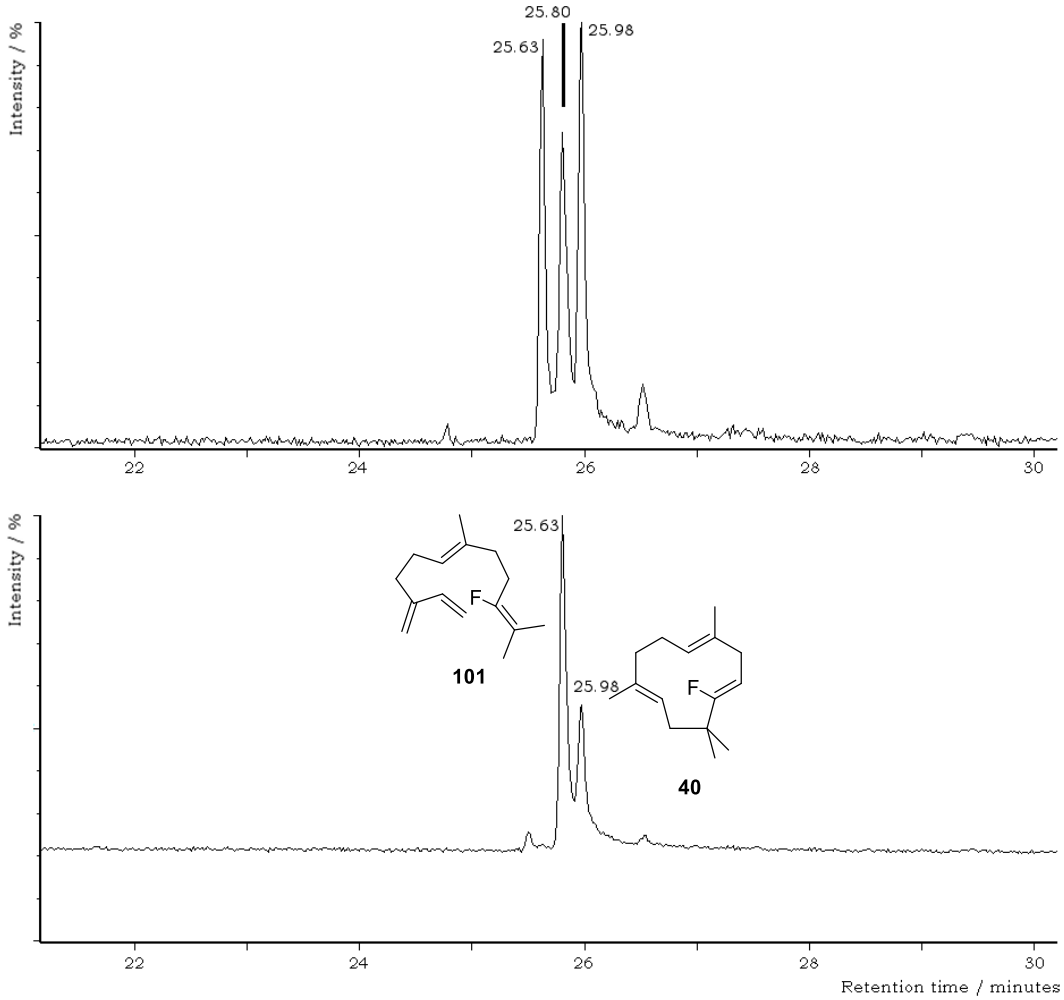
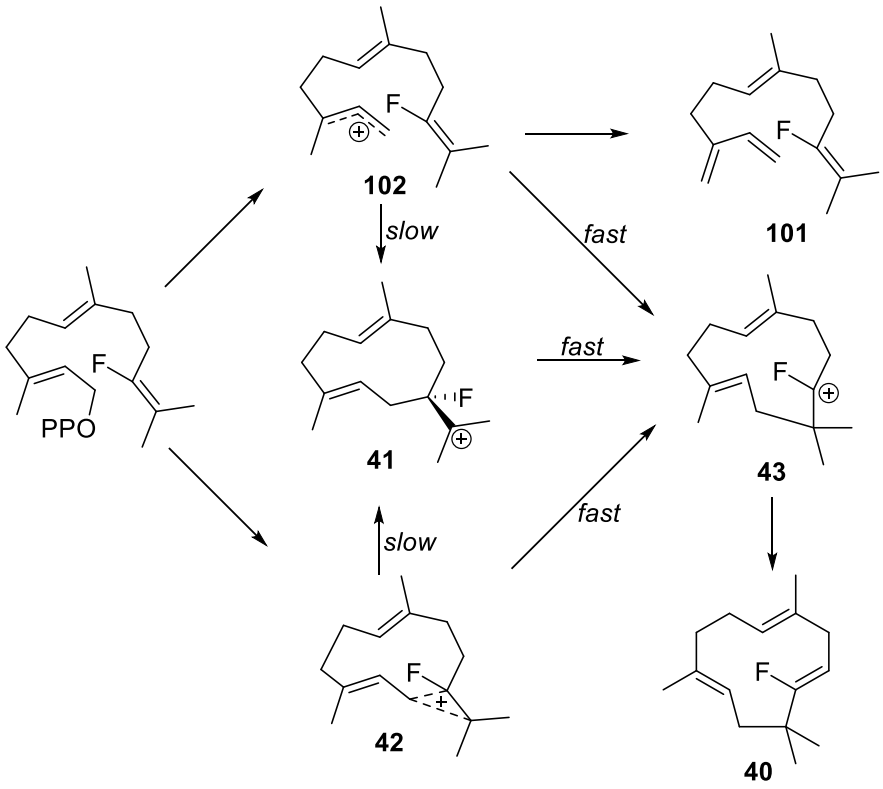


Figure 2.19 Gas chromatogram of the pentane extractable products from inubation of 10F-FDP (**39**) with GdolS, top, and GDS, bottom.

Measurement of the steady-state kinetic parameters with tritiated 10F-FDP ([1-^{3}H]-**39**) show that while the rate of the 1,11-cyclisation is slower than the natural cyclisation it is still able to turn over the 10F-FDP (**39**) with modest efficiency (k_{cat} 0.018 ± 0.0009 s^{-1}, K_M 0.6 ± 0.09 μM), Figure 2.20. The ability of the enzyme to catalyse a 1,11-cyclisation with this substrate-analogue could be explained by a number of possible mechanisms after the dissociation of the diphosphate; a direct 1,11-ring closure of the 10F-farnesyl cation (**102**) yielding the 10F-humulyl cation (**43**) which is now stabilised due to the α-effect of the fluorine substituent, initial slow 1,10-ring closure forming the destabilised 10F-geramcryl cation (**41**) followed by rapid ring-expansion to the stabilised 10F-humulyl cation (**43**) or ring closure via a bridged cyclopropyl cation (**42**) which then rapidly collapses to the 10F-humulyl cation (**43**). Deprotonation of the humulyl cation then yields the 10-F-α-humulene (**40**).



Scheme 2.8 Possible pathways for GdolS catalysed cyclisation of 10-F-FDP (**39**) to 10F-α-humulene (**40**).

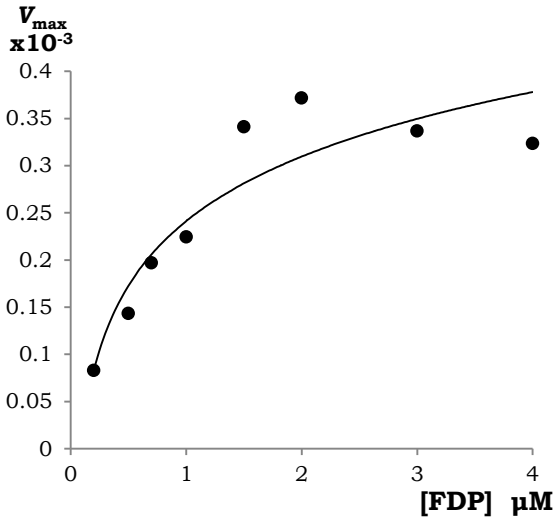
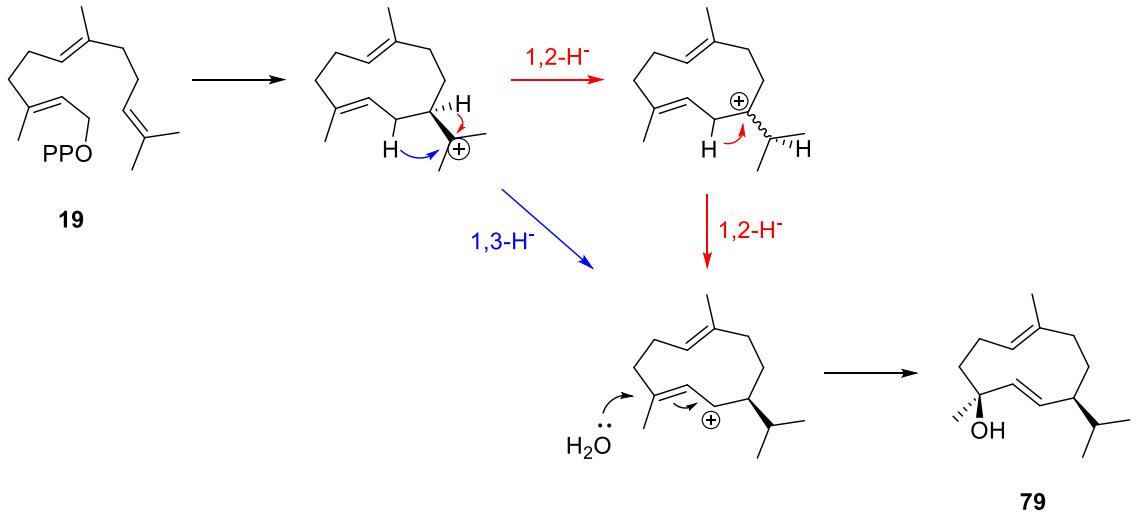


Figure 2.20 Representative graph for the calculation of steady-state kinetic parameters of GdolS with 10F-FDP (**39**)

Analysis of the GdolS catalysed turnover of deuterated-FDPs

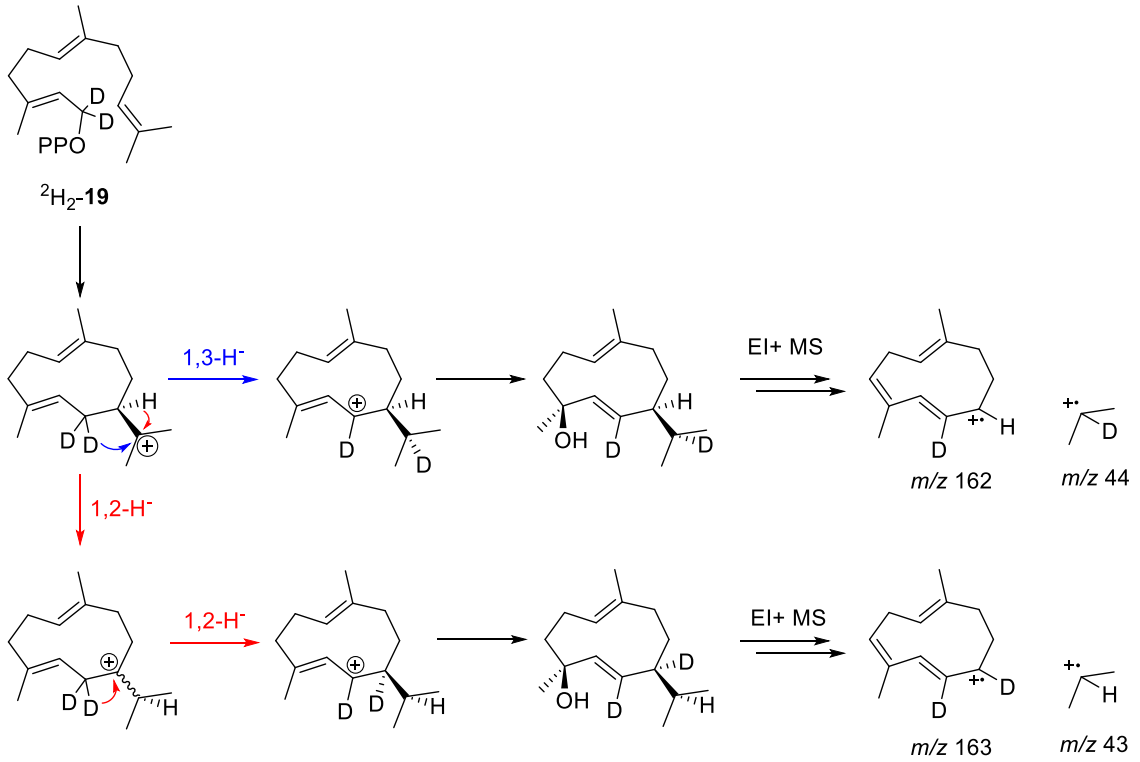
While Nakano *et al.* proposed a 1,3-hydride shift in the mechanism catalysed by germacradien-4-ol synthase,[93] it is also possible that two consecutive 1,2-hydride shifts are involved in the transfer of the carbocation from C11 to C6, Scheme 2.9. This mechanism of consecutive 1,2-hydride shifts has been shown to be that catalysed by the (*R*)-germacrene D synthase by incubating

the enzyme with the FDP substrate bearing deuterium atoms at the 1-position.[64]



Scheme 2.9 Possible 1,3- and double 1,2-hydride shifts in the mechanism catalysed by germacradien-4-ol synthase.

Germacradien-4-ol synthase was incubated with 1-$^2H_2$-FDP ($^2H_2$-1**9**) and the pentane extractable products were analysed by GC-MS showing a single peak for germacradien-4-ol. By analysing the fragmentation in the mass spectrum of the product, in particular the fragment derived from the loss of the isopropyl tail; *m/z* = 161 for the non-deuterated FDP, we can infer the positions of the deuterium atoms. If the 1,3-hydride shift is in operation, then the product (**79**) will bear one deuterium atom on the isopropyl tail and one deuterium on the macrocycle, resulting in *m/z* = 162. In the case that a double 1,2-hydride shift is in operation the isopropyl tail will contain a hydrogen atom and both deuterium atoms will be on the ring, resulting in m/z = 163, Scheme 2.10.



Scheme 2.10 Possible mechanisms for the GdolS catalysed conversion of 1-$^2H_2$-FDP.

The mass spectrum of the $^2H_2$-**79** produced by incubation of GdolS with $^2H_2$-**19** shows a molecular ion of m/z = 224, an increase of 2 amu as expected from the incorporation of two deuterium atoms. The $[M - H_2O]^+$ also shows an increase of 2 amu, m/z = 206. The fragment derived from the loss of the isopropyl tail shows m/z = 162, indicating the incorporation of a single deuterium atom on the ring and supporting the 1,3-hydride shift mechanism.

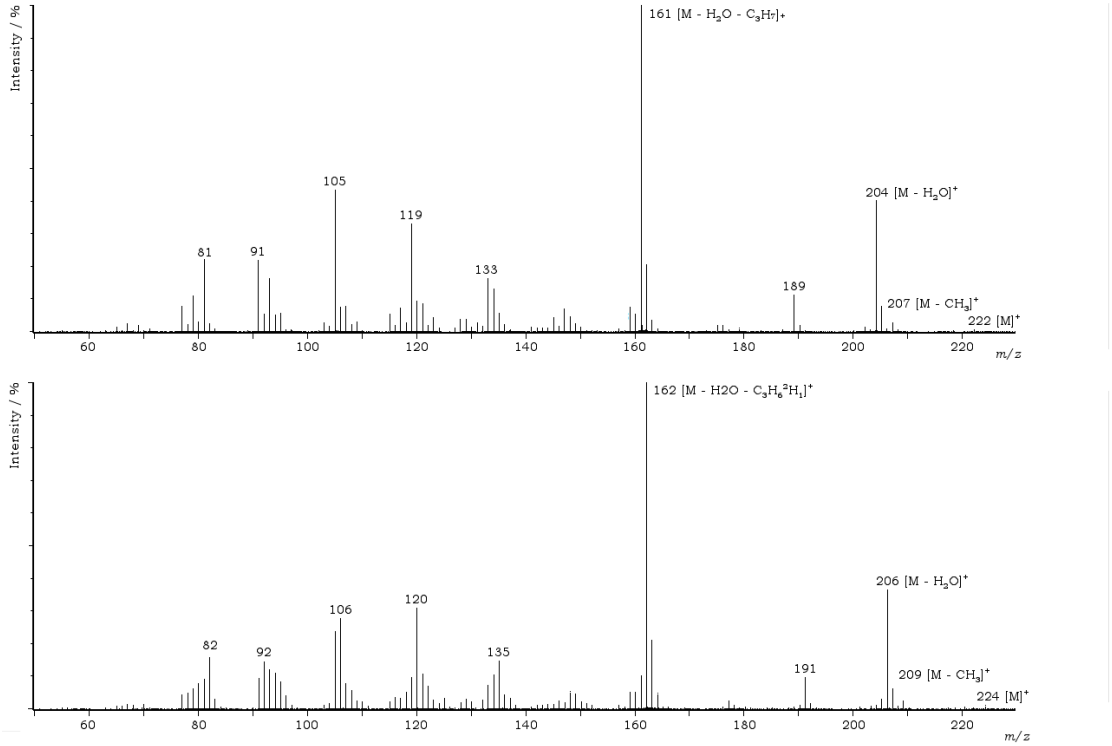
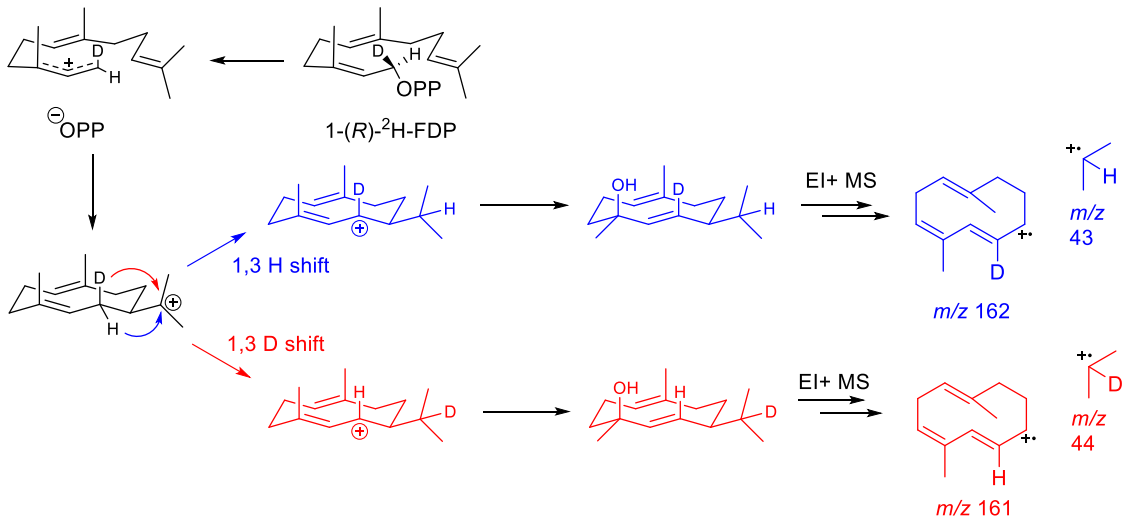


Figure 2.21 Mass spectrum of germacradien-4-ol produced by incubation of GdolS with **19** (Top) and $^{2}H_{2}$-**19** (Bottom).

This step was further investigated using a mono-deuterated-FDP, 1-^{2}H-*(R)*-FDP. The pentane extractable products of the incubation of GdolS with the deuterated-FDP were analysed by GC-MS showing a single peak for germacradien-4-ol (**79**). The mass-spectrum and in particular the fragmentation pattern imparts information about which hydride has been transferred; by looking at the fragment peak from the loss of the isopropyl tail we can see which fragment bears the deuterium. For example in the incubation of DCS with the deuterated-FDP there is clear enrichment of the peak at *m/z* = 162, indicating that the 1-*(S)*-^{1}H is transferred to the isopropyl group rather than the 1-*(R)*-^{2}H.[65]



Scheme 2.11 Possible reaction pathways for GdolS catalysed conversion of 1-*(R)*-^{2}H-FDP.

However, in the incubation of GdolS with 1-(*R*)-^{2}H-FDP there is no clear enrichment in either *m/z* 161 or 162. This indicates that either the hydride shift is not stereospecific, which seems highly unlikely due to the nature of terpene synthases, or that there is scrambling of the stereochemical information by a different mechanism.

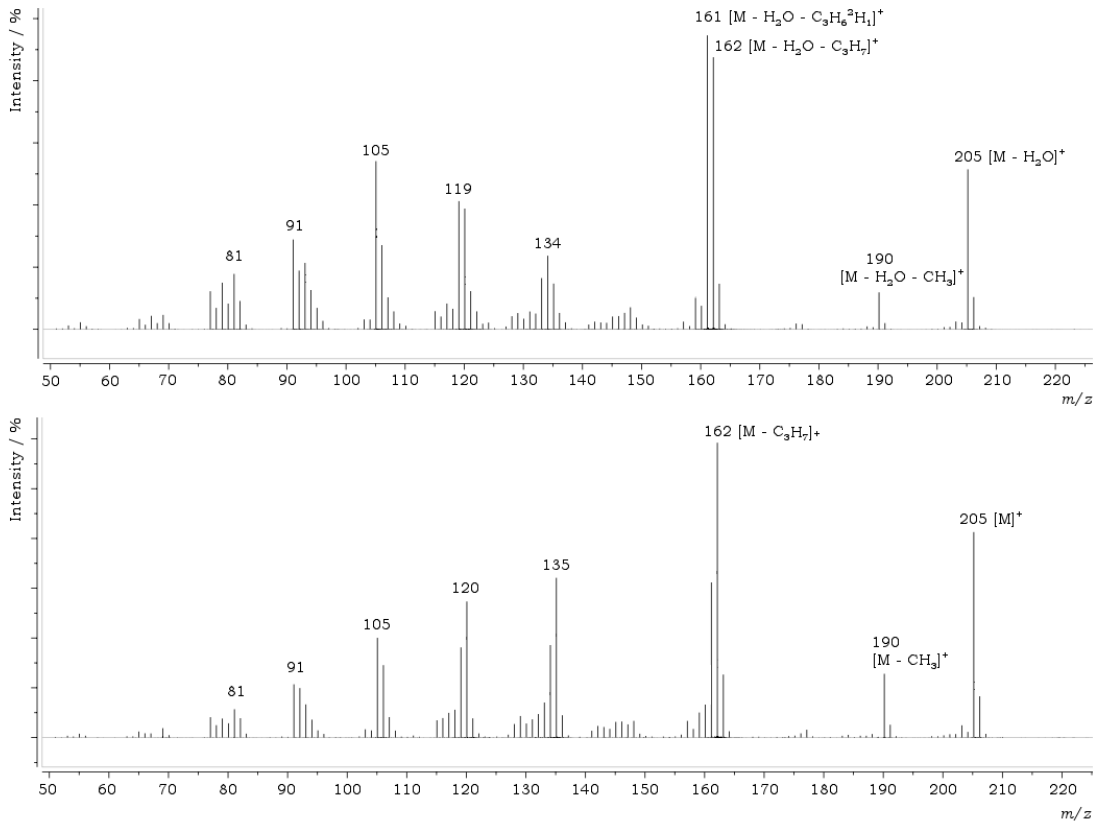


Figure 2.22 Mass spectra of the product peak, from the incubations of 1-*(R)*-^{2}H-FDP with GdolS (top) and DCS (bottom).

The carbocation formed by the hydride shift is stabilised across C1 to C3 by allylic transposition and so it could be proposed that this intermediate is relatively long lived and an equilibrium between this stabilised carbocation and the germacryl cation would scramble the stereochemical evidence of the initial hydride shift leading to no overall enrichment of either *m/z* upon fragmentation of the isopropyl group, unfortunately causing the loss of any stereochemical information about this step. The interconversion of germacrenoids between different conformers has previously been reported.[112,113,114,115] If the allylic cation is sufficiently long-lived with respect to the rate of conformational change of the 10-membered ring this could allow realignment of the vacant p-orbital and cause scrambling of the stereochemistry of the shift.

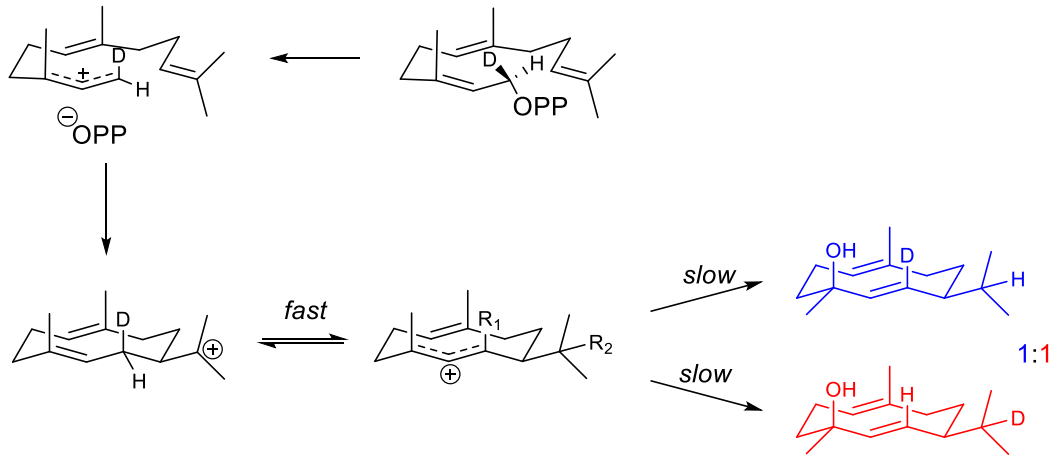


Figure 2.23 Scrambling of stereochemical information due to rapid germacryl-cations equilibria. $R_1 \neq R_2$ = D or H

2.4 Summary

In summary, this chapter has focused on a sesquiterpene synthase, germacradien-4-ol synthase, from *Streptomyces citricolor*. The enzyme was produced by overexpression in *E. coli* as previously described, and characterised in terms of stability and kinetic parameters.

Prior to the measurement of steady-state kinetic parameters the conditions for the assay were optimised and it was found that the activity was linear at concentrations below 50 nM; the concentration used for further measurements was 20 nM. In the range tested, 0.5 to 10 mM, the concentration of magnesium ions had little effect on the activity of the enzyme and 2.5 mM was used. The enzyme was structurally and catalytically stable to a range of pH values above pH 7 as determined by rate and α-helicity measured by circular dichroism spectroscopy; extended incubation at even pH 10 had no effect on

the helicity of the protein. The kinetic parameters were measured using a radiolabelled-substrate assay commonly used in the characterisation of terpene synthases; while the results obtained with this method differed from those originally reported by Nakano *et al.*[93] using a GC-MS analysis method, differences between steady state kinetic parameters based on the method used have been reported for other terpene synthases.[103,104]

The identity of the product from incubation of GdolS with FDP was analysed by GC-MS and while the mass spectrum produced differed from that previously reported, sufficient product was obtained to unequivocally determine the structure as germacradien-4-ol through the use of NMR spectroscopy. Analysis of the product by gas chromatography with a chiral column, strongly indicated that the product was stereochemically pure (-)-germacradien-4-ol, which corresponds to the (4*S*, 7*S*)-stereoisomer, according to the report by Nordin *et al.*.[105]

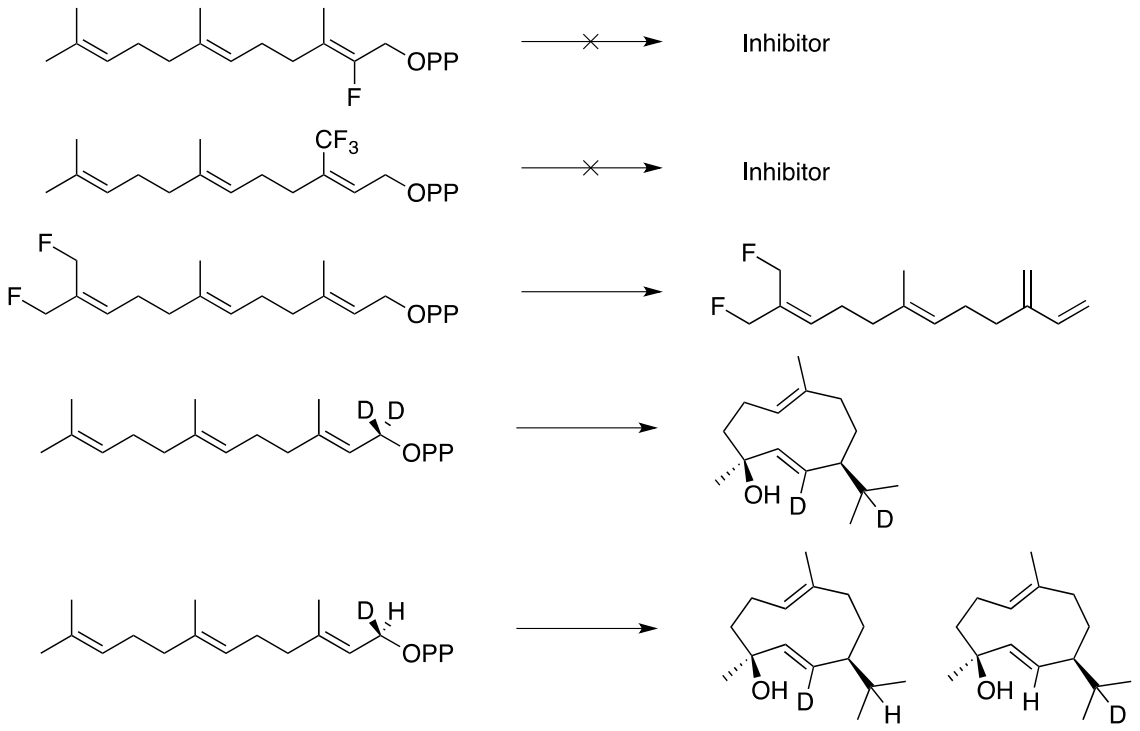
Incubations of GdolS with FDP in buffers prepared with $^{2}H_2O$ and $H_2{}^{18}O$ indicated that both the proton and oxygen in the 4-hydroxyl group originate from the bulk solvent suggesting a simple mechanism of carbocation quenching by water for the incorporation of the hydroxyl group. Due to the highly reactive nature of the carbocation intermediates in cyclisation by terpene synthases these enzymes typically function by desolvating the active site upon formation of the Michaelis complex and so the mechanism of this incorporation was probed further as described in Chapter 3.

GdolS was incubated with NDP, both racemic and the (3*R*)-stereoisomer, and the products isolated and the steady-state kinetic parameters suggested that the enzyme catalysed cyclisation may go *via* the formation of the tertiary diphosphate NDP. However, incubation with (2*Z*,6*E*)-FDP, which can directly cyclise to the *cisoid*-germacryl cation germacradien-4-ol amongst numerous other, unidentified, products.

Fluorinated FDP analogues 2F-FDP and $15CF_3$-FDP were found not to be substrates for GdolS, and proved to be potent inhibitors, pointing to the intermediacy of the farnesyl cation, and indicating a stepwise mechanism in the GdolS catalysed cyclisation of FDP. This was further supported by the ability of GdolS to turnover $12,13F_2$-FDP to the farnesene derivative, through ionisation to the farnesyl cation followed by deprotonation due to inability to then undergo cyclisation because of the destabilisation of the germacryl cation

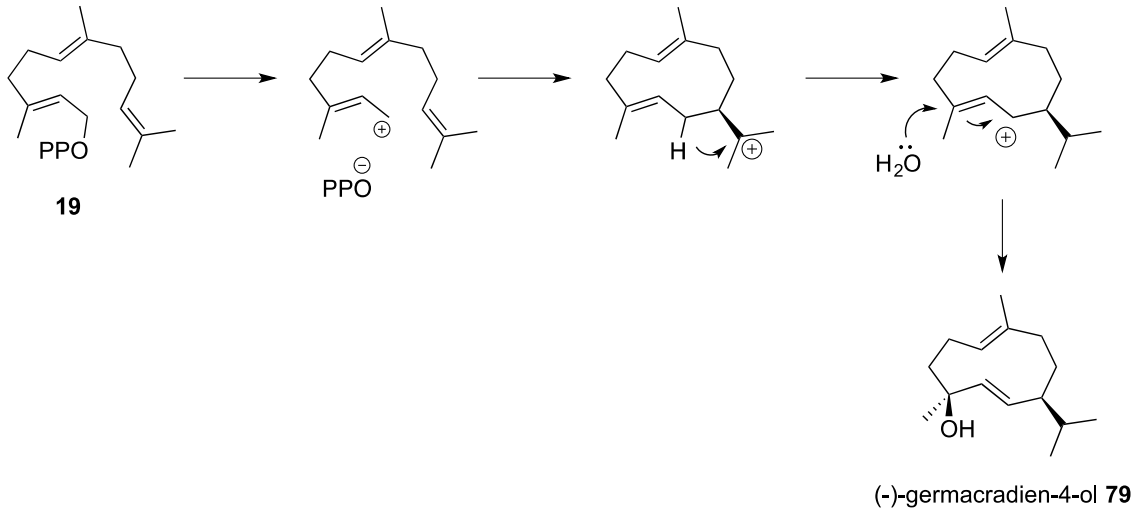
by the fluorine substituents. However, this result could also be attributed to a reduction in nucleophilicity of the 10,11-double bond. While a similar result was expected from incubation with 10F-FDP, GdolS was also able to turn this analogue over to the γ-humulene derivative, a product with an 11-membered ring. While a number of processes including a 1,11-ring closure, a 1,10-ring closure followed by rapid 1,2-alkyl shift to the more stable humulyl cation or the formation of an intermediate cyclopropyl carbocation could be proposed to explain this phenomenon, none can be favoured with the data available. Nonetheless, this ability of the enzyme to circumvent the unfavoured 1,10-cyclisation to form an 11-membered product highlights the remarkable versatility of the terpene synthases.

Analysis of the product from incubation of GdolS with 1-$^{2}H_{2}$-FDP demonstrated that the enzyme catalyses a 1,3-hydride shift subsequent to the ring closure. Incubation of GdolS with 1-(*R*-$^{2}H_{1}$)-FDP to probe the stereochemical course of the 1,3-hydride shift led, somewhat surprisingly, to a product resulting from no stereochemical bias in the 1,3-hydride shift. This was hard to reconcile with the stereospecific nature of terpene synthases and a relatively long-lived stabilised cation in rapid-equilibrium with the germacryl cation was proposed to explain this discovery; possibly with the enzyme allowing water ingress to the active site as a unintended consequence of loop movement to release the products from the active site.



Scheme 2.12 Summary of key results from incubation of GdolS with FDP analogues

In conclusion we propose that the reaction catalysed by GdolS goes *via* a stepwise mechanism for the loss of the diphosphate, to give the germacryl cation. An equilibrium between the tertiary germacra-11-yl cation and the allylic germacra-6-yl cation is mediated by a rapid 1,3-hydride shift, prior to the quenching at the 4-position by water, which originates from the solvent.



Scheme 2.13 Proposed mechanism for the cyclisation catalysed by GdolS

3 Structure–function investigation into germacradien-4-ol synthase (GdolS) by site-directed mutagenesis

3.1 Preface

This chapter will discuss the effects of a number of amino acid substitutions on the catalysis by GdolS. As discussed earlier, Section 1.5.5, it has been reported that substitutions to the metal binding domains in PR-AS have a detrimental effect on the rate of reaction and affect the distribution of products,[71] with similar effects reported in investigations with other sesquiterpene synthases.[70,72] A number of GdolS-mutants were generated to investigate the metal-binding motives and to attempt to identify residues involved in incorporating water during the catalytic cycle.

Mutants were generated using mismatched DNA primers and the polymerase chain reaction to introduce mutations to the cDNA, Section 6.1.6, the mutated and confirmed DNA was then used to express the gene with the same conditions used for the wildtype protein in the first instance, Section 6.1.10.

The use of X-ray crystal structures, particularly when substrate is bound, can be valuable in identifying active site residues involved in the catalytic cycle and used to direct the choice of residues for investigation. These data can be used to complement information gained from sequence alignments with other sesquiterpene synthases, in particular, those which have been extensively studied previously.

In the absence of a crystal structure, homology models can be generated using the protein sequence and crystal structures of terpene synthases that have previously been resolved. Due to the highly conserved tertiary structure found with the terpene synthases this can be a particularly effective starting point in the analysis of the active site in these enzymes.

3.2 Crystal structure and homology models

In collaboration with Professor David Christianson, University of Pennsylvania, we have obtained the X-ray crystal structure of GdolS in the apo-state; that is, without any ligands bound. In this structure the active-site loops, D165-T175 and S222-E248, are unrefined due to the inherent mobility of these loops in the unliganded form, and so the structure is of limited use. In addition to the crystal structure we have also made use of homology models generated using

SWISS-MODEL and I-TASSER.[116,117] Both of these methods are based primarily on the published crystal structure of a pentalenene synthase mutant (PS-N219L, pdb: 1hm7), but I-TASSER uses multiple templates to generate a model and also predicts structure in regions where there is no crystal structure to model it on. Both methods produce very similar models, in good agreement with other terpene synthases and indeed the refined regions of the GdolS-apo structure; the main differences are in the helix-termini and the loop regions. Due to the inherent flexibility necessary in these regions during the catalytic cycle, (Section 1.3.3) this mobility is unsurprising.

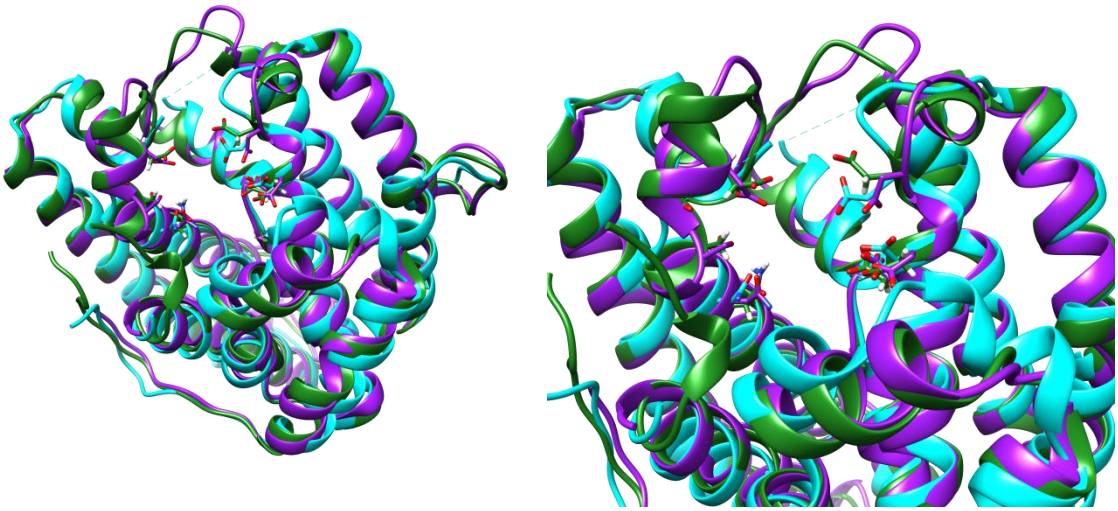


Figure 3.1 Superimposed cartoon representations of GdolS-apo (cyan) and two homology models, SWISS-MODEL (purple) and I-TASSER (green). Metal binding residues are shown with carbons in the respective ribbon colour, oxygens in red, nitrogens in blue and hydrogens in white.

Figures in the remainder of this document will primarily be based on the SWISS-MODEL homology model, or the apo-crystal structure where appropriate.

3.3 Metal-binding regions

GdolS has both conventional metal binding regions; an aspartate rich region, ^{80}**DD**QF**D**; and a NSE motif on the facing helix, ^{218}**N**DVR**S**FAQ**E**.

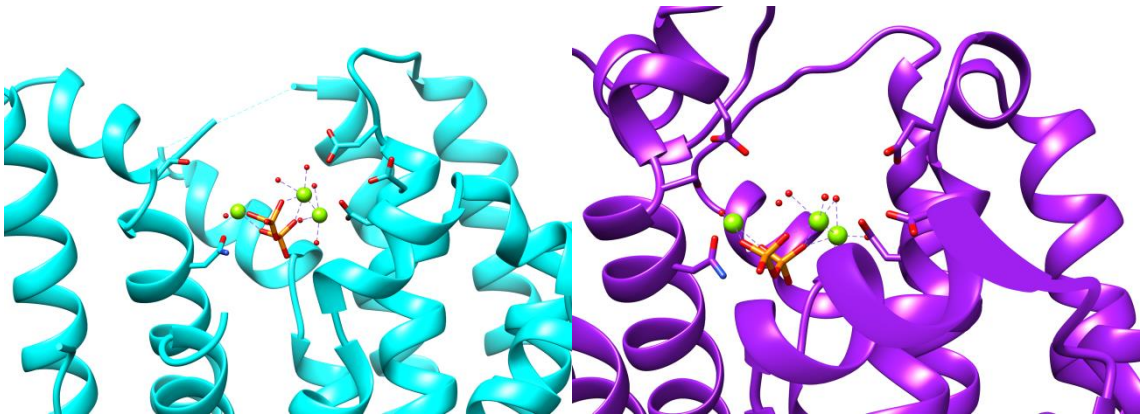


Figure 3.2 Metal binding regions shown in 'coordination' with $[Mg^{2+}]_3$-PP_i in GdolS-apo, left, and GdolS homology model, right. Both structures are overlaid with AT-AS in complex with $[Mg^{2+}]_3$-PP_i (pdb: 2OA6 chain D) with the ribbon of AT-AS not shown. Metal binding residues are shown with carbons in the respective ribbon colour, oxygens in red, nitrogens in blue and hydrogens in white. Phosphorus atoms are orange, magnesium is shown as green balls and water as small red balls.

Initially each of the metal binding residues within the metal binding domains were individually targeted with small substitutions; changing the amino acid to one of identical functionality but with a greater (D➔E, N➔Q) or lesser (E➔D) number of CH_2 units, the only exception to this rule was with serine, which was changed to alanine with a loss of functionality.

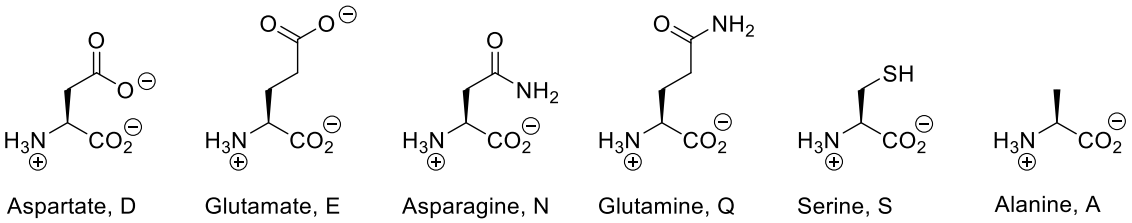


Figure 3.3 Amino acid residues involved in the metal binding mutations.

3.3.1 Aspartate-rich region

Single amino acid replacements in the aspartate rich region to glutamate residues all created functional germacradien-4-ol synthase mutants, although with a very minor effect on the product distribution. Each of these mutants produced minor, albeit noticeable, traces of germacrene A and D (**34** and **36**, total <2 %). While functional, measurement of the steady-state kinetic parameters of these mutants showed that they are much less active than the wildtype with GdolS-D80E being too slow to measure any steady-state kinetic values. GdolS-D81E and GdolS-D84E were catalytically eight to ten-fold slower while the mutation had little effect on the binding of FDP, **19**, Table 3.1.

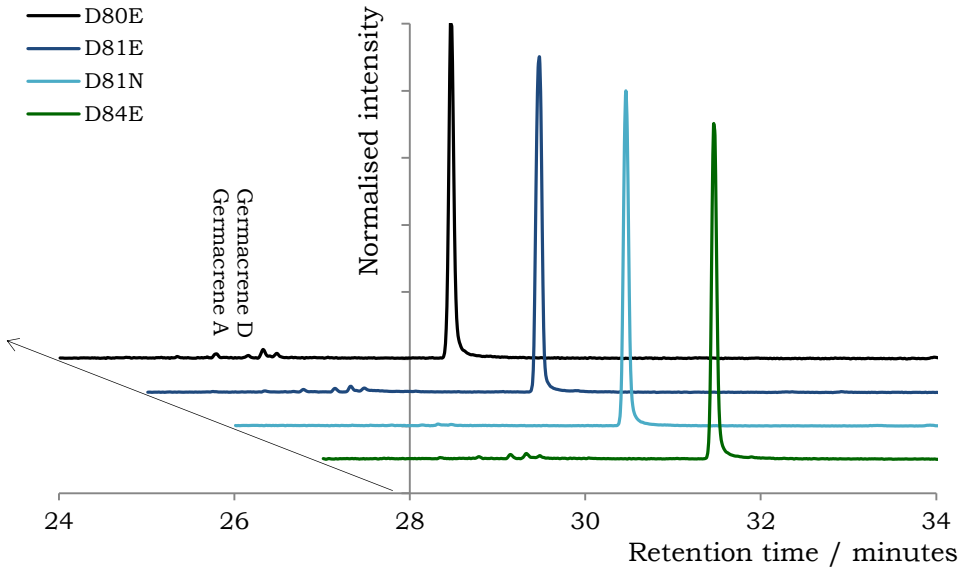


Figure 3.4 Overlaid normalised gas chromatograms of the incubations of FDP with metal-binding mutants

Further replacements of the aspartate residues to asparagine had an increasingly detrimental effect on the activity of the enzymes, yielding the inactive GdolS-D80N and GdolS-D84N mutants, and the GdolS-D81N mutant, which, while functional by GC-MS analysis, was not active enough to allow measurement of steady-state kinetic parameters.

These results are consistent with previously reported effects of amino acid replacements in the metal-binding domains, with mutation of the first aspartate in the DDxxD motif (Aspartate-80 in GdolS) causing a much greater effect on catalytic activity than mutation of the other residues in the aspartate rich region. This is consistent with the findings of Van der Kamp *et al.*[39] which indicated that upon the formation of the Michaelis complex two of the requisite magnesium ions are coordinated to this first aspartate, Section 1.3.3. Replacement of aspartate-81, the second aspartate in the DDxxD motif, displayed relatively little effect on the catalytic efficiency, with an 8-fold decrease in k_{cat} accompanied by a very minor decrease in K_M. The equivalent substitution in trichodiene synthase (TS) displays a similarly moderate, 3.5-fold, decrease in k_{cat}, but accompanied by a 1.5-fold increase in K_M;[70] while in pentalenene synthase (PS) replacement of the second aspartate with a glutamate leads to a dramatic 50-fold decrease in k_{cat}, and a 9-fold increase in K_M.[72] This indicates that while the aspartate-rich region is highly conserved in terpene synthases, the specific roles played by each of the metal-binding residues are not identical throughout all terpene synthases.

Table 3.1 Steady state kinetic parameters and product distribution (%) of GdolS-WT and metal binding mutants. Germacradien-4-ol (**79**) Germacrene A (**34**) and Germacrene D (**36**). [a] no detectable activity – activity too low for kinetic parameters to be determined. [b] not active

Enzyme	k_{cat}	K_M	Product Distribution		
			79	**34**	**36**
GdolS-WT	0.079 ± 0.003	1.07 ± 0.13	100	-	-
GdolS-D80E	nda[a]	nda[a]	98.6	0.5	0.9
GdolS-D80N	n/a[b]	n/a[b]	n/a[b]	n/a[b]	n/a[b]
GdolS-D81E	0.010 ± 0.0006	0.82 ± 0.11	99.0	0.7	0.3
GdolS-D81N	nda[a]	nda[a]	99.8	0.2	-
GdolS-D84E	0.009 ± 0.0003	0.92 ± 0.09	98.7	0.9	0.4
GdolS-D84N	n/a[b]	n/a[b]	n/a[b]	n/a[b]	n/a[b]
GdolS-E226D	nda[a]	nda[a]	99.8	0.2	-
GdolS-S222A	nda[a]	nda[a]	98.6	-	1.4
GdolS-N218Q	nda[a]	nda[a]	47.6	50.7	1.8
GdolS-N218L	nda[a]	nda[a]	99.8	0.2	-
GdolS-N218T	n/a[b]	n/a[b]	n/a[b]	n/a[b]	n/a[b]
GdolS-N218E	n/a[b]	n/a[b]	n/a[b]	n/a[b]	n/a[b]

3.3.2 NSE-motif

Asparagine-218

Substitution of the asparagine-218 of the NSE-motif with glutamine generated a functional germacradien-4-ol/germacrene A synthase, producing almost identical quantities of germacradien-4-ol and germacrene A (**79**:**34**, 55%:45%). This substitution also had a severe, detrimental impact on the catalytic efficiency, and it was not possible to measure steady-state kinetic parameters.

It was postulated that the increase in volume in the active site shifted the folded-FDP (**19**) further into the active site cleft and the conformation of the larger glutamine blocked access of water into the necessary position for the quenching to form germacradien-4-ol (**79**). In an attempt to further shift the product distribution towards germacrene A (**34**) asparagine-218 was further replaced with leucine, a residue with similar shape and volume but with no hydrogen-bonding capability. With the loss of hydrogen bond acceptors and donors and a corresponding increase in hydrophobicity it was proposed that this might further prevent ingress of water into the active site and further reduce the proportion of germacradien-4-ol (**79**) produced.

The GdolS-N218L mutant was a functional but inefficient enzyme, while reverting more closely to the product distribution displayed by wildtype enzyme, producing over 98 % germacradien-4-ol (**79**) with only a minor trace of germacrene A (**34**), Figure 3.5. Further replacements of this residue with amino acids of varying degrees of hydrogen-bond donors and acceptors (threonine, glutamate) only yielded inactive proteins.

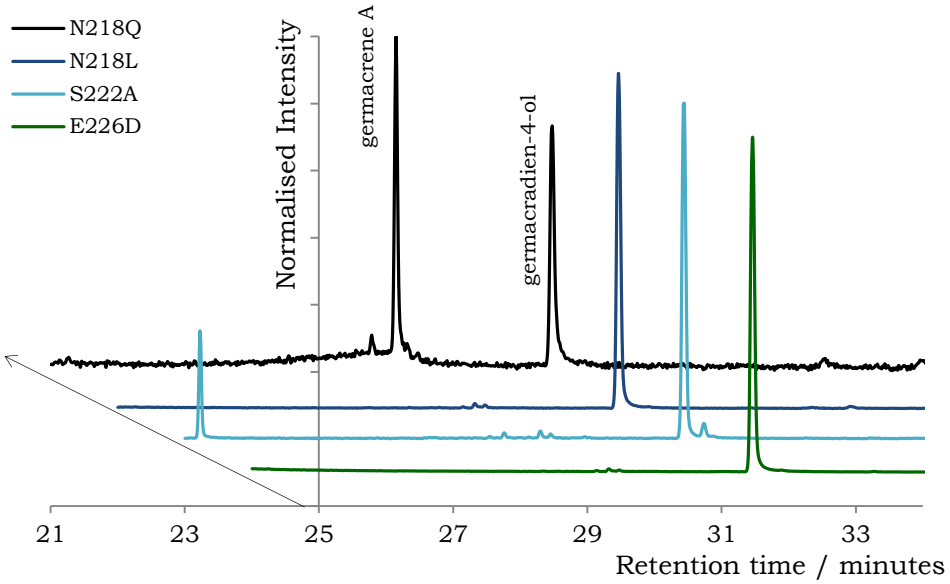


Figure 3.5 Overlaid normalised gas chromatograms of the functional NSE-motif mutants

Serine-222 and glutamate-226

Substitutions of serine-222 and glutamate-226 generated functional but inefficient mutants GdolS-S222A and GdolS-E226D, both producing germacradien-4-ol. While GdolS produced purely germacradien-4-ol (**79**), traces of germacrenes A (**34**) and D (**36**) were also produced by GdolS-S222A; both mutants are too inefficient to be able to measure steady-state kinetic parameters.

Substitutions to the NSE-motif impacted the functioning of the generated mutants to a much greater degree than those to the aspartate rich region, with the catalytic efficiencies of all NSE-mutants impaired to the extent that they could not be measured. This is consistent with results from other sesquiterpene synthases. Replacement of the arginine-219 in pentalenene synthase with hydrophobic residues leucine and alanine caused a complete loss of activity in the resulting mutants,[72] and similarly the N244L mutant of PR-AS was found to be inactive.[71] The investigation into PR-AS also generated mutants of the serine and glutamate metal-binding residues with the resulting S248A and E252D mutants displaying a 300-fold decrease in catalytic

efficiency with respect to wild-type, and the E252Q mutant a more dramatic 3500-fold decrease in catalytic efficiency.[71]

Most interesting of the DDxxD and NSE-mutants is the dramatic impact on the product distribution by the GdolS-N218Q mutant, which produces 45 % germacrene A (**34**). This mutant may indicate that, in GdolS, the metal-binding motif has a greater function than the locating of the substrate and initiation of the catalytic cycle, also actively involved in product specificity and locating of water in the active site.

Arginine-172

In response to the result with the arginine-218 mutants, the highly conserved hydrogen-bonding network was investigated. Arginine-200 in PR-AS is in close proximity to the $[Mg^{2+}]_3$-PP_i cluster, and replacement of this residue with lysine results in a 300-fold decrease in catalytic efficiency, along with a significant increase in the proportion of germacrene A (34) produced. This was attributed to arginine-200 being important for the stabilisation of the diphosphate anion in the active site, and indirectly contributing to the protonation of germacrene A.[40] The corresponding residue in GdolS is arginine-172, and it was proposed that replacement of this residue with lysine could affect the positioning and stability of the diphosphate anion as seen in PR-AS, leading to an increase in germacrene A (**34**) production. Replacement of arginine-172 with lysine results in a functional mutant with steady-state kinetic parameters showing a 200-fold decrease in k_{cat}. The product distribution was not drastically altered by this mutation, producing minor proportions of germacrenes A and D (**34** and **36**, < 6%), similar to other mutants described earlier, Table 3.2. Further replacement of arginine-178 with glutamate lead to primarily insoluble protein, and the small amount of soluble protein isolated demonstrated no cyclisation activity upon incubation with FDP (**19**). While these results clearly demonstrate the importance of this residue for catalysis, it has only a minor effect on the product distribution on GdolS. This could be due to a lesser effect on the positioning of the diphosphate anion or due to the relative rates of the competing deprotonation with the nucleophilic attack of water on the intermediate carbocation.

Table 3.2 Steady state kinetic parameters and product distribution (%) of GdolS-WT and metal binding mutants. Germacradien-4-ol (79) Germacrene A (34) and Germacrene D (36) [a] not active

Enzyme	k_{cat}	K_M	Product Distribution		
			79	**34**	**36**
GdolS-WT	0.079 ± 0.003	1.07 ± 0.13	100	-	-
GdolS-R172K	0.00035 ± 0.00002	0.82 ± 0.062	>96	<4	<2
GdolS-R172E	n/a[a]	n/a[a]	n/a[a]	n/a[a]	n/a[a]

3.4 Water-binding residues

Careful analysis of the active site using a combination of homology models and the X-ray crystal structure of the apo-enzyme highlighted two other residues on the terminal helix which were proposed to play a role in hydrogen-bonding to water in the active site, glutamate-307 and tyrosine-303. Serendipitously, Chen *et al.* in a crystallographic study on aristolochene synthase from *Aspergillus terreus* (AT-AS) found evidence of water molecules trapped in the upper active site by hydrogen bonds with three key residues: N213, N299 and S303.[38] Sequence alignment of AT-AS with GdolS showed that these three residues were the equivalent of N218, Y303 and E307; Asparagine-218 having previously been noted as a possible key residue in the incorporation of water, Section 3.3.2.

3.4.1 Tyrosine-303

As well as being highlighted as a water-bonding residue tyrosine-303 was also proposed to be likely to have importance as a residue for the stabilisation of developing carbocations, indeed in Figure 3.6 (bottom right) the proximity of the tyrosine to the C10-C11 bond can be noted and it can be postulated that it would be suitably positioned to stabilise the germacryl cation upon ring closure. Both hydrogen-bonding and stabilising functions were probed by separate replacements to the tyrosine-303. Loss of the hydroxyl in the GdolS-Y303F mutant probed the importance of the hydrogen-bonding nature of this residue, while a more drastic change to a non-aromatic, hydrophobic residue to generate the GdolS-Y303I mutant probed the importance of the aromaticity to the activity of the enzyme, a third mutation retained the hydroxyl functionality while removing the aromatic group generating the GdolS-Y303T.

```
GdolS   RSMSLDPAQRLAAERYAD-GMALWLAGYLHWESHTRRYHHG---
PR-AS   EKIASPDGCSEAAKAY-MKGLEYQMSGNEQWSKTTRRYN-----
AT-AS   RLS-AEGLETPGLAAY-VEGLEYQMSGNELWSQTTLRYSVVVD-
```

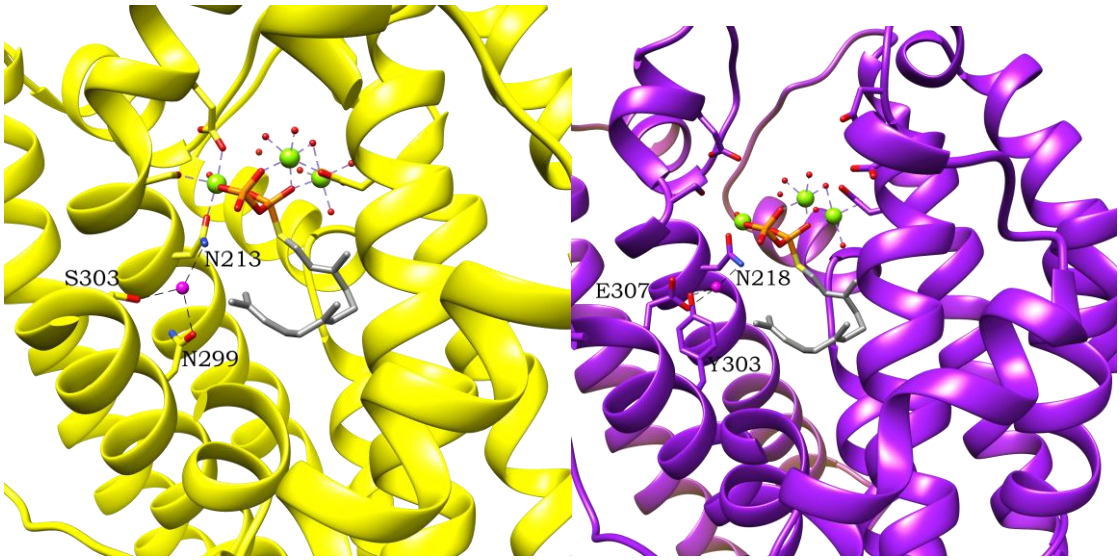


Figure 3.6 Top, Sequence alignment of the terminal helix and loop of GdolS with AT-AS and PR-AS. N299 and S303 in AT-AS and the aligning residues in GdolS and PR-AS are highlighted in yellow. Bottom left, AT-AS bound to FSDP (farnesyl-S-thiolodiphosphate) and $[Mg^{2+}]_3$-PP_i, with the active site water highlighted in magenta and interactions with N213, N299 and S303 shown as black dashes (pdb: 4KUX, chain A). Bottom right, GdolS-homology model aligned with AT-AS and the highlighted active site water (magenta) interacting with N218, Y303 and E307. Residues are shown with carbons in the respective ribbon colour, oxygens in red, nitrogens in blue and hydrogens in white. Phosphorus atoms are orange, sulfur yellow, magnesium is shown as green balls and water as small red balls, with the exception of the highlighted magenta water.

As predicted, tyrosine is an important active-site aromatic residue for the stabilising of developing carbocations. Substitution of this residue to isoleucine generated the functional GdolS-Y303I mutant, but the kinetic efficiency was severely hampered and it was not possible to determine steady-state kinetic parameters. This mutant generates primarily germacradien-4-ol (**79**), but also traces of other products including germacrene A (**34**) and D (**36**).

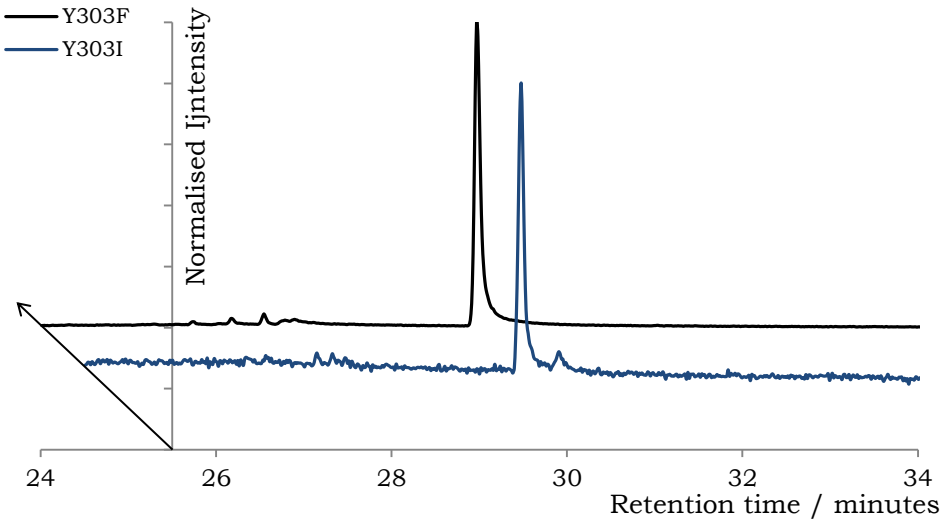


Figure 3.7 Overlaid normalised gas chromatogram of functional tyrosine-303 mutants

The smaller change to generate the GdolS-Y303F mutant has a lesser effect on catalytic efficiency, with a four-fold decrease in k_{cat} compared to wild type and a small decrease in K_M. Product distribution is only mildly affected by the loss of the hydroxyl-group in the active site with a minor increase in level of germacrene A (**34**).

Table 3.3 Steady state kinetic parameters and product distribution (%) of GdolS-WT and tyrosine-303 mutants. Germacradien-4-ol (**79**) Germacrene A (**34**) and Germacrene D (**36**). [a] no detectable activity – activity too low for kinetic parameters to be determined. [b] not active

Enzyme	k_{cat}	K_M	Product Distribution		
			79	**34**	**36**
GdolS-WT	0.079 ± 0.003	1.07 ± 0.13	100	-	-
GdolS-Y303I	nda[a]	nda[a]	> 99	trace	-
GdolS-Y303F	0.014 ± 0.006	0.66 ± 0.12	> 96	> 2	< 1
GdolS-Y303T	n/a[b]	n/a[b]	n/a[b]	n/a[b]	n/a[b]

The third mutant, GdolS-Y303T, was expressed effectively; however, upon purification it was found that the majority was expressed as insoluble inclusion bodies, which was not improved by inducing with a lower concentration of IPTG, or by expression at a lower temperature. A small amount of soluble protein was isolable, purified and found to be inactive by GC-MS analysis. Basic extraction of the inclusion bodies failed to improve the yield of soluble protein, as did attempts at purifying the enzyme in denaturing conditions and refolding the purified enzyme; all failing in producing any active protein.

Tryptophan-306

Another aromatic active-site residue, tryptophan-306, was also probed by site-directed mutagenesis. This tryptophan residue, highly conserved in terpene synthases, corresponds to tryptophan-334 in PR-AS, a residue which has been proposed to stabilise the eudesmane cation (**68**) in aristolochene biosynthesis.[74] Replacement of tryptophan-334 in PR-AS with residues of decreasing aromatic character leads to decreasing catalytic efficiency and increasing proportions of germacrene A (**34**) formed. This was proposed to be due to the importance of this tryptophan residue in stabilising formation of the eudesmane cation through quadrupolar charge-π interactions. Substitution of the corresponding residue in GdolS, tryptophan-306, with alanine generated the functional GdolS-W306A mutant; While the resulting mutant was

functional, producing germacradien-4-ol (**79**) with traces of germacrenes A and D (**34** and **36**, < 1 %), no kinetic parameters could be measured due to the very low activity of this mutant. This implicates tryptophan-306 as a vital residue for the catalytic efficiency of GdolS, necessary for the stabilisation of a carbocation intermediate in the cycle catalysed by GdolS.

3.4.2 Glutamate-307

A similar approach was taken in the investigation into the role of glutamate-307, exchanging it for residues of lesser hydrogen bonding ability while maintaining a similar volume in the active site. Functional mutants GdolS-E307Q and GdolS-E307M were generated, and the pentane extractable products of the incubations with FDP analysed by GC-MS; GdolS-E307Q produced >96 % germacradien-4-ol (**79**) with a small increase in hydrocarbon terpenes, germacrenes A (**34**) and D (**36**); while GdolS-E307M, with less hydrogen bonding ability, further decreased the proportion of germacradien-4-ol (**79**, < 95 %), with an associated increase in germacrene A (**34**, > 4 %).

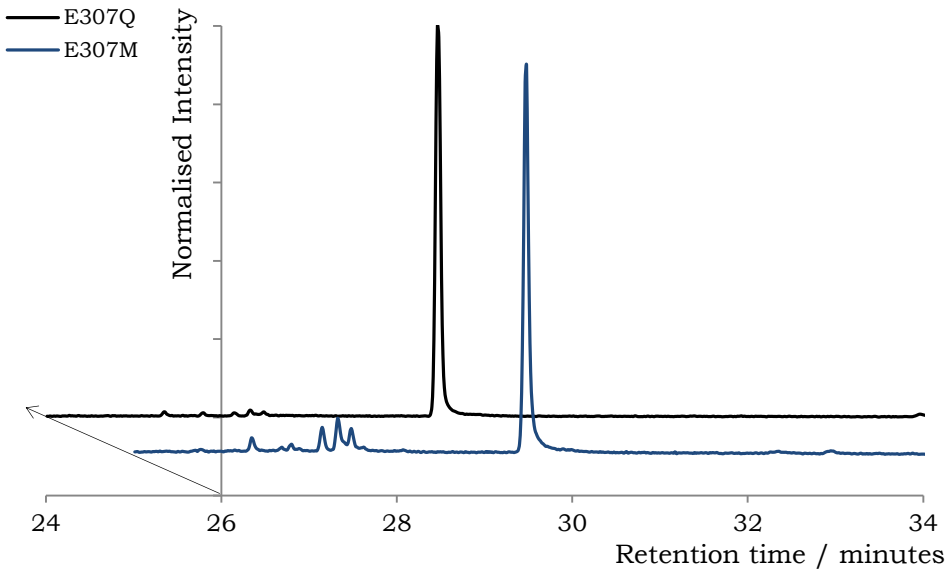


Figure 3.8 Overlaid normalised gas chromatogram of functional glutamate-307 mutants

Measurement of the steady-state kinetic parameters of these mutants found GdolS-E307Q to be moderately active with a four-fold decrease in k_{cat} coupled with a two-fold increase in K_M leading to a thirteen-fold decrease in efficiency (k_{cat} 0.015 s^{-1}, K_M 2.7 μM, k_{cat}/K_M 5,500 M^{-1} s^{-1}). Further mutation to the GdolS-E307M mutant caused such retardation in catalytic activity that it was not possible to measure kinetic parameters. These results implicate this residue as important for catalysis, possibly involved in the binding and organisation of one of the magnesium ions, leading to the increase in K_M upon

replacement with a neutral residue. While important to catalysis, the minor perturbation of product distribution displayed by the GdolS-E307Q and GdolS-E307M mutants suggest this residue is not involved in the incorporation of water in the product.

Table 3.4 Steady state kinetic parameters and product distribution (%) of GdolS-WT and glutamate-307 mutants. Germacradien-4-ol (**79**) Germacrene A (**34**) and Germacrene D (**36**). [a] no detectable activity – activity too low for kinetic parameters to be determined.

Enzyme	k_{cat}	K_M	Product Distribution		
			79	**34**	**36**
GdolS-WT	0.079 ± 0.003	1.07 ± 0.13	100	-	-
GdolS-E307Q	0.013 ± 0.0004	2.24 ± 0.16	> 98	< 1	< 1
GdolS-E307M	nda[a]	nda[a]	> 95	> 4	< 1

3.4.3 Helix-hybrid

It was noted that these two residues are in relatively close proximity to asparagine-218 of the NSE-motif and they are located on the C-terminal helix, and so it was proposed to create a hybrid bearing the C-terminal of another terpene synthase which produces a hydrocarbon terpene product. The synthase selected was δ-cadinene synthase from *Gossypium arboreum* (DCS), due to its availability within our lab and the published X-ray crystal structure, which could be used in the preparation of a protocol for the helix-switching.

```
GdolS       SMSLDPAQRLAAERYADGMALWLAGYLHWESHTRRYHHG-----------
DCS-SC      ALGLPDTVRTHADDCAASLLVWVRGHLEWGLETPRYRPGTTGTGTD----
DCS-GA      PTEMPTEVLNRSLNLARVMDVLYREGDGYTYVGKAAKGGITSLLIEPIAL
```

Figure 3.9 Sequence alignment of GdolS (Accession: BAL14866.1) with a bacterial δ-cadinene synthase (DCS-SC, from *Streptomyces clavuligerus ATCC 27064*, Accession: EFG03819.1[118]) and with DCS-GA (Accession: Q39761.1)

Sequence alignment and comparison of the GdolS homology model with the X-ray crystal structure of DCS showed that the C-terminal helix of DCS is longer than that in GdolS by approximately ten amino acids and it is also positioned slightly differently; beginning higher up the active-site cleft. Effectively leaving a significant proportion of this helix available to cover the active-site entrance and possibly more efficiently desolvating the active site.

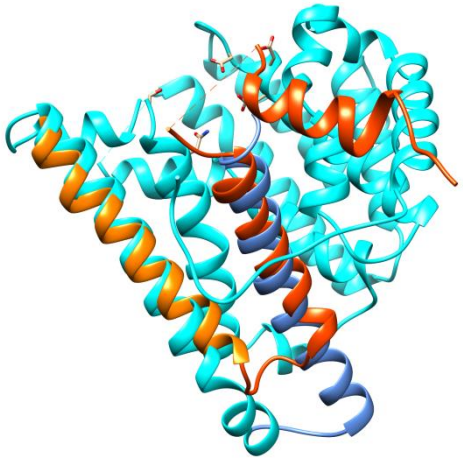


Figure 3.10 apo-GdolS (cyan) overlaid with the terminal helices of DCS (orange). The exchanged regions are highlighted by the darker colour.

Preparation

Cloning

It was decided to replace the entire C-terminal helix including the preceding loop, to minimise the effects on the folding of the helix. Both cDNA for GA-DCS and GdolS bear a *Bam*HI restriction site at the 3' end and this site was used for the removal and insertion of the fragments. A second unique restriction site prior to the methionine-279 of the loop in GdolS was needed; unfortunately no unique restriction sites were already present and no silent mutations were possible to generate a unique site. A unique *Sma*I restriction site could be generated by mutation of a single base-pair causing the single point mutation of aspartate-274 to a glycine and so this mutation was introduced using site-directed mutagenesis. The desired section of the cDNA for GA-DCS was amplified using the polymerase chain reaction with the forward primer bearing the desired loop sequence and the *Sma*I restriction site, Figure 3.11.

After confirmation of the successful mutation, the gene encoding the GdolS-D274G mutant was digested sequentially with *Sma*I followed by *Bam*HI (Section 6.1.6), and purified by gel electrophoresis (1 % agarose, Figure 3.12 C) to isolate the open plasmid. The DCS-fragment was isolated from the pET21d-DCS expression vector by PCR amplification, and the PCR product was then isolated by gel electrophoresis (2 % agarose, Figure 3.12 A) and directly ligated into a StrataClone™ Blunt PCR cloning kit. The resulting plasmid containing the DCS-fragment was then sequentially digested with *Sma*I followed by *Bam*HI and purified by gel electrophoresis (2 % agarose, Figure 3.12 B) to isolate the digested fragment.

A.

^{272}L P D A → ^{272}L P ***G*** A

ctc ccg gac gcc → ctc ccg g***g***c gcc

B.

L P G A C R S | F L K P T E

5'- ctc ccg ggc gcc tgc cgc tcg|ttt ttg aaa cca aca gaa -3'

-*Sma*I- GdolS helix end|DCS section start

C.

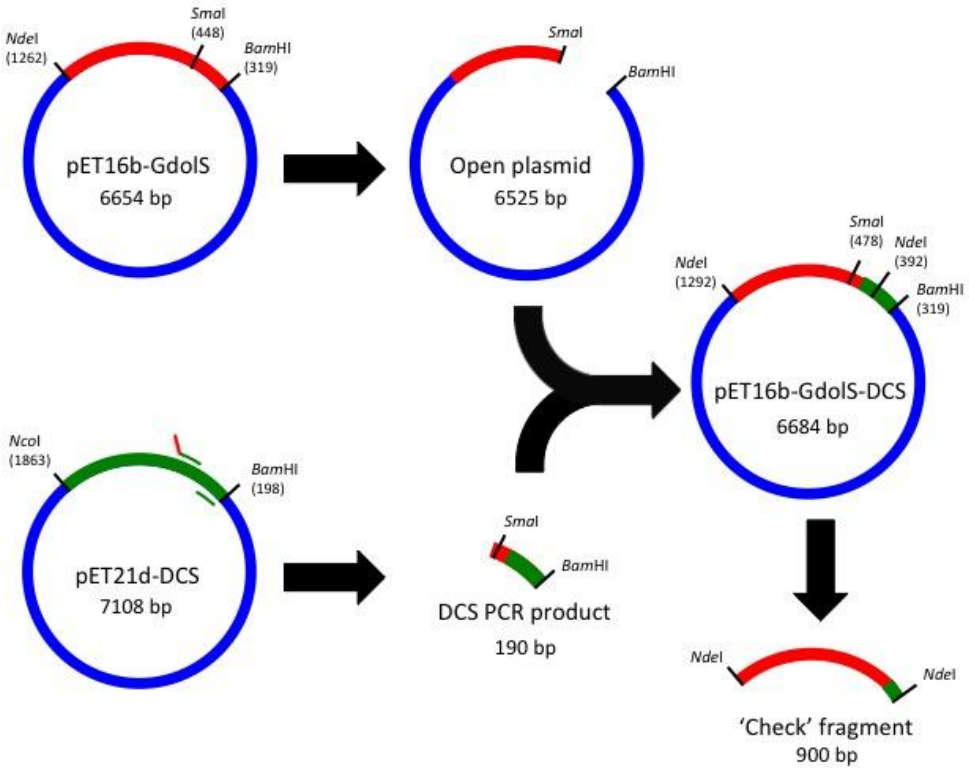


Figure 3.11 A. Mutation of GdolS to insert *Sma*I restriction site, with the mutated base and resulting amino acid change in bold italics; B. forward primer to generate DCS loop helix cDNA fragment with GdolS section in red and DCS complementary fragment in green; C. Diagram of subcloning strategy.

The fragment and open plasmid were then ligated with T4 DNA ligase (16 °C, 14 hours) and supercompetent XL1-blue cells then transformed with 5 µL of the resulting mixture. A selection of colonies were grown up and the plasmid isolated and tested for successful ligation by digestion with *Nde*I restriction enzyme; there is an *Nde*I site at the beginning of the cDNA for GdolS and another within the inserted DCS-fragment yielding a 900 basepair (bp) fragment when digested. Digested samples were analysed by gel electrophoresis (2 % agarose, Figure 3.12 D) and sample showing a band at ~900 bp were sequenced to confirm generation of the GdolS-DCS hybrid plasmid.

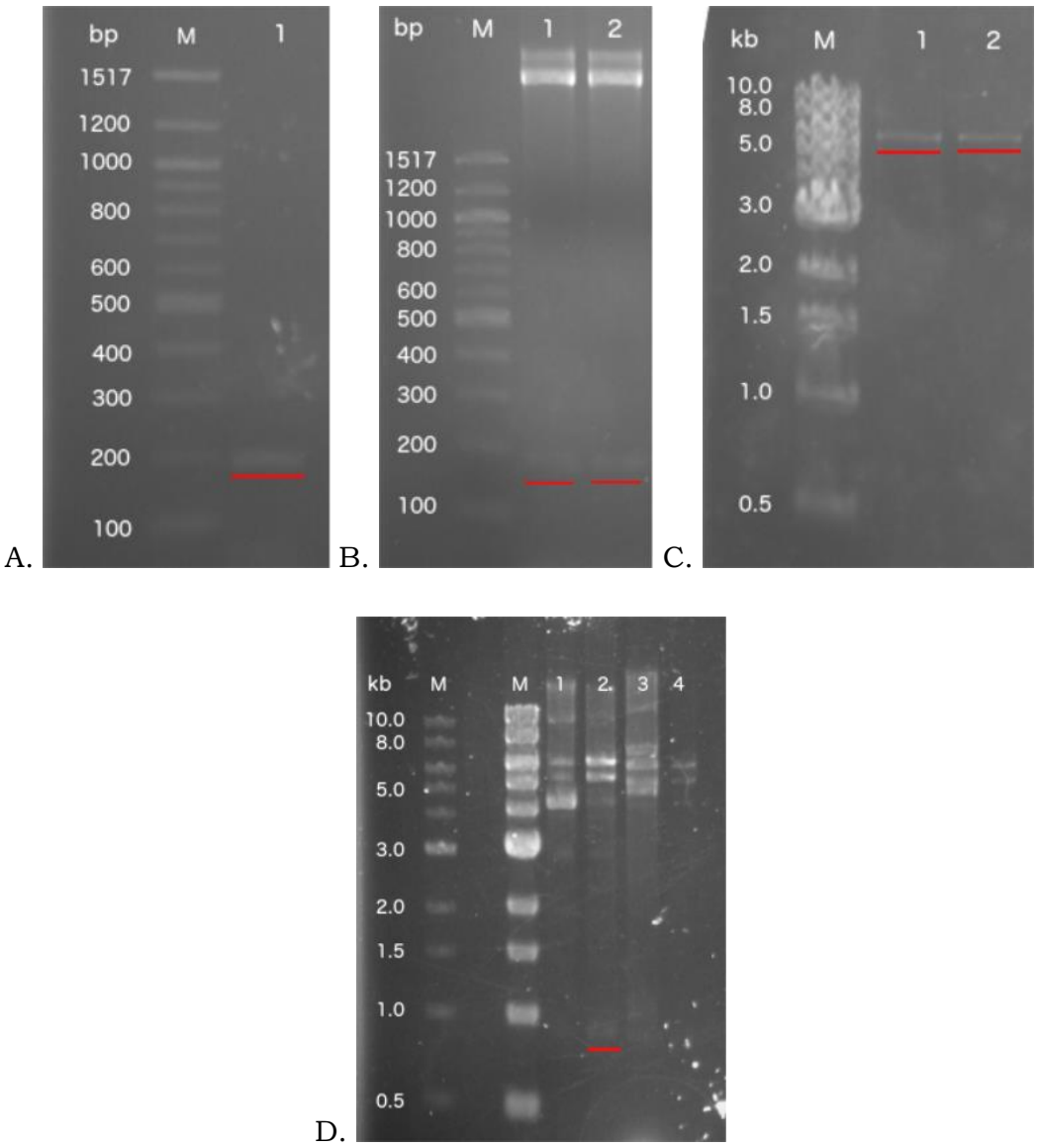


Figure 3.12 Agarose gels for the preparation of the GdolS-DCS hybrid gene; A. Purification of the PCR product of the DCS loop-helix fragment (1); B. Purification of the DCS loop-helix fragment digested from cloning vector (1 & 2); C. Purification of the GdolS-SmaI vector fragment digested from pET16b (1 & 2); D. *Nde*I digests of ligated plasmids (1-4 from different colonies). In all images lanes labelled 'M' are DNA ladder with sizes shown to the left, where 'bp' is number of basepairs and 'kb' is number of kilobases. In all images the desired fragments are underlined in red.

Expression and purification

E. coli BL21-(DE3) cells were transformed with the constructed plasmid and the gene was overexpressed at 16 °C for 21 hours, producing almost exclusively insoluble protein, Figure 3.13 A.; while a small quantity of protein was isolated by affinity chromatography this protein was inactive when incubated with FDP. A test expression was performed to follow the production of protein, inducing with 0.1 mM IPTG and growing at 37 °C for 6 hours before reducing the temperature to 16 °C for a further 15 hours (overnight). Samples were taken before induction, 2.5, 4, 6 and 21 hours post-induction and analysed by SDS-PAGE, Figure 3.13 B., indicating the greatest proportion of

protein at 2.5 hours post induction. GdolS-DCS hybrid was then produced using these conditions; unfortunately despite the high yield of protein, this was solely as insoluble inclusion bodies and so efforts were directed towards solubilising and purifying the protein rather than improving the level of soluble protein produced. The standard method of basic extraction was initially tried and failed to yield any soluble protein and so a number of modifications were trialled; extending the period allowed for refolding from one hour to 15 hours; adding $MgCl_2$ (Final concentration 10 mM) after adjusting the pH to 12 and before reducing the pH to 8; and adding detergent (0.1 % Tween-20) to the buffer used for resuspension of the pellet. In the cases of $MgCl_2$ and Tween-20 additives refolding was also left for 15 hours.

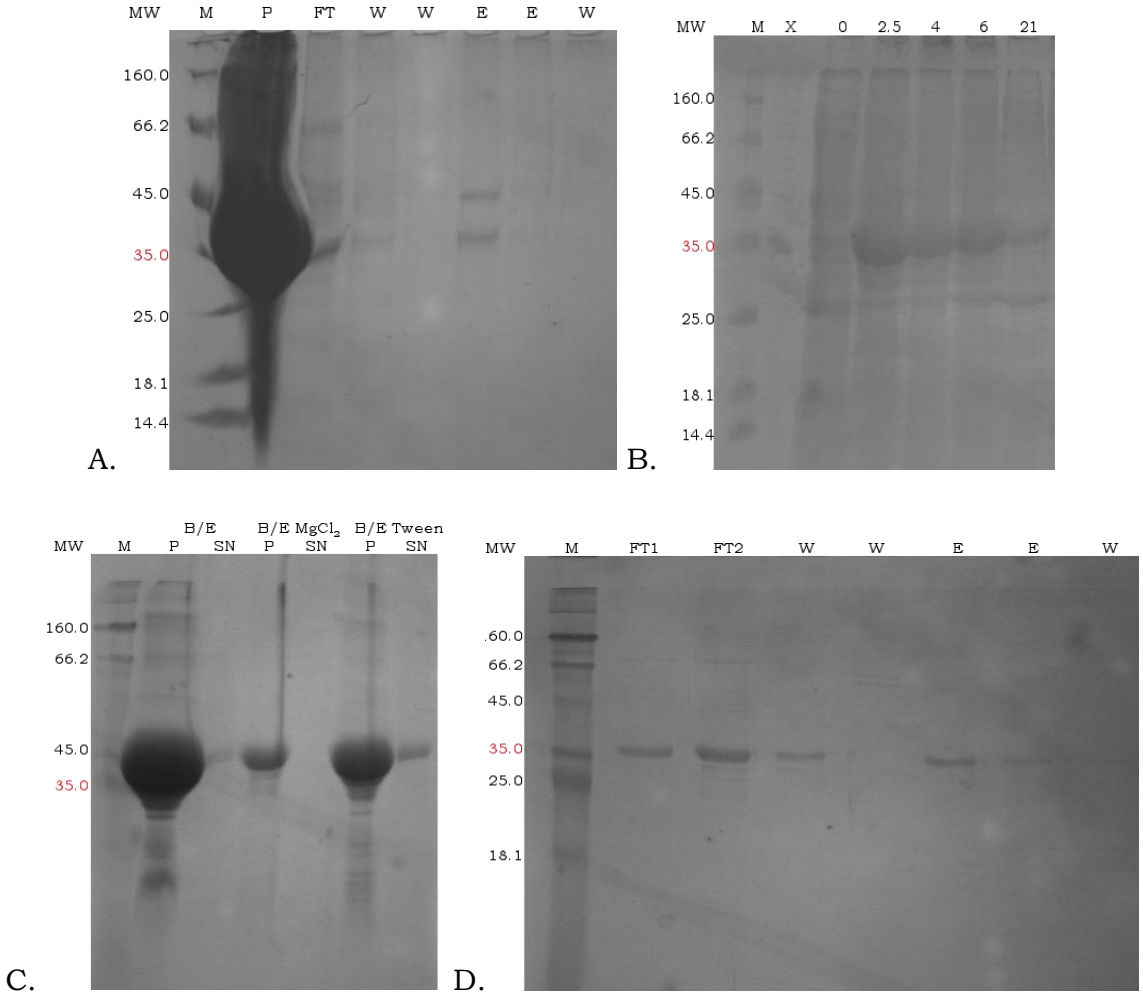


Figure 3.13 SDS-polyacrylamide gels of the production and purification of the GdolS-DCS hybrid. A) Purification from the expression of pET16b-GdolS-DCS at 16 °C, showing the pellet (P), flowthrough (FT), washes (W) and elution (E). B) Test expression of pET16b-GdolS-DCS in BL21(DE3), analysis of samples taken at periods after induction, where the lanes are labelled by the number of hours after induction (X is an empty lane). C) Analysis of the pellets (P) and supernatant solutions (SN) after basic extraction of the insoluble inclusion bodies in buffer (B/E), buffer with 10 mM $MgCl_2$ (B/E $MgCl_2$) and buffer with 0.1% Tween 20 (B/E Tween). D) Purification of the soluble GdolS-DCS;; Showing the flowthroughs from basic extractions in buffer (FT1) and buffer with 0.1% Tween 20 (FT2), washes (W) and elution (E). MW, molecular weight x10^3, approximate expected mass of GdolS higfhlighted in red.

The addition of $MgCl_2$ was proposed to aid in the refolding by directing the organisation of the metal binding region; this method failed to solubilise any protein, instead causing a precipitate to form rapidly after addition of the $MgCl_2$. The other two methods were marginally more successful with the extended refolding time yielding a small amount of soluble protein and the addition of Tween-20 further increasing the soluble yield, Figure 3.13 C. These soluble fractions were both applied to the same Nickel-affinity column and purified by affinity chromatography (Section 6.1.11). Analysis of the fractions by SDS-PAGE showed the majority of the soluble protein failed to bind to the column, and was found in the flowthrough fractions, Figure 3.13 D. A small proportion of GdolS-DCS hybrid protein was successfully purified and concentrated by ultrafiltration.

Analysis

The pentane-extract of the incubation of purified GdolS-DCS with FDP showed no peaks and neither did that of the incubation with the soluble, unpurified hybrid.

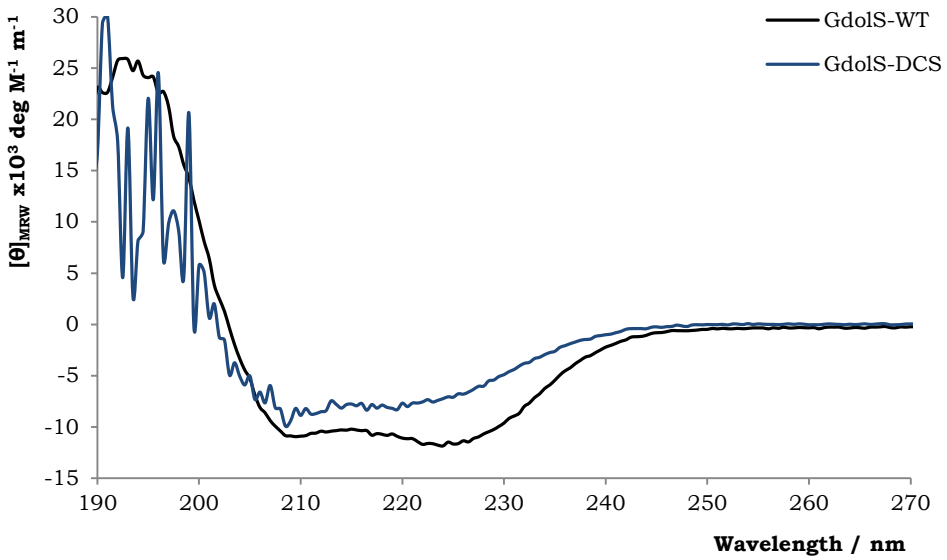


Figure 3.14 Overlaid CD spectrum of GdolS-WT and the GdolS-DCS hybrid

In order to understand if the inactivity was due to an unproductive active-site conformation or primarily due to incorrect or poor folding the circular dichroism spectrum of the GdolS-DCS hybrid was measured and compared against wildtype GdolS, Figure 3.14. Due to the low concentration of the GdolS-DCS sample a neat sample was used for measurement of the CD spectrum, and the Tris present (10 mM) creates noise in the spectrum at lower wavelengths, normally below 200 nm but due to the higher concentration here,

relative to the protein, it has more of an effect. Due to the presence of Tris-buffer the sample is quite noisy, but clearly lacks the definite minima shown in the wildtype sample, in particular that at 222 nm. The magnitude of the typical α-helical minima at 209 and 222 nm are clearly less than those of the wildtype. The shape and magnitude of the spectrum strongly suggest that the protein is not fully folded, and the lack of terpene synthase activity is due to this incorrect folding.

3.5 H-1α-loop

In the X-ray crystal structure of apo-GdolS the H-1α loop, the loop succeeding the NSE-motif, and the E-F loop are not resolved due to the high mobility of these regions. Previous research on the H-1 α loop in fungal sesquiterpene synthases from *Coprinus cinereus* found that mutations to this region could markedly influence the product profile in some cases,[77] and so the H-1 α loop in GdolS was studied using sequence alignments and homology model of GdolS to identify residues that may be important in the movement and locating of the loop during catalysis. The previous study by Lopez-Gallego *et al.*[77] highlighted the role of ionic interactions between the H-1a loop and the metal-binding motives (Section 1.1) and so residues with this ability were identified in the H-1α loop of GdolS; glutamate-228, arginine-229, aspartate-231 and asparagine-234. Of these residues, glutamate-228 does not appear to be positioned suitably for interaction with the active site, situated on the outer edge of the H-1α loop facing towards the solvent. Arginine-229 and aspartate-231 are facing into the active site, while in the homology model they are too far away to interact with the $[Mg^{2+}]_3$-PP_i complex and metal binding residues, 6 – 8 Å, it could be imagined that the mobility of the loop will bing them into closer proximity during the catalytic cycle. Asparagine-234 is facing towards the H-helix and suitably positioned to interact with serine-222 and glutamate-226 of the NSE-motif, with distances of 4.7 Å and 2.6 Å respectively in the homology model, Figure 3.16.

```
GDOLS       NDVRSFAQESERGDVA NLVMIVRR
COP6        NDLLSFYKEELDCETV NFIS
COP4        NDVYSTDMEQAKGHTGNNVVTVLMK
COP3        NDIFSYNVEQSKGDTH NMIIILM
COP2        NDIFSFNREQSRHDSF NMVSIVM
COP1        NDIFSFNVEQSKGDTH NMIPVVM
EABS        NDLCSLQREEGLPAVQFNMVMTLQRTH
G11OL       NDLFSYQREVEDEGENANCVLVLERFLNV
PS          NDVYSLEKEEARGDMD NLVLVLVLEHTR
PSCANE      NDIASLEKEEARGEQN NMVMILRREH
DCSSC       NDLMTVDKEAAHGDLH NLVLVTEHDR
PRAS        NDIYSYDKEEEASRTGHKEGAFLCSAV
```

Figure 3.15 Alignment of the H-1α loop regions of bacterial and fungal sesquiterpene synthases; COP1-6, various sesquiterpene synthases from *Nostoc* cyanobacteria;[46] EABS, epi-α-bisabolol synthase from *Streptomyces citricolor*;[93] G11OL, germacradien-11-ol synthase from *Streptomyces coelicolor*;[119] PS, pentalenene synthase from *Streptomyces UC5319*;[120] SC-DCS, δ-cadinene synthase from *Streptomyces clavuligerus*;[118] PRAS, Aristolochene synthase from *Penicillin roquefortii*.[60] The NSE-motif is highlighted in yellow, with other conserved residues highlighted in other colours.

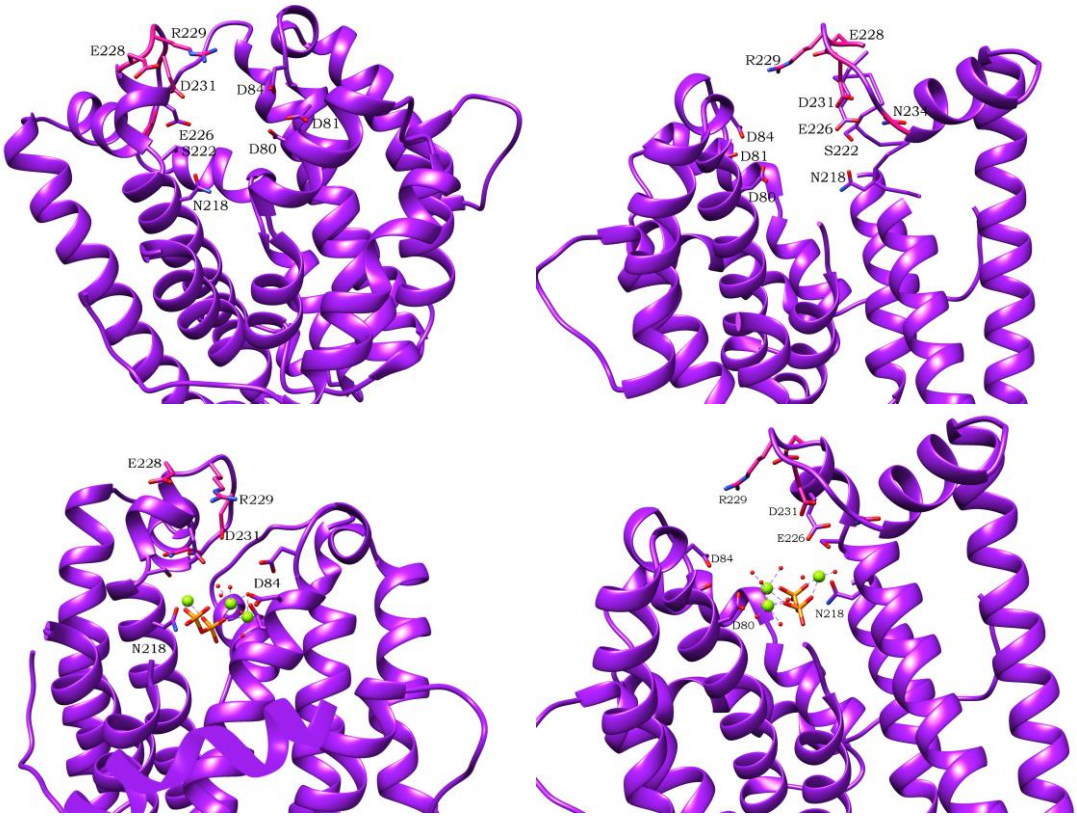


Figure 3.16 GdolS homology model showing charged residues of the H-1α loop and, below, overlaid with $[Mg^{2+}]_3$-PP_i (AT-AS, pdb: 2OA6). Residues are shown with carbons in the respective ribbon colour, oxygens in red, nitrogens in blue and hydrogens in white. Phosphorus atoms are orange, sulfur yellow, magnesium is shown as green balls and water as small red balls.

As the residues with the greatest potential for interactions with the metal-binding regions and, as such it was proposed, the greatest effect on catalysis arginine-229 and aspartate-231 were selected for study through the replacement of these residues. Mutants GdolS-R229E and GdolS-D231L were

generated with the aim of detrimentally affecting any electrostatic coordination to other active site residues, both mutants were expressed and purified as previously described, and the purified proteins were incubated with FDP and the pentane extractable products analysed by GC-MS. The arginine-229 mutant showed terpene synthase activity, producing solely germacradien-4-ol (**79**), while the aspartate-231 mutant showed to be inactive, generating no pentane extractable products upon incubation with FDP. This implies that this aspartate residue is vital for catalytic activity; suggesting it has a role in the catalytic cycle. Given the likely positioning of this residue, as can be seen in the homology model, Figure 3.16, it is unlikely to play a direct role in the cyclisation of FDP. However, it is suitably positioned to play a role in the binding of the $[Mg^{2+}]_3$-PPi cluster, and a role in the formation of the Michaelis complex would account for the complete loss of activity upon replacement of this residue with one of a hydrophobic nature.

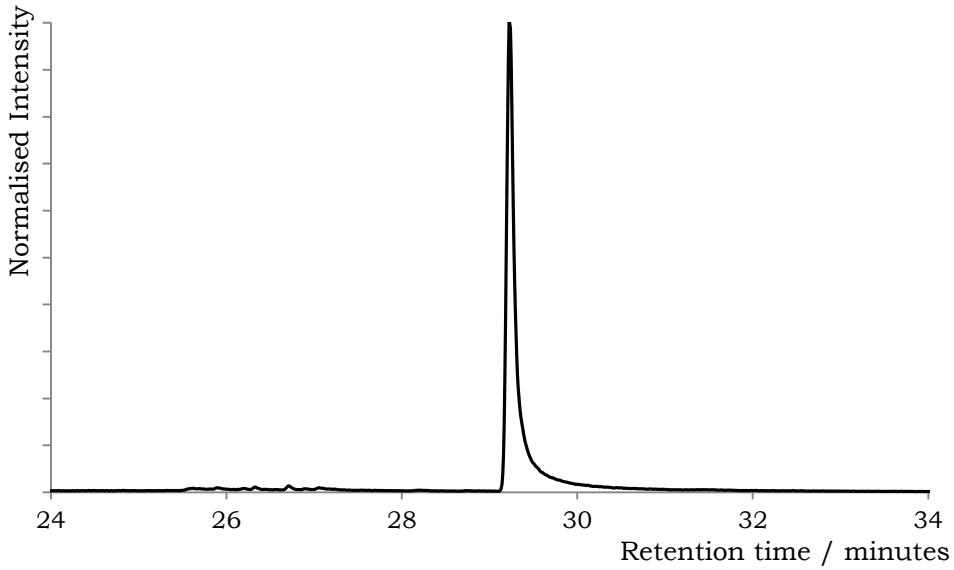


Figure 3.17 Normalised chromatogram of arginine-229 mutant

3.6 Monomer/dimer

The X-ray crystal structure of the apo-structure of GdolS was comprised of two GdolS monomers in very close proximity, suggesting that GdolS may form a dimer in solution. Close analysis of the interface between the two chains highlighted a possible contact point between the two monomers comprising a salt bridge and a disulphide bridge. The possible disulphide bridge was not shown in the crystal structure with the two cysteine-141 residues facing away from each other and a distance of >4.5 Å, rotation of the cysteine only

minimises the distance to <3.4 Å with a dihedral angle of <45°, making formation of a disulphide bond in this orientation unlikely.

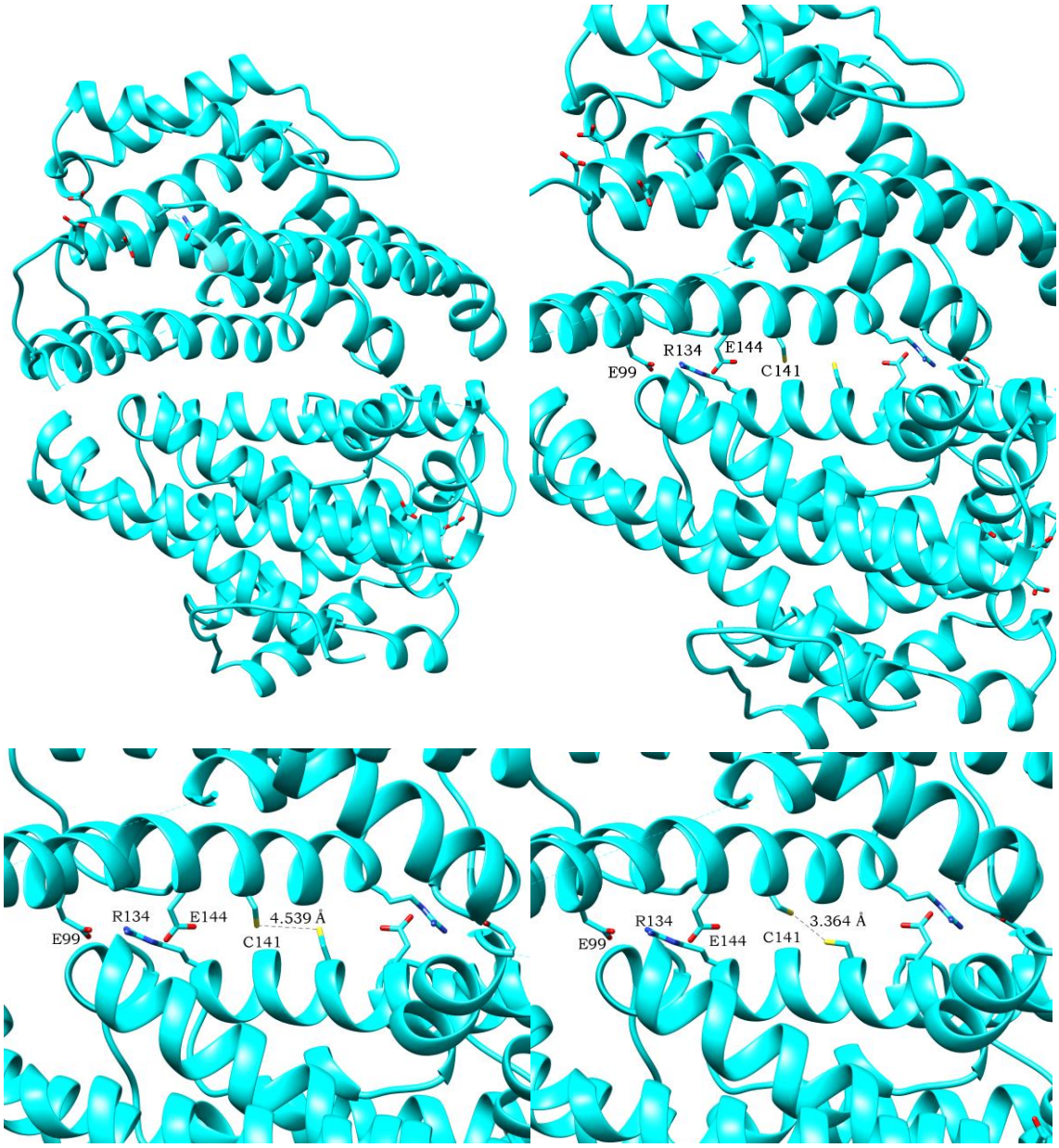


Figure 3.18 apo-crystal structure of GdolS (A) showing dimer interface with salt bridges and cysteine residues (B). Bottom, cysteine-cysteine distance is highlighted (C), minimised cysteine-cysteine distance (D).

The residues highlighted as a possible salt bridge were glutamates-99 and -144 on one monomer with the arginine-134 of the second monomer effectively creating two salt bridges, one at either end of the contact face, with distances of <2.7 Å. To investigate the possibility of a solution state dimer the central residue of the saltbridge trio, arginine-134 was mutated to an aspartate and an alanine, and the resulting mutants analysed by size-exclusion chromatography and native-PAGE in concert with the wildtype protein.

3.6.1 Arginine-134

The GdolS-R134E mutant was expressed primarily as insoluble inclusion bodies but a small amount of soluble protein was isolated. Basic extraction of the inclusion bodies yielded a further sample of soluble protein. Both samples of GdolS-R134E produced solely germacradien-4-ol prior to purification, however, less than 24 hours later after purification and dialysis the activity was severely reduced and no attempt at measuring steady-state kinetic parameters was made. Repeated attempts at preparing a sample with maintained activity highlighted rapid aggregation as the probable cause of the loss of activity. Expression and purification of GdolS-R134A was straightforward, yielding the functional mutant which, upon incubation with FDP, produced solely germacradien-4-ol (**79**). Measurement of the steady-state kinetic parameters found this mutant to be 5-fold less active than wildtype (k_{cat} 0.0013 s^{-1}), although with a reduced K_M (0.42 μM).

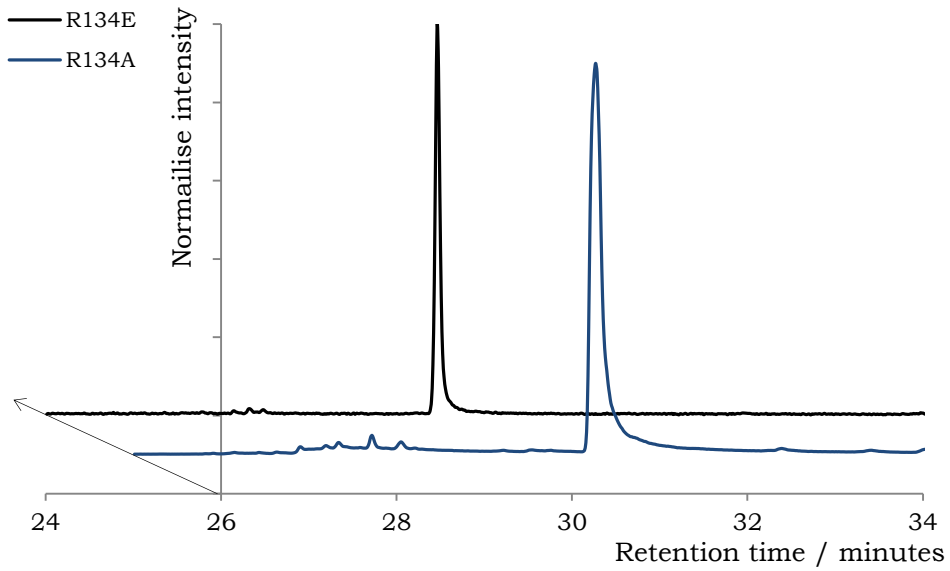


Figure 3.19 Normalised overlaid gas chromatogram of incubations with arginine-134 mutants.

Analysis of GdolS-WT by size-exclusion chromatography showed a small amount of aggregated protein eluting in the void volume (~7 mL) and a single peak eluting at 14.4 mL; Comparison against proteins of known size indicated that this is a monomer. Despite the low concentration of the R134E samples they were analysed by size-exclusion chromatography and it was found that the majority of the protein was aggregated and ran through the column in the void volume. There were two other peaks, poorly resolved due to the low intensity of the sample, one of these peaks eluted at the same time as the GdolS-WT, while the second peak eluted slightly earlier, suggesting monomer

and dimer. In the sample of R134E solubilised by basic extraction (Figure 3.20 left, dark green) the proportion of monomer appears to be greater than that seen in the soluble protein obtained directly from expression; Similarly, repeating the size-exclusion analysis in the presence of salt (150 mM NaCl) increases the proportion of the monomer over the dimer, while in the wildtype enzyme it causes the appearance of a second peak at earlier retention time, thought to be dimer. Analysis of the R134A mutant by size-exclusion chromatography shows two peaks similar to that shown by the GdolS-R134E mutant, however with a lesser proportion of aggregation. This mutant was also less prone to rapid aggregation than the R134E mutant. The relative instability of the R134E mutant compared to the GdolS-WT and the R134A mutant could be attributed to the close proximity of three negative charges.

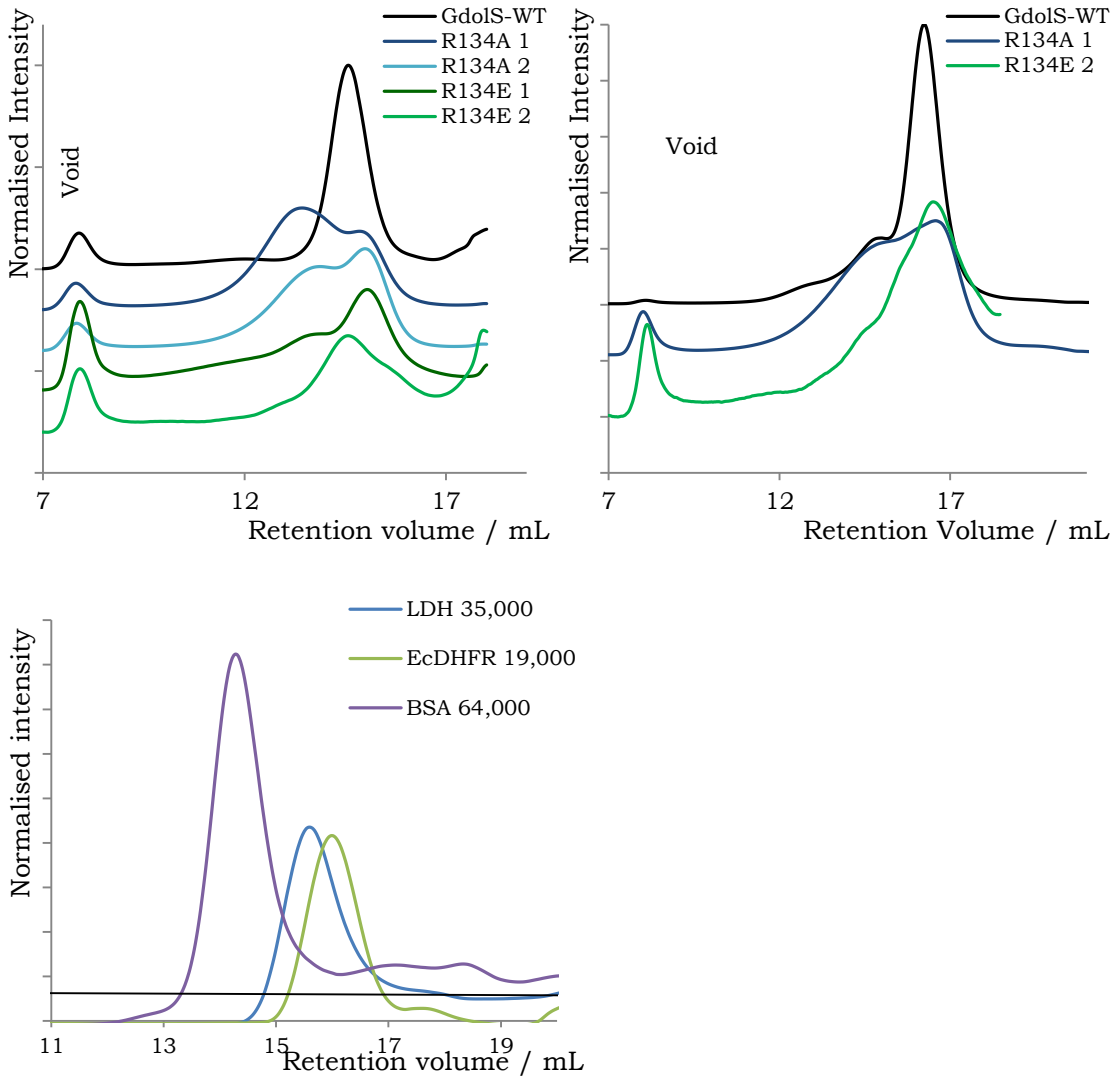


Figure 3.20 Normalised and overlaid size-exclusion chromatograms of GdolS-WT, GdolS-R134A (GdolS-R134A stored at 4 °C for 4 weeks, R134A 1; and freshly defrosted prior to injection, R134A 2) and GdolS-R134E (GdolS-R134E solubilised by basic extraction, R134E 1; and soluble protein directly from purification of expression, R134E 2), in the absence (Left) and presence (Right) of 150 mM NaCl.

In addition to the size-exclusion chromatography, non-denaturing gel electrophoresis was used to analyse the protein. Blue native polyacrylamide gel electrophoresis (BN-PAGE) was used, following the protocol developed by Schägger *et al.* which uses Coomassie brilliant blue dye as the charge shift agent.[121,122] The charge shift agent is necessary so that the mobility of the proteins is dependent on their charge-to-mass ratio, rather than shape of natural charge state. Coomassie brilliant blue dye is used because it complexes to the protein without denaturing it, allowing the protein to run according to size while also maintaining levels of higher order structure such as the formation of complexes. The GdolS-R134A was analysed by BN-PAGE in addition to GdolS-WT, other proteins and NativeMark™ protein standard from Life Technologies, Figure 3.21.

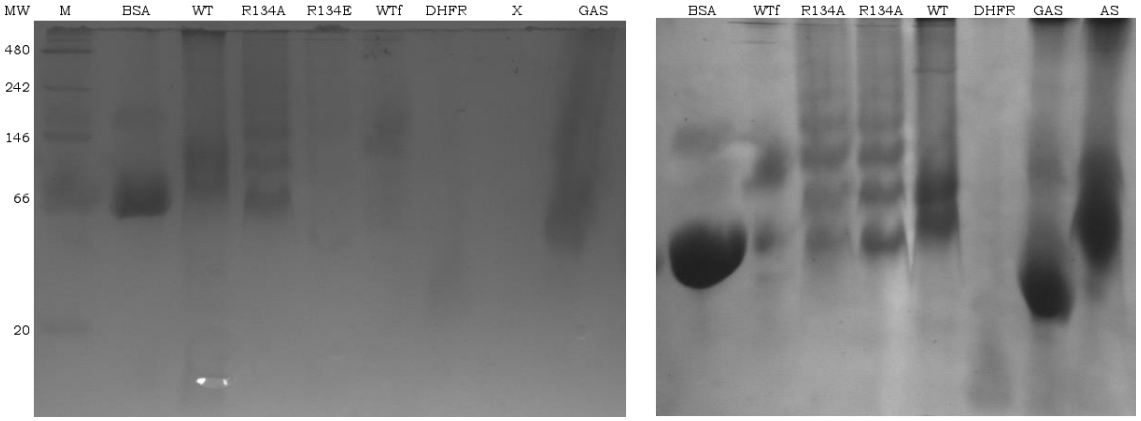


Figure 3.21 BN-PAGE analysis of GdolS-WT and GdolS-R134A against marker (M) and other proteins. BSA, bovine serum albumen, MW 66.4; WT, freshly defrosted GdolS-WT, MW 38.7; R134A, GdolS-R134A, MW 38.7; R134E, GdolS-R134E, MW 38.7; WTf, fridge stored GdolS-WT, MW 38.9; DHFR, dihydrofolate reductase from *E. coli*, MW 18.0; GAS, germacrene A synthase, MW 64.0; AS, PR-AS, MW 38.0. MW, molecular weight x10^3.

The BN-polyacrylamide gels do not definitively show the molecular weight of neither the GdolS-WT nor the GdolS-R134A mutant (the sample of the GdolS-R134E mutant is too dilute to be visible on the gel), with both samples running at just below the 66,000 MW band on the protein ladder. However the gel clearly shows bands of increasing MW in the samples of GdolS-R134A. In the samples of GdolS-WT, there are two bands with the second band at approximately a MW of 120,000, the same molecular weight as the second band in the samples of R134A. In the sample of WTf the second, heavier molecular weight band, correlates to the third band in the R134A samples; with a molecular weight of approximately 150,000. This suggests that the GdolS-WT is present in solution as an equilibrium of two forms, which tends towards aggregation upon long-term storage, this is consistent with the observed formation of a precipitate over the long term storage of the protein in

the fridge. The BN-polyacrylamide gel shows that the GdolS-R134A mutant is present in a greater number of forms, tending towards aggregation much quicker than the wild-type protein.

The result of the BN-PAGE in conjunction with the size-exclusion data suggests that the solution state of GdolS-WT is a monomer, and that while the dimer in the crystal state may be directed by the saltbridges; in the solution state these residues aid in the solubility of the enzyme and prevent dimerization and aggregation.

3.7 Summary

In this chapter we have used site-directed mutagenesis to probe, primarily the active site of, GdolS. Initial point mutations were directed at the metal binding motives, which gave results in good agreement with previous studies on other terpene synthases; with mutations to the NSE motif and the first aspartate in the aspartate rich region dramatically affecting catalytic efficiency while mutations to the remaining residues in the aspartate rich region caused on minor decreases in efficiency. The most interesting mutation was that to the arginine of the NSE-motif. Mutation of this residue to a glutamine caused, along with a severely impaired efficiency, a dramatic shift in the product profile of GdolS, to yield 45 % of germacrene A (**34**) in addition to the natural product, germacradien-4-ol (**79**). Further mutations to this residue caused a return to wildtype product profile (N218L), or inactive mutants (N218E, N218T). Nonetheless, this mutation indicates that, in GdolS, the metal-binding motif may also be actively involved in product specificity and locating of water in the active site.

A crystallographic study on AT-AS by Chen *et al.* highlighted an active site bound water molecule. This molecule was coordinated to the arginine of the NSE-motif, arginine-213; in addition to arginine-299 and serine 303. The corresponding residues in GdolS were tyrosine-303 and glutamate-307 which had serendipitously, been mutated and shown a slight moderation in product profile although not to the extent seen in the N218Q mutant. Both these residues showed more of a role in maintaining enzyme activity than in guiding product selection, with minor mutations (tyrosine to phenylalanine, glutamate to glutamine) causing small decreases in catalytic activity while greater mutations (tyrosine to alanine, glutamate to threonine) severely impacted catalytic efficiency.

To further investigate interactions between arginine-218 and proximal residues a helix-switch approach was adopted; changing the terminal loop-helix region of GdolS with the same, slightly longer region, from DCS; a δ-cadinene synthase available in our lab. Unfortunately this generated hybrid showed only sparing solubility due to incomplete folding and was completely inactive.

Following a report on the effect of the H-1 α loop on product distribution in fungal sesquiterpene synthases, this region was analysed and mutated in GdolS. Two mutants were generated to disrupt ionic interactions between the H-1 α loop and the active site, resulting in one wildtype-like enzyme (R229E) and an inactive mutant (D231L); suggesting that the aspartate residue is key in closing the active site to form the Michaelis complex.

The X-ray crystal structure of GdolS appeared to show a dimer and closer analysis of the interface between the two units of GdolS appeared to show a salt bridge comprised of two glutamate residues and an intercalating arginine. Replacements of this arginine with glutamate and alanine generated two functional mutants, which were analysed by size-exclusion chromatography and native-polyacrylamide gel electrophoresis. The R134E mutant was prone to rapid aggregation and the size-exclusion chromatography showed aggregation in addition to peaks for monomer and dimer. This was also the case for the R134A mutant, but with less aggregation. Due to the low concentration of the R134E sample it was not visible on the BN-polyacrylamide gel, but the R134A sample showed clearly the presence of a number of different size forms suggesting that while in the crystal structure these residues direct the dimer formation, in solution they are important for solvation of the protein and help to prevent aggregation.

4 Synthesis and evaluation of farnesyl diphosphate analogues

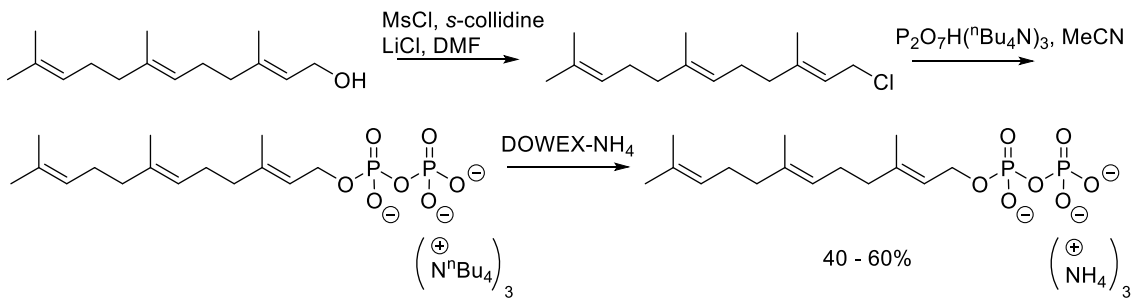
4.1 Preface

This chapter will describe the rationale behind and synthesis of, several FDP-analogues and their incubation with a range of sesquiterpene synthases. The sesquiterpene synthases used in this study have been described in Section 1.4.

4.1.1 Diphosphorylation of farnesol analogues

To generate FDP-analogues the corresponding farnesol-analogues were prepared and then diphosphorylated. Two methods of diphosphorylation were used in the preparation of the FDP-analogues described herein; a direct displacement of the farnesyl chloride (**103**) with a diphosphate salt and the condensation of farnesol (**104**) with inorganic phosphate using trichloroacetonitrile (**105**) as an activating agent.

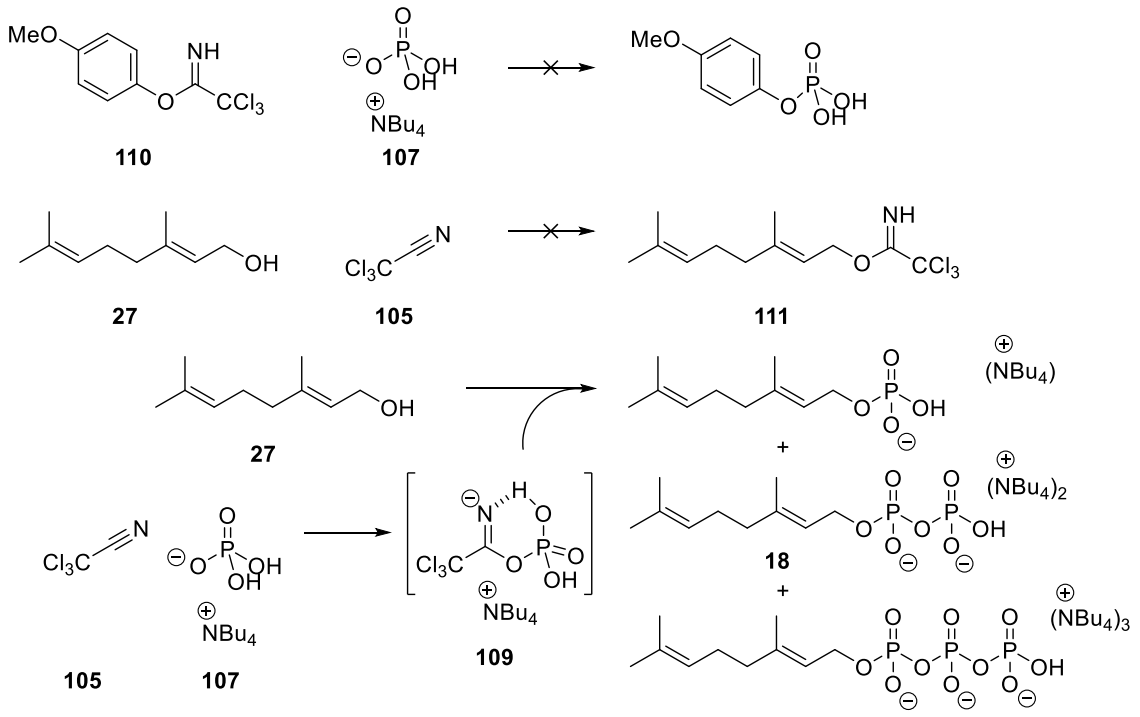
The displacement method used was that described by Davisson *et al.*[123] This method involves the generation of the farnesyl chloride (**103**) from farnesol (**104**) *via* an intermediate mesylate; the farnesyl chloride (**103**) is isolated and used without further purification, stirring with tris-(tetrabutylammonium)hydrogen diphosphate (**106**) overnight (15 h) in dry acetonitrile before removal of the solvent. The crude farnesyl diphosphate tris(tetrabutylammonium) salt is then passed through a column of DOWEX 50W-X8 cation exchange resin (ammonium form) to generate the trisammonium farnesyl diphosphate salt which is purified by HPLC and pure fractions lyophilised to yield the product in, typically, 40 to 60 % yield from the farnesol (**104**), Scheme 4.1. While the tris(tetrabutylammonium) farnesyl diphosphate salts may be used without ion exchange they are harder to handle, being a sticky solid, whereas the trisammonium salts are isolated as a dry, white powder.



Scheme 4.1 Diphosphorylation of farnesol *via* the intermediate farnesyl chloride as developed by Davisson *et al.*[123]

The condensation method followed was as described by Keller *et al.*,[124] a modification of the procedure originally described by Cramer[125] and modified by Cornforth,[126] involving mixing an alcohol with trichloroacetonitrile (**105**) and inorganic phosphate. The original method as described by Cramer[127] produces a complex mixture of mono-, di- and triphosphates; Keller *et al.*[124] have reported the optimisation of the protocol to improve the yield of the diphosphate and also describe the purification of the diphosphate using flash column chromatography on silica. Farnesol (**104**) was incubated with inorganic phosphate and trichloroacetonitrile (**105**) in acetonitrile for 15 minutes at 37 °C (three additions at five minute intervals) with mixing on a rotary evaporator and then the whole reaction mixture applied to a silica column pre-equilibriated with running buffer (^{i}PrOH:NH_4OH:H_2O, 6:2.5:0.5), fractions were analysed by TLC (^{i}PrOH:NH_4OH:H_2O, 6:3:1) and the diphosphate isolated and lyophilised to yield the product in 20 – 40 % yield. The inorganic monophosphate used can be tetrabutylammonium dihydrogen phosphate (**107**), which is available commercially, or bis(triethylammonium) hydrogen monophosphate (**108**) which can be prepared prior to use as a solution in acetonitrile.[128] When using the bis(triethylammonium) hydrogen phosphate (**108**), it was found during this study that the triethylammonium counter ions are readily exchanged during the silica purification to yield the trisammonium farnesyl diphosphate salt as a white powder directly. Use of the tetrabutylammonium dihydrogen monophosphate (**107**) results in the tetrabutylammonium farnesyl diphosphate salt, which requires an additional ion-exchange column. Lira *et al.* proposed a mechanism for this reaction *via* the formation of a mixed anhydride (**109**) formed between the trichloroacetonitrile (**105**) and the monophosphate, Scheme 4.2.[129] Following experiments by ^{13}C-NMR spectroscopy no reaction was detected when 4-methoxybenzyl chloroacetimidate (**110**) was mixed with tetrabutylammonium dihydrogenphosphate (**107**), nor upon the premixing of geraniol (**27**) with

trichloroacetonitrile (**105**). These results rule out a mechanism through activation of the alcohol by trichloroacetonitrile (**105**) to form a trichloroacetimidate intermediate (**110**, **111**) that is then displaced by the phosphate. When geraniol (**27**), trichloroacetonitrile (**105**) and tetrabutylammonium dihydrogenphosphate (**107**) were mixed in succession, the disappearance of the ^{13}C-signals of trichloroacetonitrile (**105**) were noted, in addition to the appearance of the signals corresponding to the phosphorylated alcohol (**18**).



Scheme 4.2 Diphosphorylation using the trichloroacteonitrile coupling, *via* a mixed anhydride.[129]

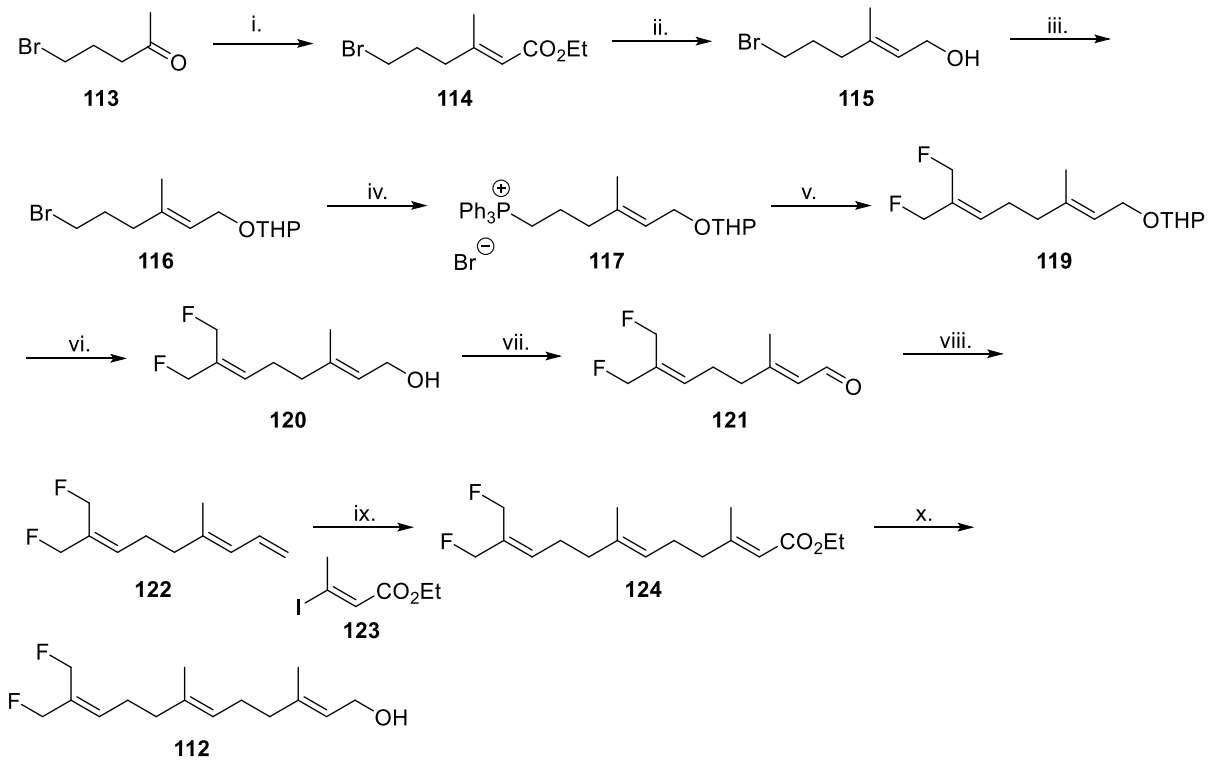
While the direct displacement method is superior for the preparing the diphosphate salts in higher yield and purity necessary for use in kinetic studies, it is a time consuming and protracted method. The trichloroacetonitrile coupling is a time-efficient method for the preparation of diphosphate salts suitable for the exploratory work described in this chapter. Both methods have been used for the diphosphorylation of the farnesol analogues described herein.

4.2 FDP analogues

4.2.1 12,13-Difluorofarnesyl diphosphate (73)

The 12,13-difluoro-FDP (**73**) was prepared as part of the mechanistic investigation of GdolS as described earlier (Section 2.3.2). While the 12,13-difluoro-FDP (**73**) has been synthesised previously, it was necessary to repeat

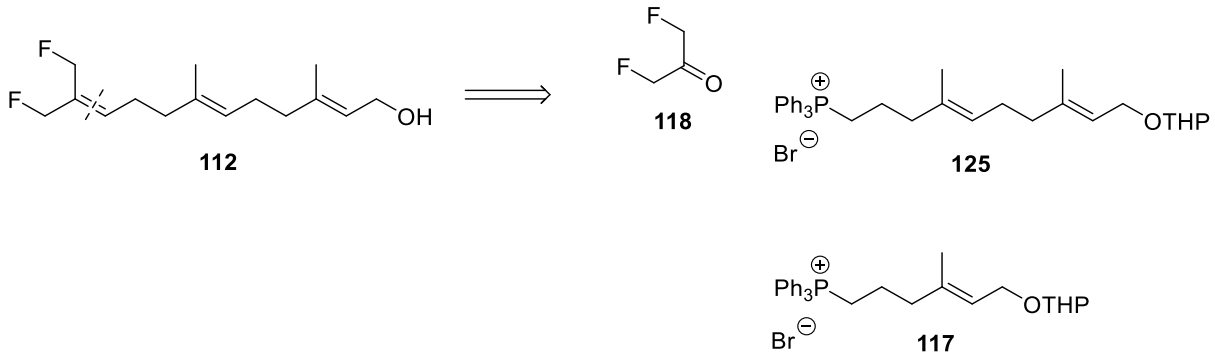
the preparation and a new synthesis was devised. The previous synthesis of 12,13-difluorofarnesol (**112**) by Miller and Yu was completed from bromoketone **113** using three chain extension steps.[67] A Horner-Wadsworth-Emmons reaction was used to generate the bromoester **114** with good stereoselectivity (*E*/*Z* 11:1); Functional group interconversions then yielded the allylic alcohol protected as the tetrahydropyran ether (OTHP) **116**. This was used to prepare the phosphonium bromide **117** which underwent a Wittig reaction with difluoroacetone **118** to yield the THP protected difluorogeraniol **119**. The final chain extension was completed using a Heck coupling with diene **122** and crotonyl iodide **123** to yield the ester **124** that, upon reduction, yielded the desired 12,13-difluorofarnesol (**112)**.



Scheme 4.3 Original synthesis of 12,13-difluoroarnesol (**112**) by Miller and Yu. *Reagents and conditions: i. triethyl phosphonoacetate, NaH, THF, 0 °C, 68%; ii. DIBAL-H, THF, -78 °C, 98%; iii. 3,4-DHP, DCM, p-TsOH, 98%; iv. PPh_3, MeCN, Δ; v. LiHMDS, THF, -78 °C, then **118**, 87%; vi. PPTS, EtOH, 55 °C, 97%; vii. TPAP,NMO, MeCN, 82%; viii. CH_3PPh_3Br, n-BuLi, THF, 78%; ix. 9-BBN, THF then **123**, PdCl-dppf, NaOH, $AsPh_3$,THF, 50 °C, 28%; x. DIBAL- H, THF, -78 °C, 68%.*

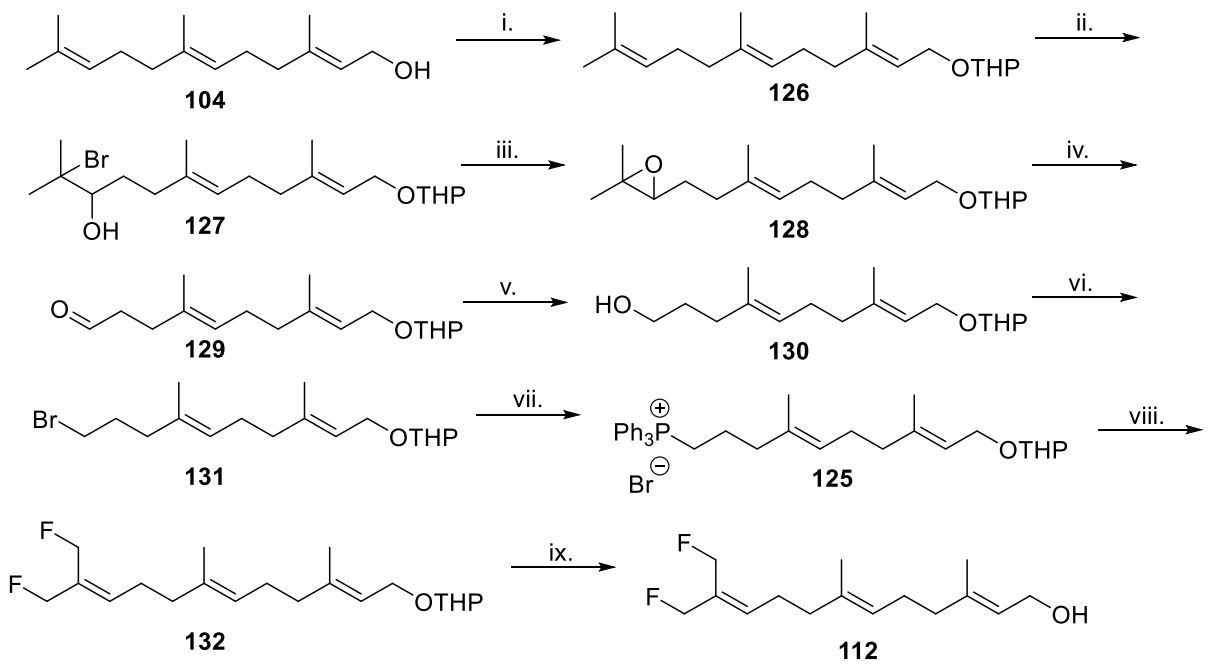
A new synthesis was devised to reduce the number of chain extension steps using farnesol as the starting material, which also removed any issues of stereoselectivity in the synthesis. The new method was based upon the using a similar Wittig reaction to Miller and Yu, with difluoroacetone for a single key extension step, but using the longer phosphonium bromide (**125)** bearing an additional isoprene unit, Scheme 4.4. Phosphonium bromide **125** was

prepared from farnesol (**104**) using a number of literature procedures,[67,130,131] and generated intermediate compounds that have been utilised in the preparation of a number of other targets.



Scheme 4.4 Retrosynthesis of 12,13-difluoroarnesol (**112**)

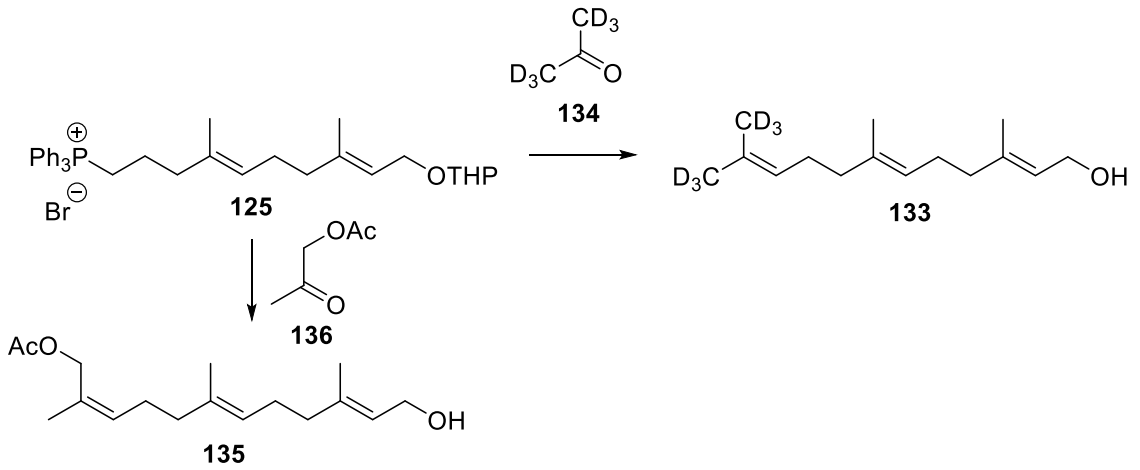
Farnesol (**104**) was first protected using 3,4-dihydropyran and catalysed by pyridinium *para*-toluene sulfonate (PPTS) to generate the tetrahydropyran ether (**126**) in 98% yield.[130] The THP protecting group was chosen despite the likely complication of the ^{1}H-NMR spectra due to the introduction of an asymmetric carbon. The distal (10,11)-double bond of farnesol can be regioselectively epoxidised in a two-step procedure *via* the bromohydrin. Reaction of **126** with *N*-bromosuccinimide (NBS) in aqueous saturated THF yields the bromohydrin **127** in 70% yield, the epoxide **128** is then formed by base catalysed ring closure. Epoxide **128** then undergoes oxidative cleavage with HIO_4, to yield aldehyde **129** and acetone.[131] Aldehyde **129** is readily reduced to the alcohol **130** with $NaBH_4$, which can then be converted to bromide **131** *via* a mesylate intermediate. Heating bromide **131** under reflux in acetonitrile with triphenylphosphine yielded the desired phosphonium bromide **125**, which was partially purified by repeated trituration with diethyl ether to give a white solid.



Scheme 4.5 Redesigned synthesis of 12,13-difluorofarnesol. *Reagents and conditions: i. 3,4-DHP, DCM, PPTS, 0 °C, 98 %; ii. NBS, THF/H_2O, 70 %; iii. K_2CO_3, MeOH, 99 %; iv. H_5IO_6, $NaIO_4$, THF, 0 °C, 79 %; v. $NaBH_4$, MeOH, 0 °C, 69 %; vi. Et_3N, MsCl, -45 °C then LiBr, RT, 79 %; vii. PPh_3, MeCN, Δ; viii. n-BuLi, THF, -45 °C, then **118**, 61 %; ix. pTsOH, EtOH, 91%.*

The ylid was formed by the dropwise addition of n-butyl lithium to a solution of phosphonium bromide **125** in THF at -45 °C, generating a deep, bright yellow colour. Dropwise addition of **118** to the ylid caused a lessening in the intensity of the colour, and the reaction was allowed to warm to room temperature over 2 hours.[67,132] Removal of the THP-group with p-TsOH in methanol from the resulting THP-ether (**132**), yielded the 12,13-difluorofarnesol **112**.

The symmetry of the aldehyde makes this a simple and effective method for the generation of farnesol analogues with identical modifications at the C12 and C13 positions and has also been used for the preparation of [12,12,12,13,13,13-$^2H_6$]-farnesol (**133**) using d_6-acetone (**134**).[79] This method was also used in an attempt to prepare the 13-hydroxy-farnesol **135** using acetoxyacetone **136** as the aldehyde, Scheme 4.6. While this method demonstrated high selectivity for the (*Z*)-isomer (86-90 %, Figure 4.1) the yield of the Wittig reaction was poor and further work by a project student, Simon Robinne, indicated that this problem was due to the nature of the ketone used, acetoxyacetone (**136**).[133,134] Similarly poor yields (< 10 %) of product were isolated from reactions of **136** with methyl and phenyl phosphonium halides.[135] With further development and a change in protecting group this method may provide access to the 13-hydroxy-FDP but this was not pursued further in the course of this study.



Scheme 4.6 Generation of other FDP-analogues through alternative Wittig reactions using phosphonium bromide **125**

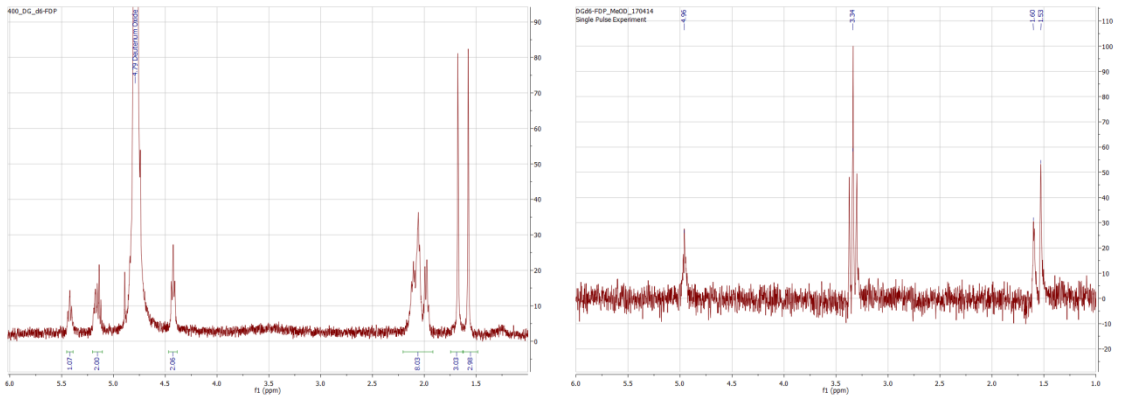
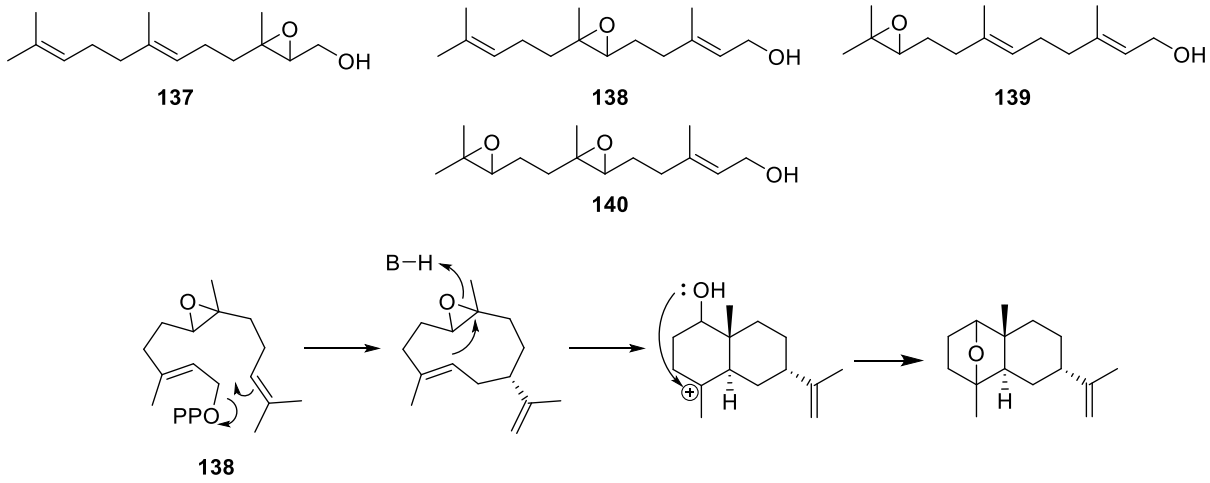


Figure 4.1 ^{1}H-NMR spectra (300 MHz, $CDCl_3$) of 12-OAc-farnesyl THP ether (upper) and 13-OAc-farnesol **135** (lower) with the CH_2O peaks picked and integrated. ^{1}H-NMR (300 MHz, MeOD, bottom left) and ^{2}H-NMR spectra (300 MHz, MeOH, bottom right) of [12,13-$^{2}H_6$]-FDP.

4.2.2 Epoxy-FDPs

Preparation

A series of epoxy-farnesols were prepared to investigate the ability of a range of sesquiterpene synthases to turn over FDP-analogues with the aim of opening the epoxide in the cyclisation cascade and trapping carbocation intermediates, generating novel terpenoids, Scheme 4.7. To this end, each of the double-bonds in farnesol was individually epoxidised to prepare the 2,3-, 6,7- and 10,11-epoxy-farnesols (**137**, **138** and **139,** respectively). The (6,7-10,11)-diepoxy-farnesol (**140**) was also prepared, Scheme 4.7. The epoxides were prepared as racemates due to the exploratory nature of this investigation.



Scheme 4.7 Epoxy-farnesol target compounds (top) and conversion of 6,7-epoxy-FDP to a novel terpenoid by AS (below).

While it is possible to regiospecifically epoxidise the terminal 10,11-double bond as described earlier (**128**, Scheme 4.5),[131] through formation of the bromohydrin **127**, direct regiospecific epoxidation of C6,C7 double bond is not known. Lichtor and Miller have been able to increase the levels of C6,C7 epoxidation of farnesol through their development of a short peptide catalyst;[136,137] Using the peptide catalyst they were able to produce up to a 4:1 mixture of the desired 6,7-epoxy-farnesol (**138**) with the 10,11-epoxy-farnesol (**139**) in a 43 % yield, however this would require preparation of the necessary peptide and would still need further purification to procure the desired 6,7-epoxy-farnesol (**138**). Because separation of the epoxy-farnesols was possible we instead chose to prepare a mixture of the epoxides using 3-chloroperbenzoic acid (mCPBA) and to perform the separation on the epoxy-farnesol mixture; Epoxidation of farnesyl acetate (**141**) with one equivalent of mCPBA for 1 hour yielded a mixture of unreacted starting material (**141**, 13

%), the monoepoxides (**142** & **143**, 75 %) in a ratio of 58:42, and a small amount of the diepoxide (**144**), Scheme 4.8. The identification of the two epoxides was made through comparison with the literature and with NMR-spectra of **128** showing that the major epoxide was the 10,11-epoxyfarnesyl acetate (**142**).

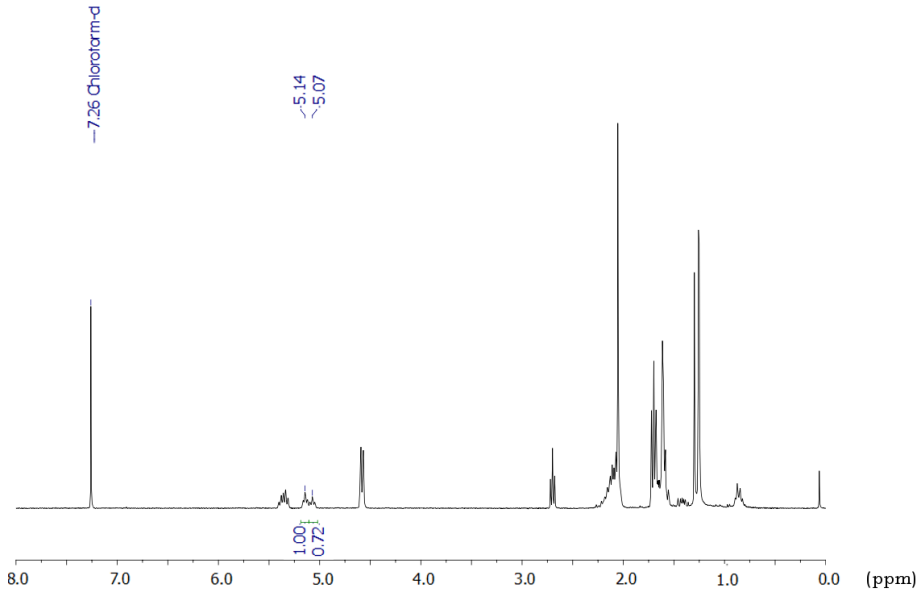


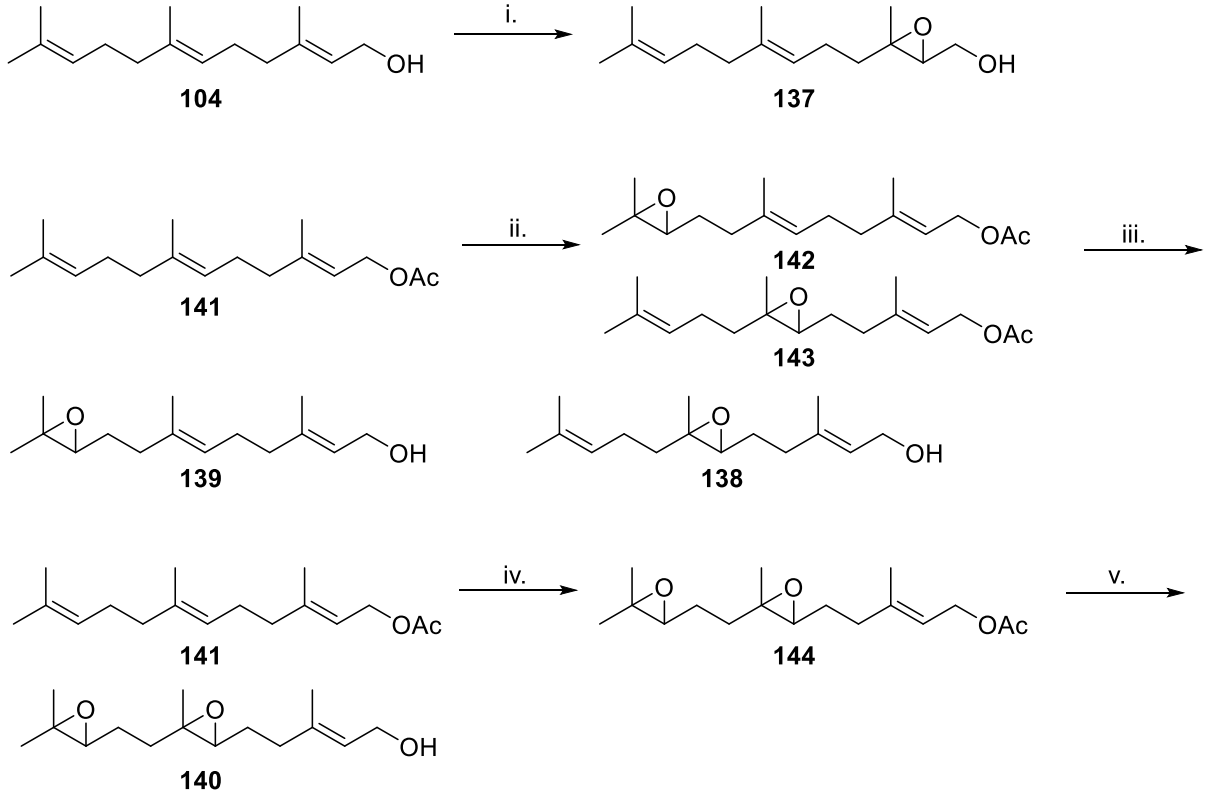
Figure 4.2 ^{1}H-NMR spectrum (300 MHz, $CDCl_3$) of the mixed epoxyfarnesyl acetates

These compounds were separated by column chromatography on silica gel and the mixture of the mono-epoxyfarnesyl acetates (**142** & **143**) was deprotected quantitatively with potassium carbonate. The separation of the epoxyfarnesols (**138**, **139**) was effected using an automated chromatography system over a gradient of 10 – 60 % diethyl ether in petroleum ether yielding 21% of the 6,7-epoxyfarnesol (**138**), 32 % of the 10,11-epoxyfarnesol (**139**) and 20 % of the unresolved mixture. Resolution of the two epoxides by thin layer chromatography required multiple developments across the same gradient, with the 10,11-epoxyfarnesol (**139**) marginally less polar than the 6,7-epoxide (**138**).

The diepoxyfarnesol (**140**) was prepared in the same fashion as the mono-epoxides, oxidising farnesyl acetate (**141**) with 2-equivalents of mCPBA for 1 hour at 0 °C and hydrolysis of the acetate group to yield **140** in 65 % yield over two steps, Scheme 4.8.

The 2,3-epoxyfarnesol (**137**) could be prepared efficiently and in high yield, directly from farnesol (**104**) using the Sharpless epoxidation,[138] Scheme 4.8, this method would also allow stereospecific synthesis of the two enantiomers

by using the Sharpless asymmetric epoxidation, with titanium tetraisopropoxide and (*R*,*R*)- or (*S*,*S*)-diethyl tartrate to impart the stereospecificity.[139]



Scheme 4.8 Synthesis of epoxy-farnesols. *Reagents and conditions: i. $VO(acac)_2$, Toluene, t-BuOOH, 0 °C, 95 %; ii. mCPBA 1 eq., DCM, 0 °C, 1 h, 75 %; iii. K_2CO_3, MeOH, 100 %; iv. mCPBA 2 eq., DCM, 0 °C, 1 h; v. K_2CO_3, MeOH, 65 % over two steps.*

The diphosphate salts were prepared from the farnesol analogues as discussed at the start of this chapter (Section 4.1.1), to yield the epoxy-FDP analogues in 30 – 68% yield. Analytical incubations were prepared with the range of sesquiterpene synthases available (Section 1.4).

Table 4.1 Methods of diphosphorylation for the epoxyfarnesols

Analogue	Method	Yield (%)
2,3-epoxyfarnesol (**137**)	Kramer	68
6,7-epoxyfarnesol (**138**)	Poulter	58
10,11-epoxyfarnesol (**139**)	Poulter	66
diepoxyfarnesol (**140**)	Kramer	30

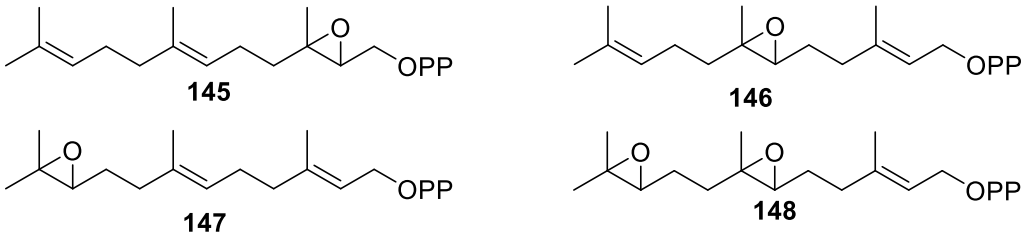
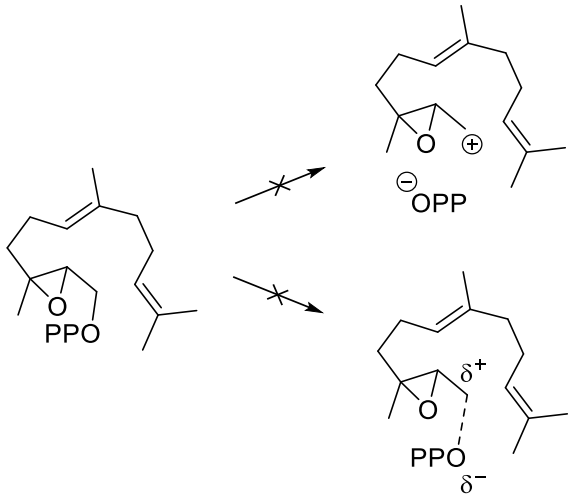


Figure 4.3 Epoxy-FDP analogues prepared in this study

2,3-Epoxy-FDP (145)

The 2,3-epoxy-FDP (**145**) was found not to be a substrate for any of the enzymes tested; with no pentane extractable products detected in the GC-MS analysis of the product mixture even after extended incubation of **145** with the enzymes. Formation of the farnesyl cation is thought to be the first step in the catalytic mechanisms of GAS, GDS, DCS, ADS, EBFS and GdolS. Stabilisation of the 2,3-epoxyfarnesyl cation is not possible, and without stabilisation formation of the primary carbocation is highly disfavoured. The cyclisation of FDP by PR-AS has been proposed to proceed through a concerted ring-closure and loss of the diphosphate; the failure of PR-AS to turnover **145** might suggest that this step is not a symmetrical S_N2 reaction. Instead the loss of the diphosphate group might begin to occur before the attack by the 10,11-double bond leading a developing positive charge on C1, Scheme 4.9.



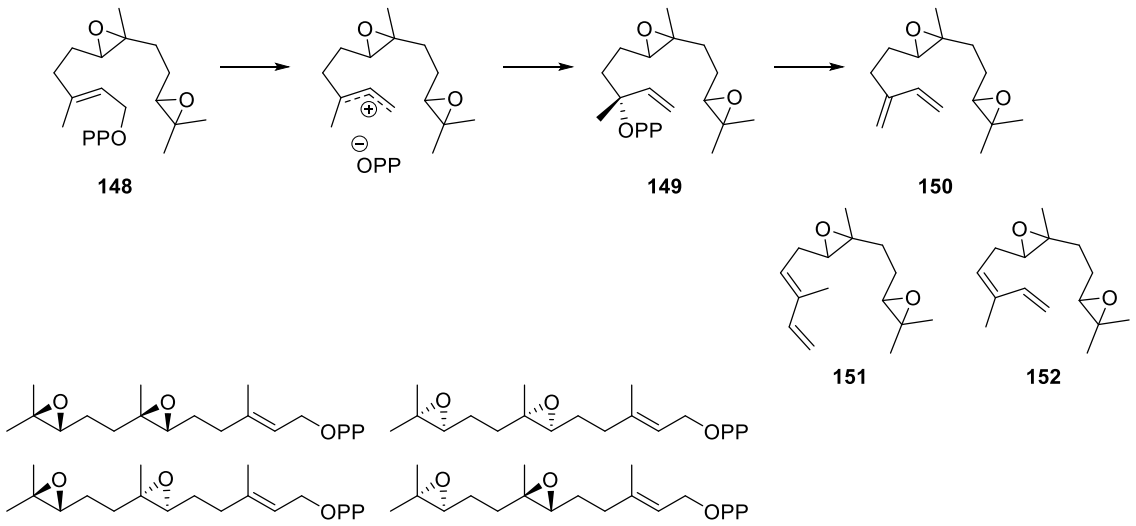
Scheme 4.9 Possible mechanisms for the failed turnover of **145**

In addition to the lack of stabilisation the introduction of the epoxide has changed the configuration at C2 and C3. In FDP, C1 to C4 and C15 are all co-planar due to the sp^2 nature of C2 and C3; In the 2,3-epoxy-FDP, C2 and C3 are now sp^3 in nature and tetrahedral with the added complexity of two

stereocentres. This change in conformation may alter substrate folding in the active site and does not provide productive alignment between the C10 and C1 for cyclisation.

Diepoxy-FDP (148)

The diepoxy-FDP (**148**) was found to be a substrate for EBFS, but not for any of the other enzymes tested. Analysis of the gas chromatogram of the pentane extracted products from incubation of **148** with EBFS showed four main products, although the total ion count for the gas chromatogram (TIC) was relatively low (1 x 10^4), indicating that the turnover of this substrate by EBFS was poor compared to FDP (TIC 6 x 10^5). Analysis of the mass spectra of the products is consistent with formation of diepoxy-farnesenes (**150** - **152**), which shows that EBFS is able to accept the analogue into the active site and catalyse the conversion to the NDP-derivative (**149**) and deprotonation. The presence of other farnesenes might suggest that the presence of the epoxides has an effect on the contouring of the farnesyl chain in the active site leading to an alternative deprotonation to form diepoxy-α-farnesenes (**151**, **152**) in addition to the diepoxy-(*E*)-β-farnesene (**150**), Scheme 4.10. Due to the racemic nature of the epoxidations the diepoxy-FDP (**148**) is a mixture of 4-stereoisomers, two pairs of enantiomers, and while the enantiomers are in separable by the gas chromatography system used, it is possible to separate the diastereomers further complicating analysis of the products by GC-MS.



Scheme 4.10 Turnover of **148** by EBFS, and the stereoisomers present in **148**

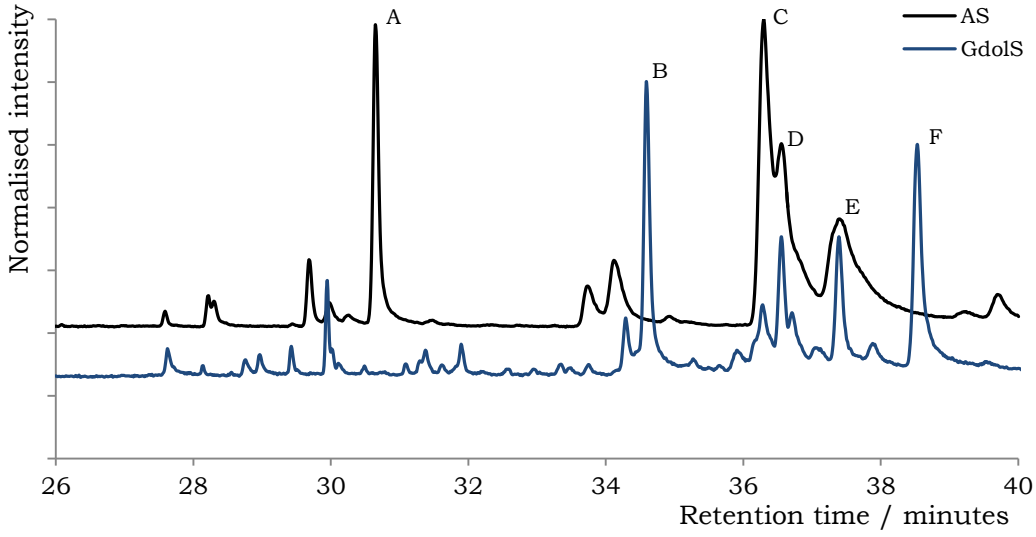
6,7-Epoxy-FDP (146)

The 6,7-epoxy-FDP (**146**) was not a substrate for GAS, GDS, DCS or ADS. That analogue was a good substrate for GdolS, PR-AS and EBFS, as judged by total

ion count (1-2 x 10^5). Incubation of **146** with EBFS produced 5 products, and analysis of the mass spectra was consistent with formation of farnesenes. Incubation with PR-AS produced two major products and a number of minor products, while incubation with GdolS produced multiple trace products and 4 products in greater proportions, Figure 4.4. While eluting with similar retention times, comparison of mass spectra showed none of the products produced by AS or GdolS were farnesenes.

All of the products show a molecular ion with m/z = 220, corresponding to $C_{15}H_{23}O$ products and the retention of the epoxide moiety. None of the products produced by GdolS display a m/z = 238, which would be consistent with the incorporation of an additional hydroxyl group.

While the mass spectra of compounds A and F, Figure 4.4, have a fragment of m/z = 161 indicative of loss of an isopropylidene group (-41 m/z, C_3H_5) the other peaks; B,C,D and E; have a major fragment at m/z = 159, -43 m/z. This loss of m/z = 43 is indicative of loss of an isopropyl fragment (C_3H_7), suggesting that while all of these peaks are 10-membered macrocycles peaks A and F may be germacrene A type products, with a double bond in the isopropyl tail. None of these products were isolated or definitively identified at this time due to the lack of a single overriding product.



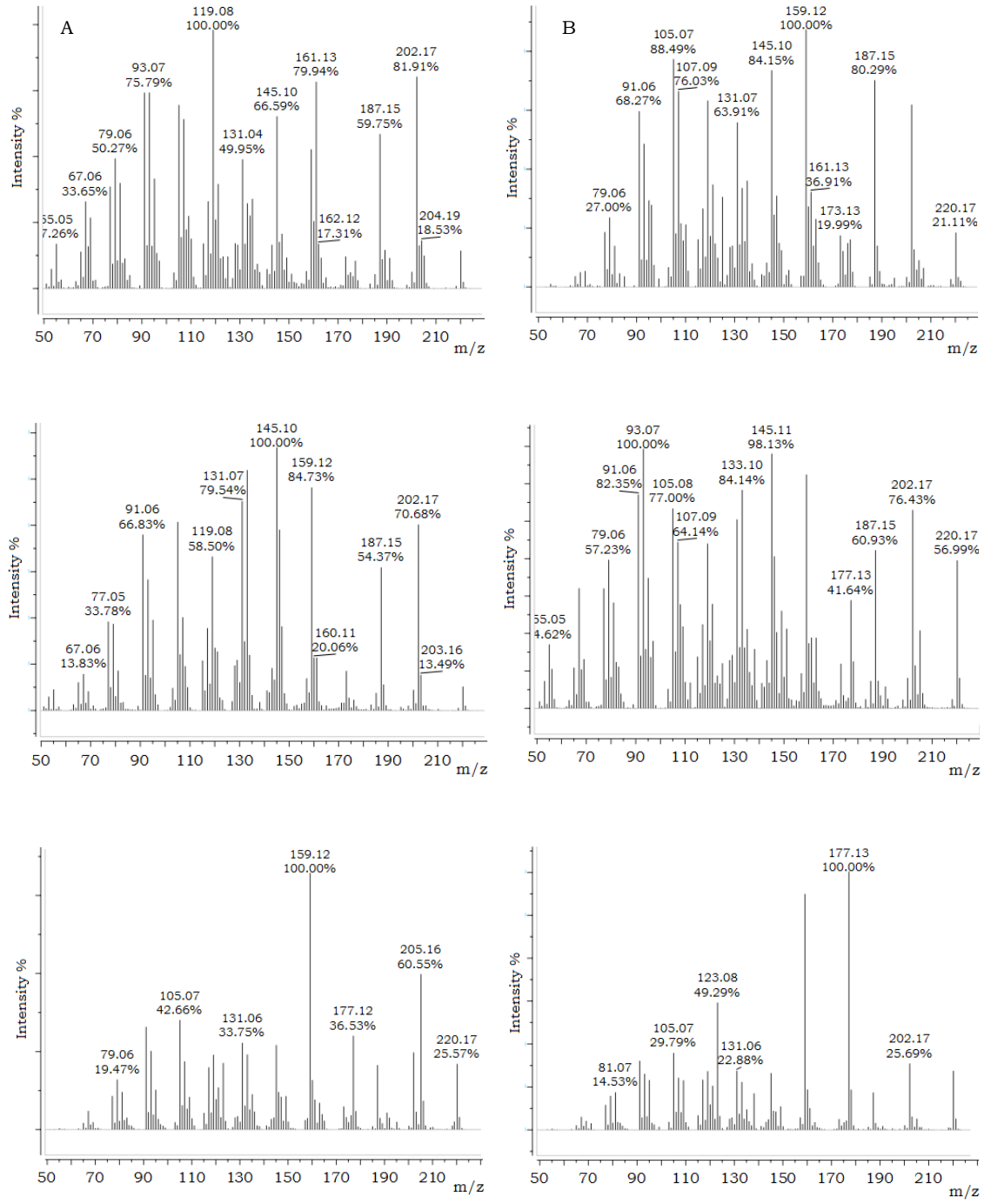


Figure 4.4 Overlaid normalised gas chromatograms of the pentane-extractable products arising from incubation of 6,7-epoxy-FDP with AS and GdolS. Mass spectra of the major compounds as labelled in the chromatogram (A-F).

10,11-Epoxy-FDP (147)

Incubation of the 10,11-epoxy-FDP (**147**) with EBFS produced a single product with a mass spectrum consistent with the 10,11-epoxyfarnesene. The TIC (9 x 10^5) suggested that this substrate is turned over by EBFS with a similar efficiency as FDP (3 x 10^5).

Incubation of **147** with GdolS yielded a single major product with a trace of the 10,11-epoxyfarnesene, **153**, Scheme 4.11. The mass-spectrum of the major

product had a molecular ion with m/z = 220, corresponding to a $C_{15}H_{23}O$ product.

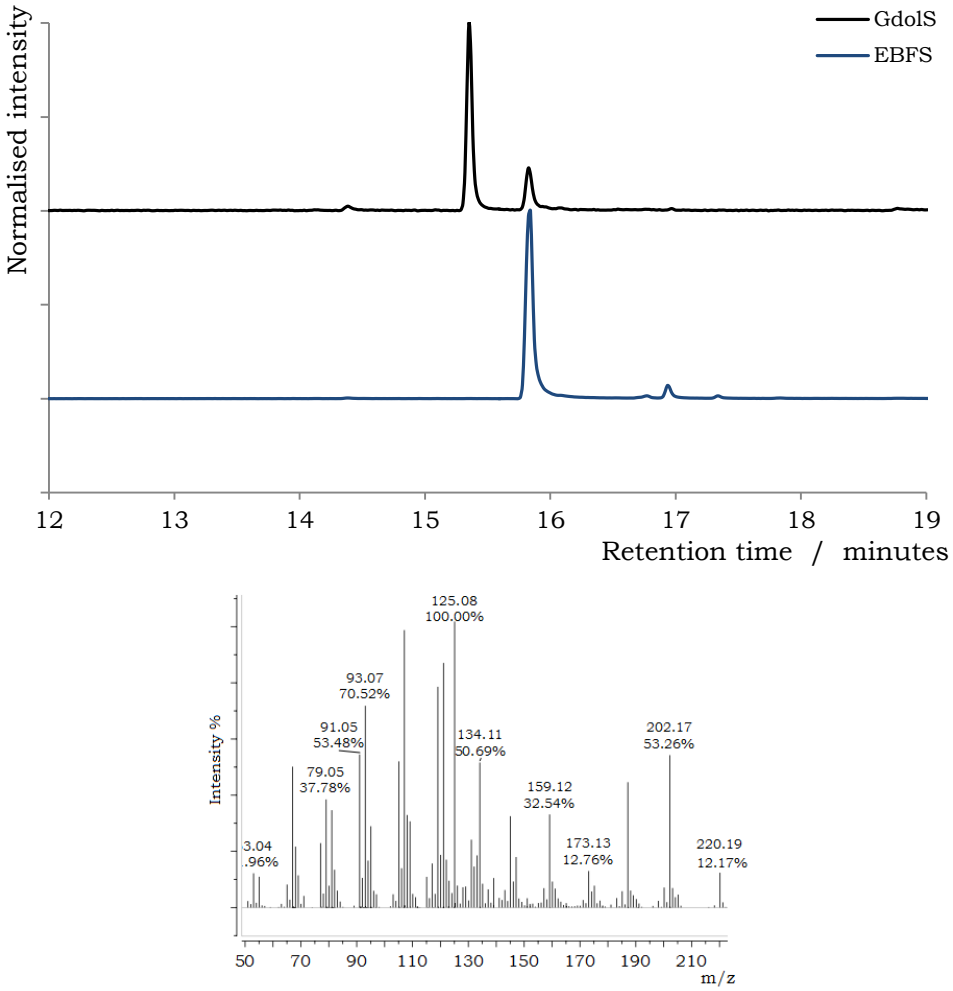
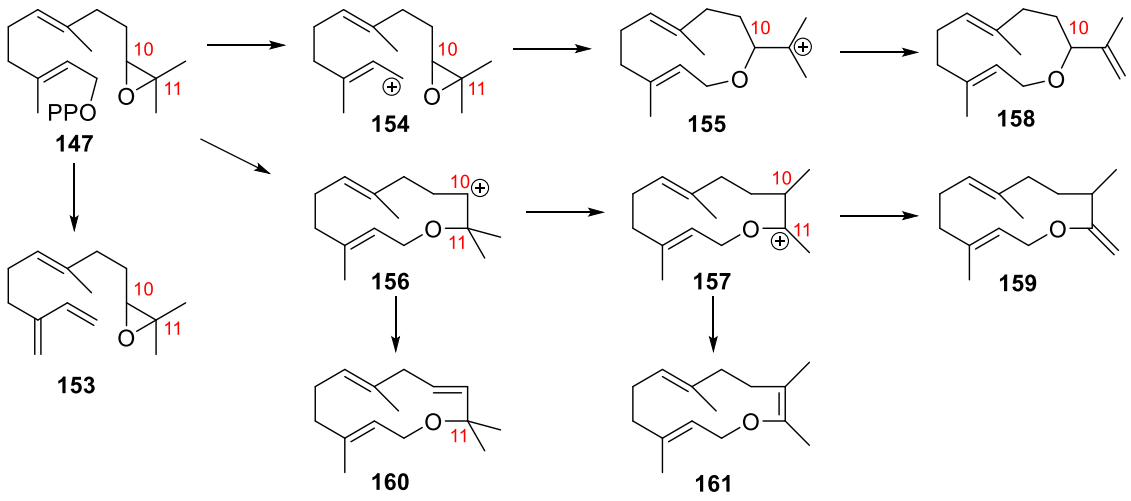


Figure 4.5 Gas chromatogram (top) of the pentane extractable products arising from incubation of 10,11-epoxy-FDP with GdolS and EBFS, and mass-spectra (bottom) of the major compounds.

The putative catalytic mechanism of GdolS was used to propose possible products. Assuming loss of the diphosphate occurs first, forming the 10,11-epoxyfarnesyl cation **154**, the epoxide can open at C10 or C11 to attack the cation and yield a macrocyclic ether. Opening of the epoxide at C10 would yield a tertiary carbocation **155**, more stable than the secondary carbocation **156**, generated by opening in the 11-position. However, a rapid C11-C10 methyl shift from **156** would yield the carbocation stabilised by conjugation with the oxygen, potentially leading to a long-lived intermediate (**157**). Further methyl- and hydride shifts followed by deprotonation of any of the intermediate carbocations could lead to a number of potential products (**158** – **161**), Scheme 4.11.



Scheme 4.11 Possible products of the cyclisation of 10,11-epoxy-FDP

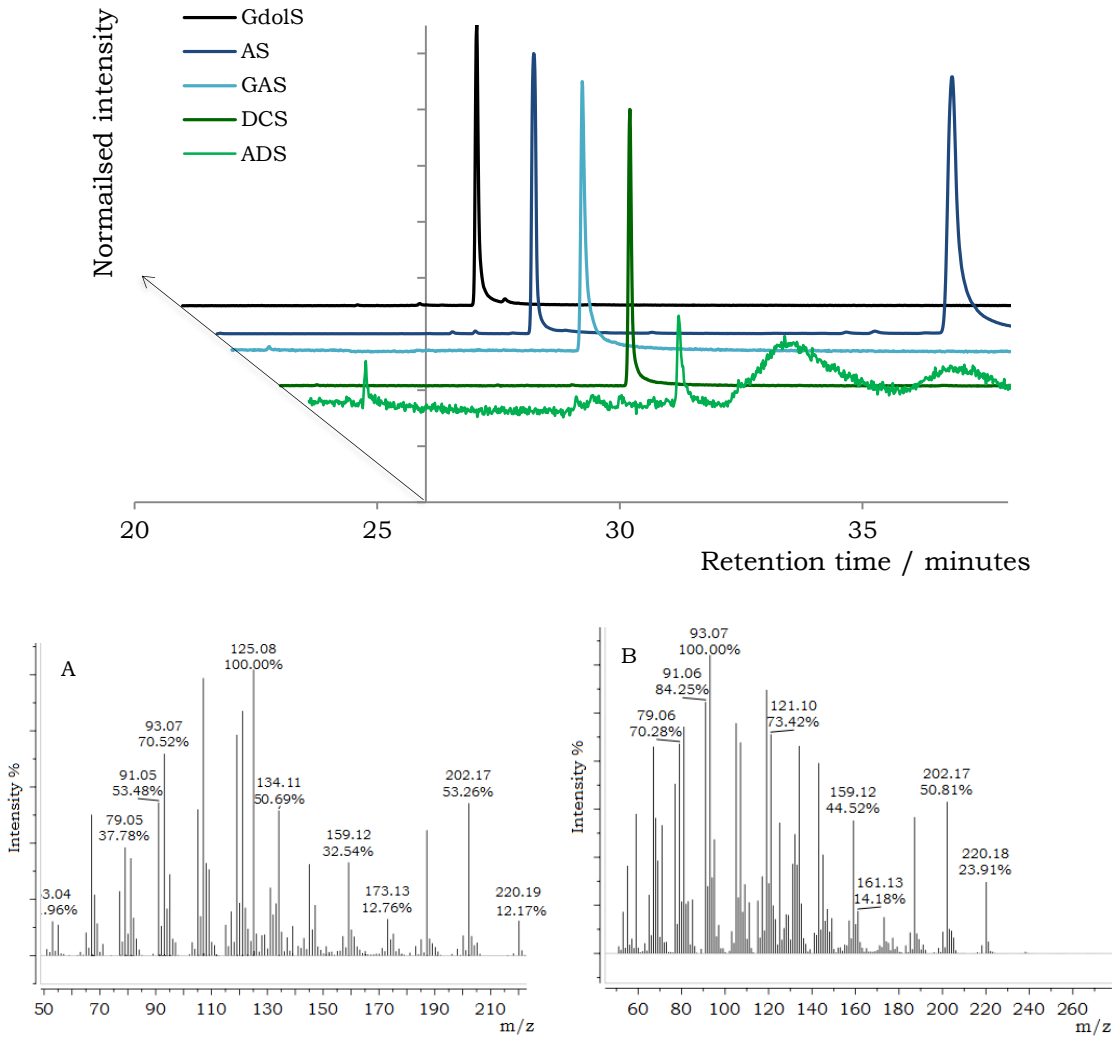


Figure 4.6 Overlaid normalised gas chromatogram the pentane extractable products arising from incubation of 10,11-epoxy-FDP with all enzymes. Mass spectra of the two compounds arising from incubation of 10,11-epoxy-FDP with AS as labelled in the chromatogram (A and B).

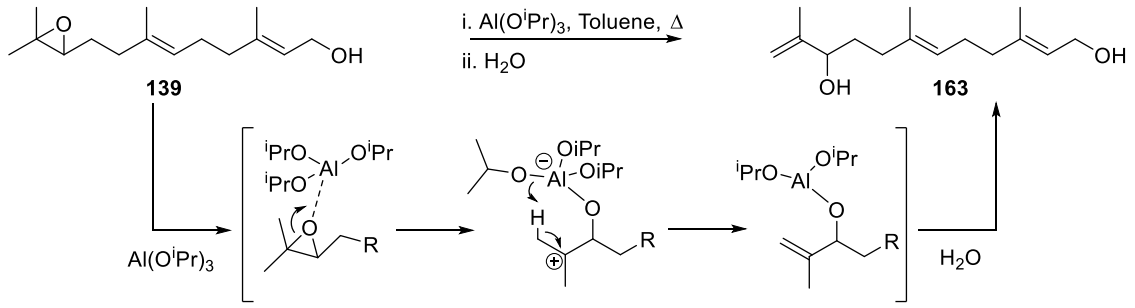
Incubations with AS, GAS and DCS also showed **147** to be a good substrate for these enzymes. With GAS and DCS the same compound was found to be the

sole product, while PR-AS also yielded this product in addition to a second product with a longer retention time. Analysis of the mass-spectrum of the second product with PR-AS showed a compound with *m/z* = 238, possibly suggesting a diol.

Incubations with ADS found 10,11-epoxy-FDP **147** to be an extremely poor substrate, with a small amount of the macrocyclic ether **158** discernible from the baseline. This analogue was not turned-over by GDS.

4.2.3 Second generation analogue

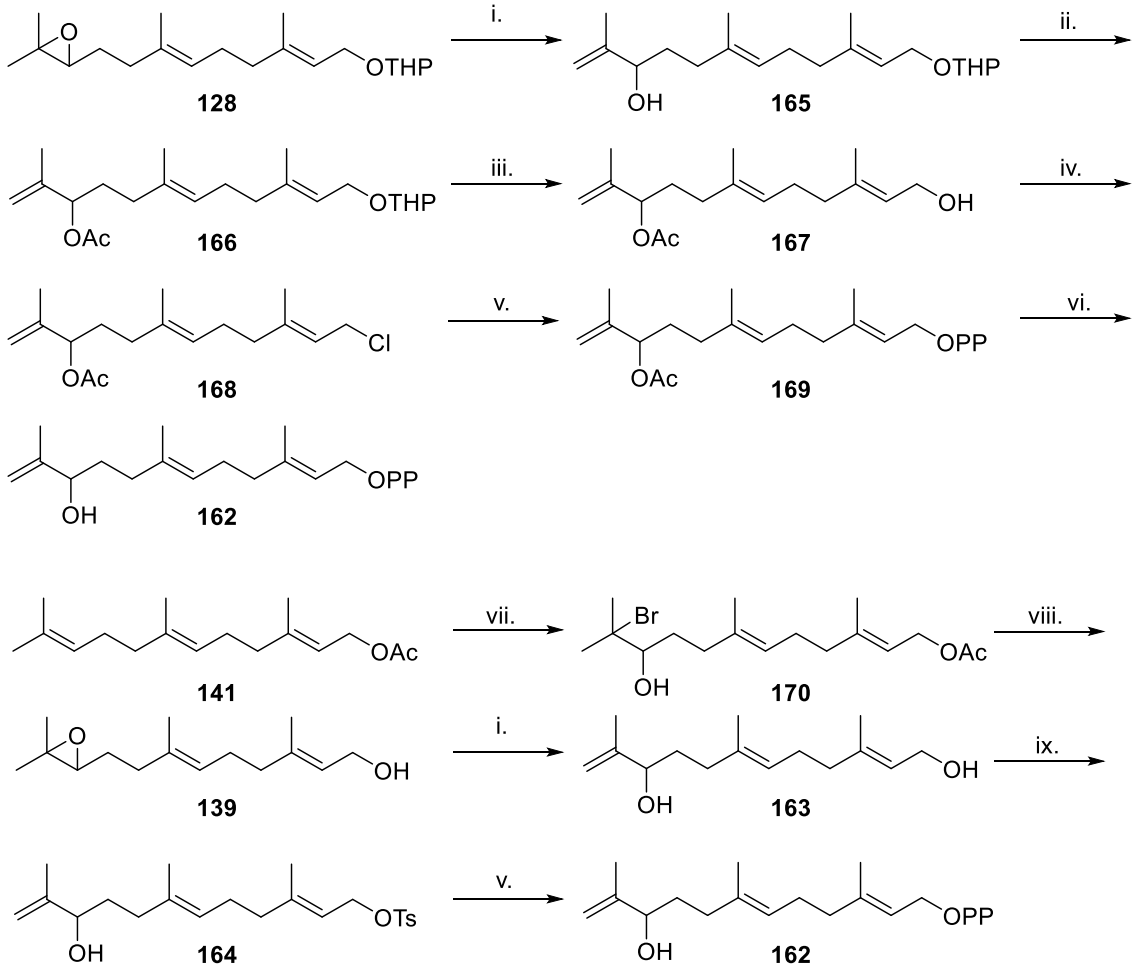
The 11-membered macrocyclic ether **158** was proposed to be the most likely product from incubation of 10,11-epoxy-FDP (**147**) with all enzymes due to the short and simple, germacrene A-like, catalytic mechanism from which it arises; generating only a single tertiary carbocation intermediate. To test this proposal another analogue was designed which, if a substrate, would trap the farnesyl cation and generate the proposed macrocyclic ether directly without the generation of further intermediate carbocations. The designed analogue, 10-hydroxy-11-ene-FDP (**162**), could be prepared by heating a 10,11-epoxyfarnesol (**139**) derivative with aluminium isopropoxide in toluene in a sealed tube; this variation of the Meerwein-Ponndorf-Verley reduction gave the desired product in 65% yield.[140,141]



Scheme 4.12 Meerwein-Pondorf-Verley reduction of the epoxide **139.**

Initially this analogue was prepared from the previously prepared **128** by protecting the secondary 10-hydroxy group as the acetate **166** prior to exposing the primary allylic alcohol **167** for the diphosphorylation step, the acetate group was then removed after ion exchange and HPLC to yield **162**. The deprotection of the acetate was carried in D_2O to allow the reaction to be monitored by NMR spectroscopy. A second look at the method highlighted the difference in reactivities of the two-alcohols and so an improved synthesis took advantage of this to shorten the synthesis and remove the necessity of orthogonal protecting group use. Use of farnesyl acetate (**141**) yielded **139** in

two steps; treatment of the bromohydrin **170** with K_2CO_3 in methanol resulted in both the closure of the epoxide and saponification of the acetate group. After Meerwein-Pondorf-Verley reduction of the epoxide the primary alcohol of **163** was selectively activated with 1-equivalent of tosyl chloride and the diphosphate generated from the tosylate (**164**). It was also possible to selectively diphosphorylate the primary alcohol directly using the trichloroacetonitrile coupling.[124]



Scheme 4.13 Syntheses of **162**. *Reagents and conditions: i. Al(O^{i}Pr)$_3$, toluene, Δ, 65 %; ii. Ac$_2$O, Py; iii. PPTS, EtOH, 65 % over two steps; iv. s-collidine, MsCl, LiCl, 0 °C; v. (Bu$_4$N)$_3$HP$_2$O$_7$, MeCN then DOWEX ion exchange, HPLC, 56 %; vi. NH$_4$OD, D$_2$O: vii. NBS, THF/H$_2$O, 65 %; viii. K$_2$CO$_3$, MeOH, 92 %; ix. Et$_3$N, pTsCl.*

With the exception of EBFS, incubation of 162 with the range of sesquiterpene synthases gave the same product as with 10,11-epoxy-FDP (147) to varying extents,

Table 4.2, Figure 4.7. Interestingly, incubation with AS once again produces a second product with a longer retention time. This second product is different to that produced by AS with **147**, Figure 4.8.

Table 4.2 Relative efficiencies of sesquiterpene synthases with 10-hydroxy-11-ene-FDP (**162**) and 10,11-epoxy-FDP (**147**) based on TIC. nas*, not a substrate.

Enzyme	10-hydroxy-11-ene-FDP	10,11-epoxy-FDP
EBFS	nas*	Good
GdolS	Good	Good
AS	Good (2 products)	Good (2 products)
DCS	Poor	Average
GAS	Poor	Average
GDS	Average	nas*
ADS	Average	V.poor

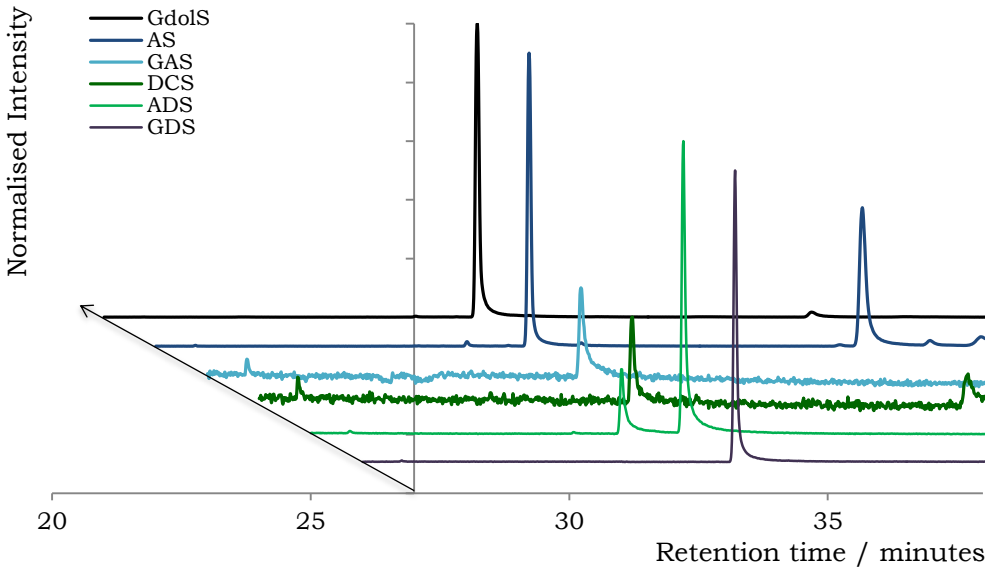


Figure 4.7 Overlaid normalised gas chromatogram of the pentane extractable products arising from incubation of **162** with all enzymes

In order to definitively identify the product generated by all enzymes upon incubation with **147** and **162**, a preparative-scale incubation of GdolS with **162** was used to generate sufficient compound for NMR spectroscopic analysis; generating 12.2 mg of compound which was identified as being the 11-membered cyclic ether, **158**, in 41.3 % yield. The resulting compound was analysed by GC-FID using a chiral column. While the epoxide **139**, and therefore the resultant diol **163**, was prepared in a non-stereoselective manner we wanted to analyse whether there was any preference in stereochemistry for the ring closure. Integration of the resolved peaks shows a ratio of 48:52, indicating that there is no stereochemical preference for this ring closure. This might be due to the relative proximity of the attacking hydroxyl group to the highly reactive farnesyl cation, **171**, in either stereochemistry, the carbocation is rapidly quenched with very little involvement of the enzyme other than the

folding of the substrate in the active site. Incubation of the analogue with the acetate-protected secondary alcohol gave no pentane-extractable no products with any enzymes.

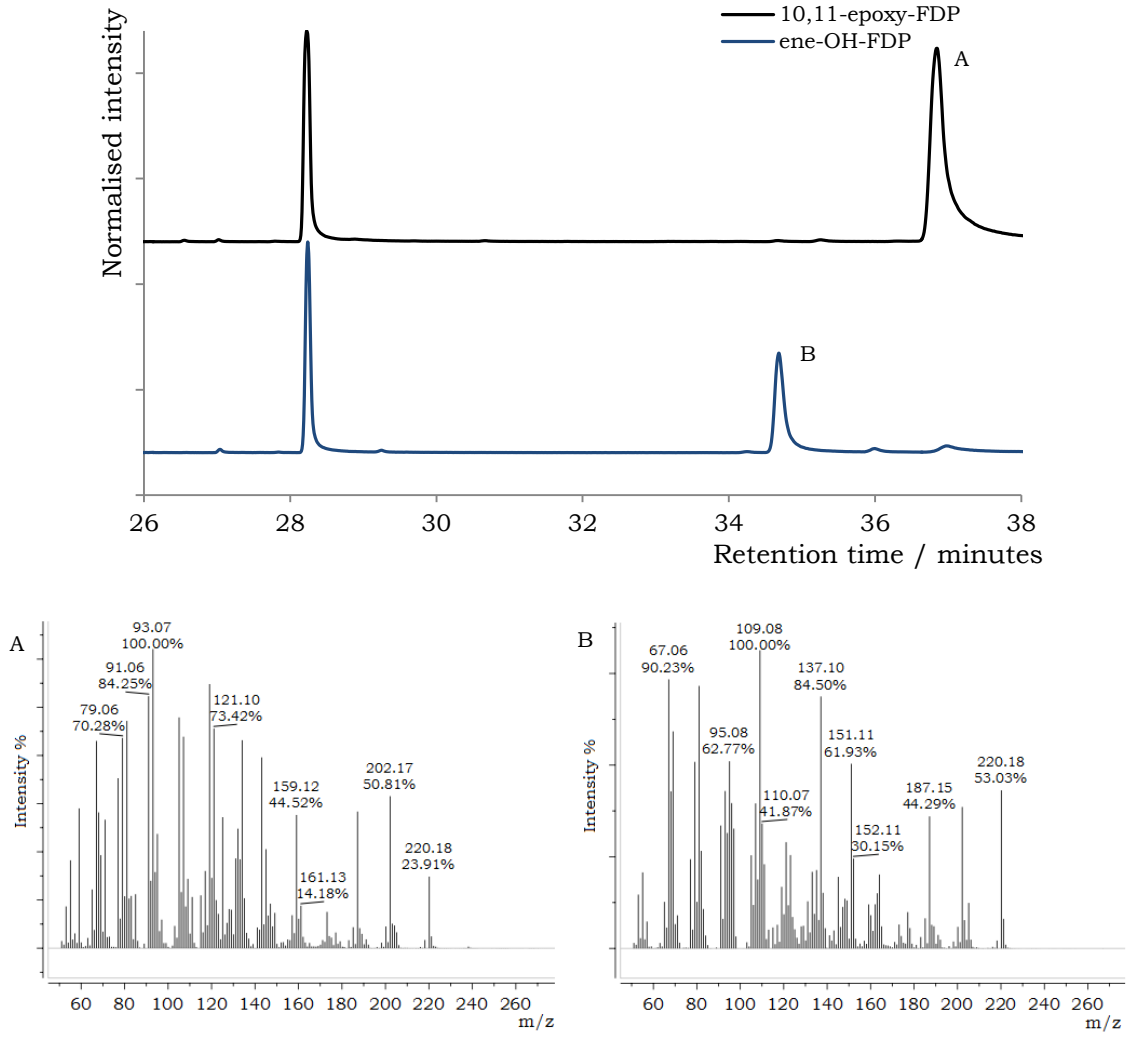
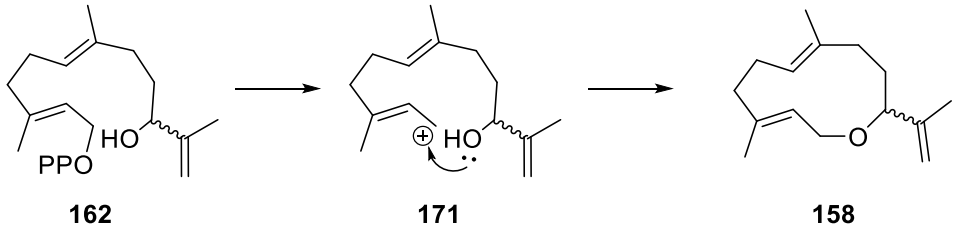


Figure 4.8 Overlaid normalised gas chromatogram of the pentane extractable products arising from incubation of 10,11-epoxy-FDP (**147**) and 10-hydroxy-11-ene-FDP (**162**) with AS (top) and mass spectra of the labelled compounds (below).



162 **171** **158**

Scheme 4.14 Cyclisation of **162** by sequiterpene synthases

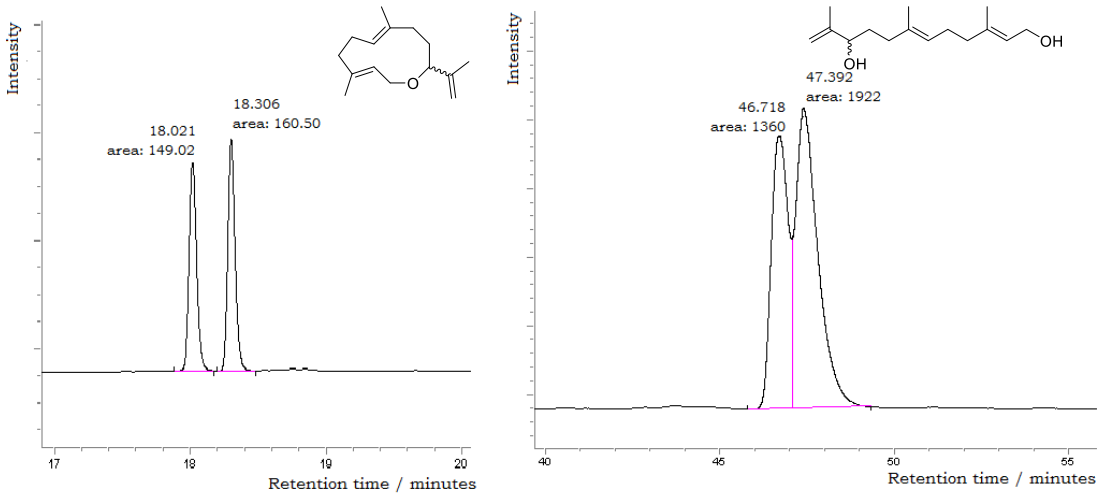


Figure 4.9 GC-FID chromatograms of macrocyclic ether **158** (Left) and diol **163** (Right)

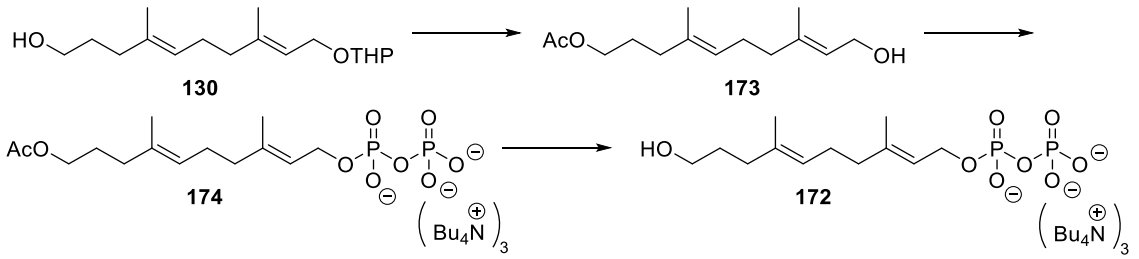
4.2.4 Third generation analogue

Following on from the success of **162** to directly quench the farnesyl cation we decided to extend this principle and investigate the minimum substrate recognition motif necessary for sesquiterpene synthases for cyclisation of the substrate. The isopropyllidine tail was envisioned as being removed from **162** to create a 'truncated' 12-carbon-hydroxy-diphosphate with a 10-carbon chain, **172**, to test if the enzymes are able to accept this as a substrate and still able to fold the full length of the analogue into a cyclisation ready configuration in the active site, Figure 4.10.

162 **172**

Figure 4.10 Rational design of truncated FDP analogue by removal of isopropylidine tail (red)

The synthesis of this analogue used an intermediate from the synthesis of the 12,13-F_2-FDP (**73**), the alcohol **130**. The free terminal alcohol was protected as an acetate and the tetrahydropyran ether removed to expose the allylic alcohol, **173**, for diphosphorylation. The diphosphate was prepared using the Keller method to generate the tetra-n-butyl ammonium salt with the acetate intact, **174**. A sample of the terminal alcohol was prepared by the removal of the acetate by stirring in NH_4OD/D_2O for 12 hours as previously described, and the reaction progress monitored by NMR spectroscopy. The resulting tetra-n-butyl ammonium salts were used without further ion exchange.



Scheme 4.15 Synthesis of 172. *Reagents and conditions: i. Ac_2O, Py; ii. PPTS, EtOH, 61% over 2 steps; iii. Cl_3CCN, $(Bu_4N)H_2PO_4$, MeCN, 31%; iv. NH_4OD, D_2O, 96%.*

GC-MS analysis of the pentane-extractable products after overnight incubation of the 10-*OH*-DP (**172**) with all sesquiterpene synthases showed that it was a substrate for all cyclases and the same product was generated, eluting at 23.1 minutes Figure 4.11. Compound **172** was not a substrate for EBFS. As judged by TIC, it was a poor substrate for ADS, GAS and GDS; while for the other enzymes it was a good substrate.

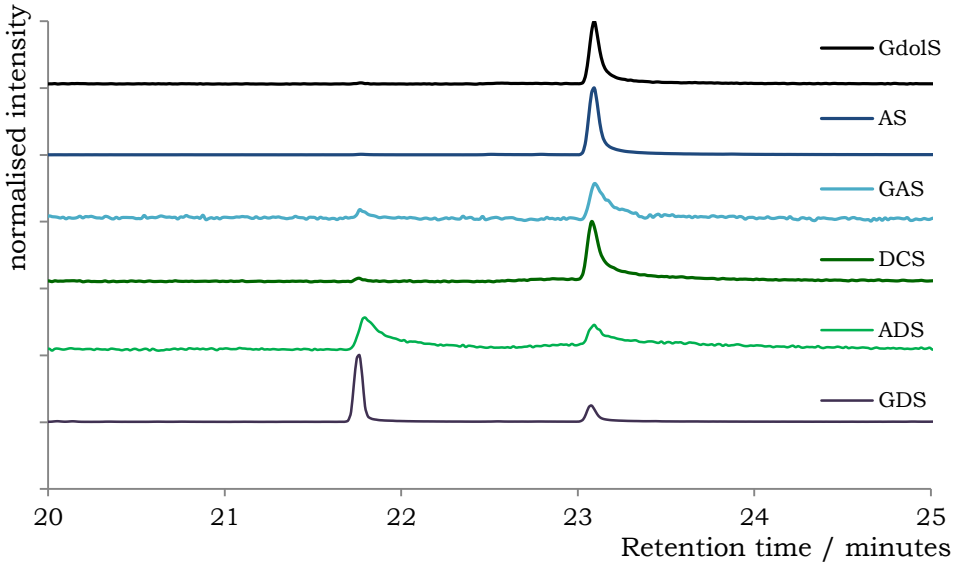


Figure 4.11 Overlaid normalised chromatogram of the pentane extractable products arising from incubation of 10-OH-decadiene-DP **172** with enzymes

The mass spectrum of the product formed by all cyclases with **172** (retention time 23.1 minutes) possessed a molecular ion with m/z = 180, and the fragmentation pattern also shows loss of progressive CH_2 units, Figure 4.12 B. This product may be the farnesene-type product **175** formed by deprotonation, or the macrocyclic ether **176**, formed by quenching of the allylic cation, **177**. Alternatively, another possibility would be the ionisation of the diphosphate with rearrangement to the tertiary diphosphate, **178**, followed by 1,6-ring closure and deprotonation to yield a bisabolyl-type product, **179**. While the formation of the bisabolyl-type product may be less likely due to the lack of

any evidence of aristolochene synthase following a mechanism through NDP, at present none of these products can be ruled out with the available data.

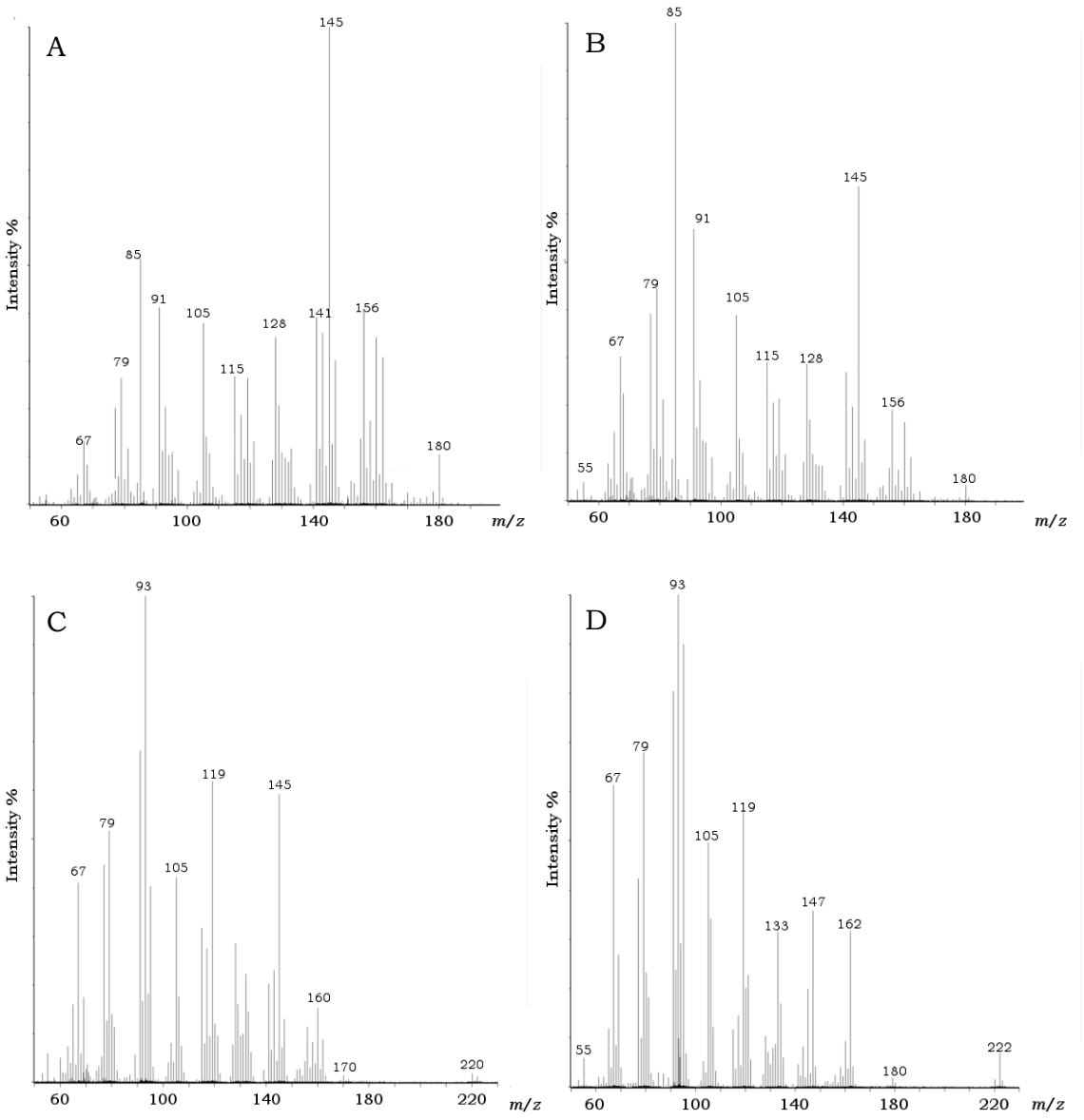
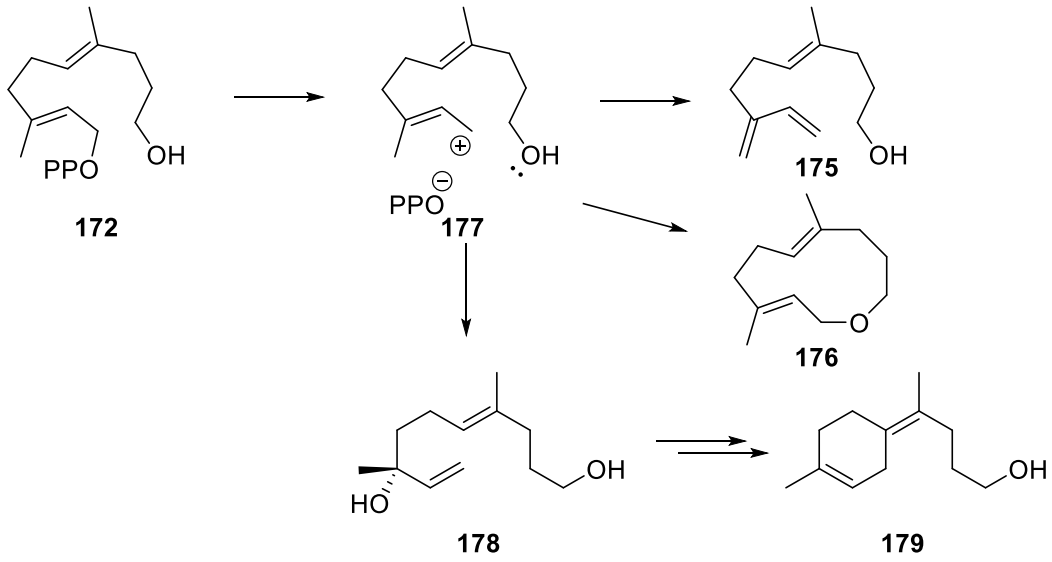


Figure 4.12 Mass spectra of the products arising from incubation of **172** with GDS (A) and AS (B), and incubation of **174** with AS(C) and EBFS (D).

Incubation with GDS, ADS and GDS also produces a second compound which is the major product formed by GDS, eluting at 21.8 minutes Figure 4.11. The mass spectrum of this product possessed a molecular ion with m/z = 180, and the fragmentation pattern shows loss of progressive CH_2 units, Figure 4.12 A. However, the fragment of m/z = 145 is the most abundant for this product, compared to the mass spectrum for product A, in which the fragment with m/z = 93 is the most stable fragment.



Scheme 4.16 Conversion of **172** with sesquiterpene synthases and possible products.

The 10-OAc-DP, **174**, was only a substrate for PR-AS and EBFS, while the products have almost identical retention times, Figure 4.13, analysis of the mass spectra clearly shows that different compounds are produced,Figure 4.12 C and D. The compound produced by EBFS possesses a molecular ion with *m/z* = 222, congruent with loss of the diphosphate to form a linear or monocyclic product. The mass spectrum produced no consistent hits with the NIST spectral library, but the fragmentation pattern is congruent with a 'farnesene'-type product. That EBFS is able to turn over the longer 10-*OAc*-DP **174**, but not the 10-*OH*-DP **172** is of interest, possibly demonstrating that in EBFS the longer chain is needed to correctly orientate the molecule in the active site for turnover.

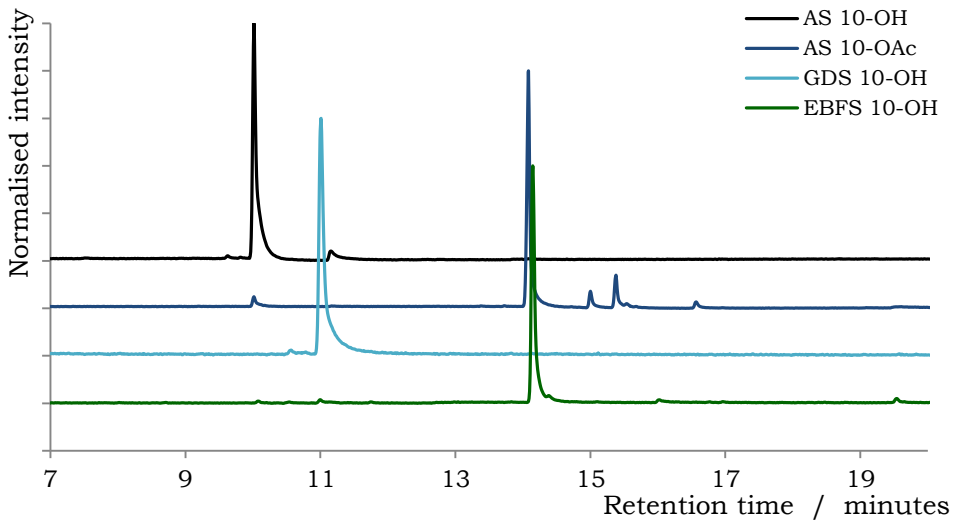
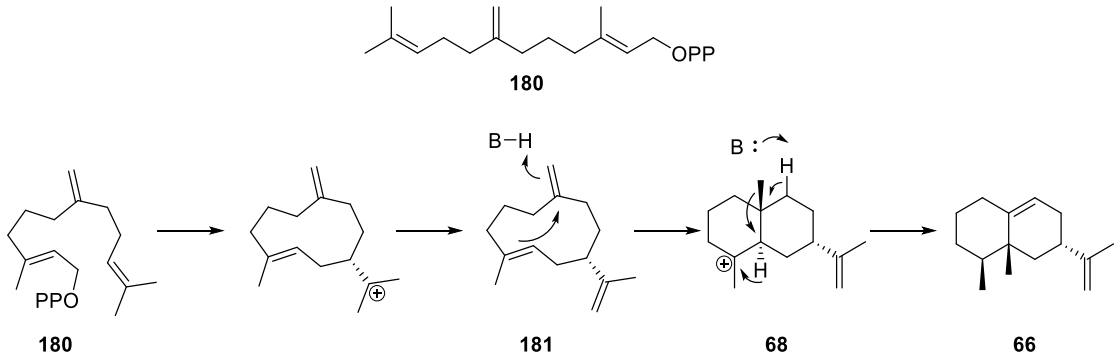


Figure 4.13 Overlaid normalised gas chromatogram of the pentane extracted products from incubation of EBFS and PR-AS with **172** and **174**

4.2.5 7-methylene-FDP

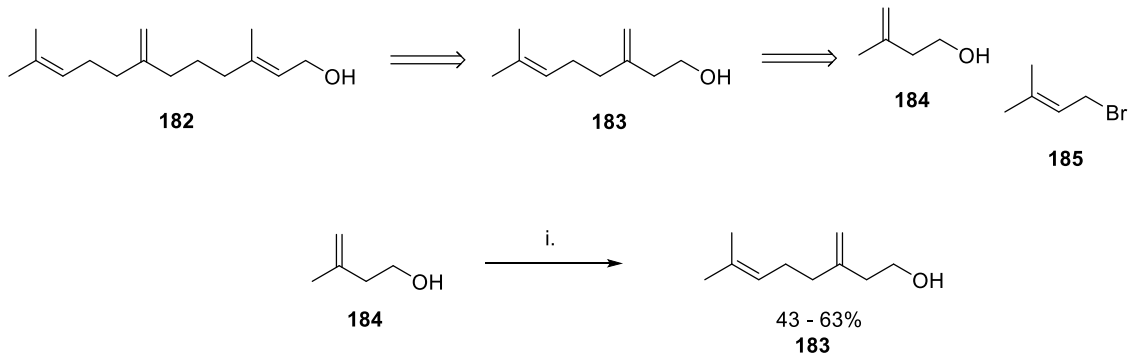
Aristolochene synthase was found to be the one of the most versatile of the enzymes tested, being the only cyclase able to turn over the 10-*OAc*-DP, **174**, and the more specific of the two cyclases (PR-AS and GdolS) able to turn over the 6,7-epoxy-FDP, **146**. The presence of the second products in the incubations with the 10,11-epoxy-FDP, **147**, and the 10-hydroxy-11-ene-FDP, **162** indicates that in addition to the simple quenching of the farnesyl cation to produce the cyclic ether **158**, AS is able to perform another function generating the second products in almost equal proportion.

To further probe the versatility of PR-AS and the flexibility of the active site, an FDP-analogue was designed whereby the 6,7-double bond was shifted to the 7,14-position to create the 7-methylene-FDP, **180** Scheme 4.17. It was postulated that this alteration would have no effect on the initial cyclisation, to yield the intermediate germacrene A analogue, **181**; if the second ring-closure was then able to proceed and protonation to occur at the 7-methylene to yield the eudesmane cation **68**, the carbocation cascade could continue; producing aristolochene **66**.



Scheme 4.17 Structure of 7-methylene-FDP (**180,** top) and proposed cyclisation of **180** to **66** by PR-AS (bottom)

The 7-methylene-farnesol (**182**) was prepared using a combination of established procedures.[142–145] γ-Geraniol (**183**, 7-methyl-3-methyleneoct-6-en-1-ol) was prepared via the dilithiation of 3-methylbut-3-en-1-ol (**184**) and trapping with prenyl bromide (**185**). While the reported yield of 84% was never achieved, extending the period over which **185** was added, from 15 minutes to 50 minutes, improved the yield significantly, from 43 to 63%.[142]

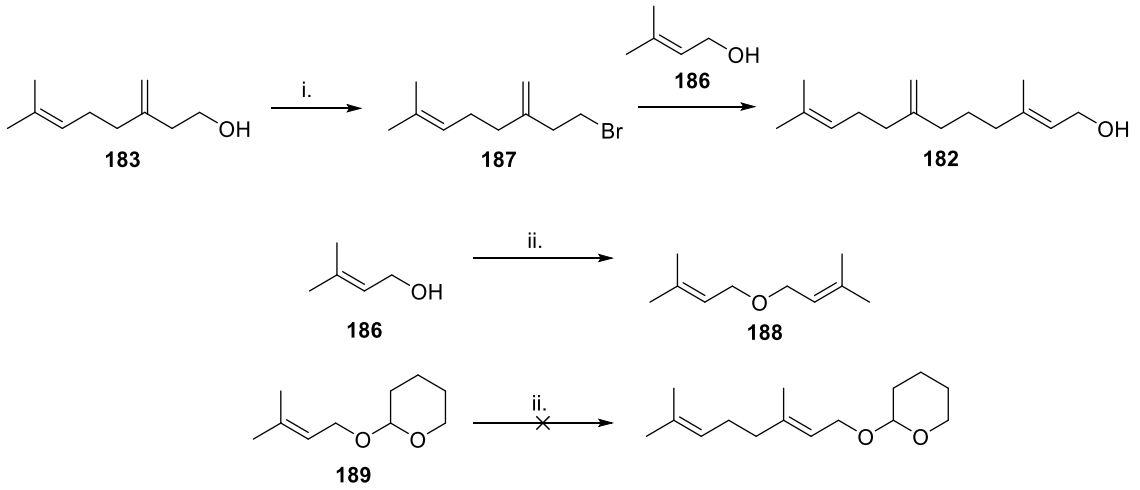


Scheme 4.18 Retrosynthesis of 7-methylene-farnesol **182** (top) and synthesis of γ-geraniol **183** (bottom) *Reagents and conditions: i. TMEDA, n-BuLi, THF, 0 °C, then **184**, 0 °C to RT, 6 h followed by **185** -78 °C to RT, 16 h.*

Due to the relative simplicity of the reaction for the extension of the chain it was proposed that the γ-geranyl chain could be extended to **182** in a similar manner by dilithiation of prenol (**186**) and the γ-geranyl bromide (**187**). This reaction was first tested using **186** and **185**, yielding only the diisoprenyl ether (**188**, in good yield, proposed to arise from formation of the singly lithiatiated alkoxide species, Scheme 4.19. In order to counter this, the alcohol was protected as the THP-ether (**189**) and the reaction repeated, this time yielding only unreacted starting material (**189**).

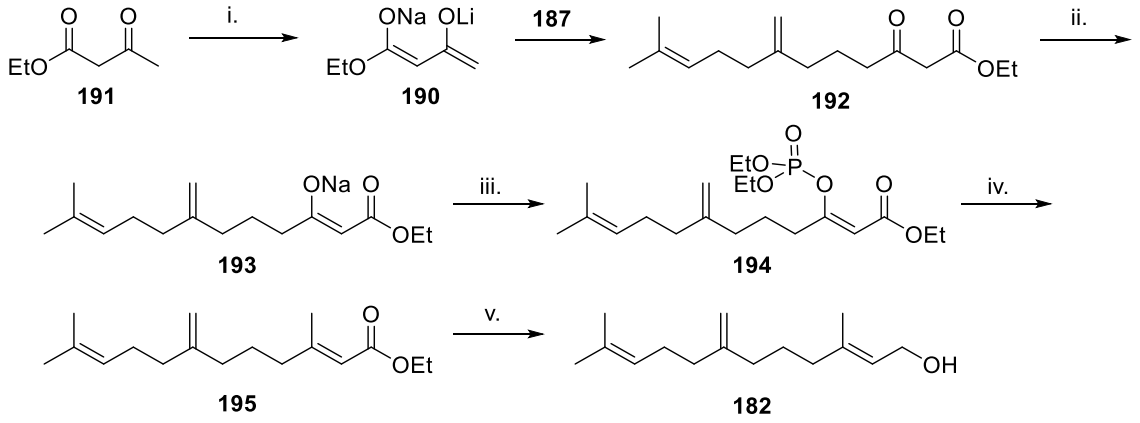
After numerous failed attempts it was decided to use the reliable and established, although longer, method of chain extension using the acetoacetate dianion 190**,**

Scheme 4.20.[144,145] The acetoacetate dianion **190**, was formed by successive deprotonations of ethyl acetoacetate (**191**) with sodium hydride and n-butyl lithium; alkylation with **187** then forms β-keto ester **192**. Deprotonation of **192** with sodium hydride formed enolate **193**, which was trapped as the enol phosphate (**194**) with diethylchlorophosphate and then used without purification for the methylation with lithium dimethylcuprate to yield ester **195**. Reduction with DIBAL-H yields the allylic alcohol (**182**).[143] The farnesol was then diphosphorylated using the Keller procedure in 45 % yield.[124]



Scheme 4.19 Proposed elongation of γ-geraniol **183** (above) Failed elongation tests (below). *Reagents and conditions: i. Et_3N, MsCl, -45 °C then LiBr, RT, 72 %; ii. TMEDA, n-BuLi, THF, 0 °C, then **186**, 0 °C to RT, 6 h followed by **185** -78 °C to RT, 16 h.*

While initially designed to probe the versatility and active site flexibility of AS, this analogue was also incubated with the other available sesquiterpene synthases and the pentane-extractable products analysed by GC-MS.



Scheme 4.20 Elongation of γ-geranyl bromide to 7-methylene-farnesol. *Reagents and conditions: i.NaH, THF, 0 °C, then n-BuLi, followed by γ-geranyl bromide, 65%; ii. NaH, THF, 0 °C, then diethylchlorophosphate; iv. Me_2Cu, -78 °C, then addition of enolphosphate, -78 °C for 2 h, -45 °C for 2 h, followed by addition of MeI, -45 °C, 0.5 h, 30 % over two steps; v. DIBAL-H, toluene, -78 °C, 85 %.*

Incubation of the 7-methylene-FDP (**180**) with AS shows that, as proposed, AS is able to turn over the analogue to generate aristolochene (**66**) as the major product (A, 26.01 minutes, 37 %), along with a number of other products; one other major product (B, 26.88 minutes, 31 %) and four minor products (27 % total).

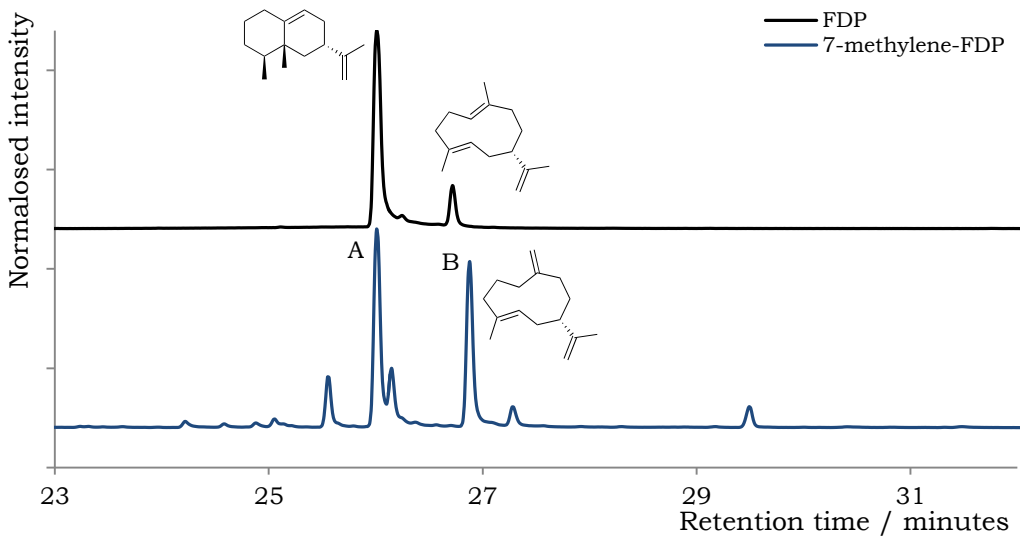


Figure 4.14 Normalised chromatogram of the pentane extractable products arising from incubation of FDP and 7-methylene-FDP with AS

Compound **180** is also a substrate for all the other sesquiterpene synthases used in this study. With the exception of EBFS, all of the enzymes generate product B (retention time, 26.88 minutes) in varying proportions along with other products. Incubation of **180** with EBFS generated 3 products, and analysis of the mass spectra is consistent with production of the farnesenes, Figure 4.15. Incubation with GAS generated almost exclusively product B (>90 %); the mass-spectrum possesses a molecular ion with m/z = 204, while the fragmentation pattern is near identical to germacrene A suggesting that this product is the 7-methylenegermacrene A (**181**), Figure 4.16.

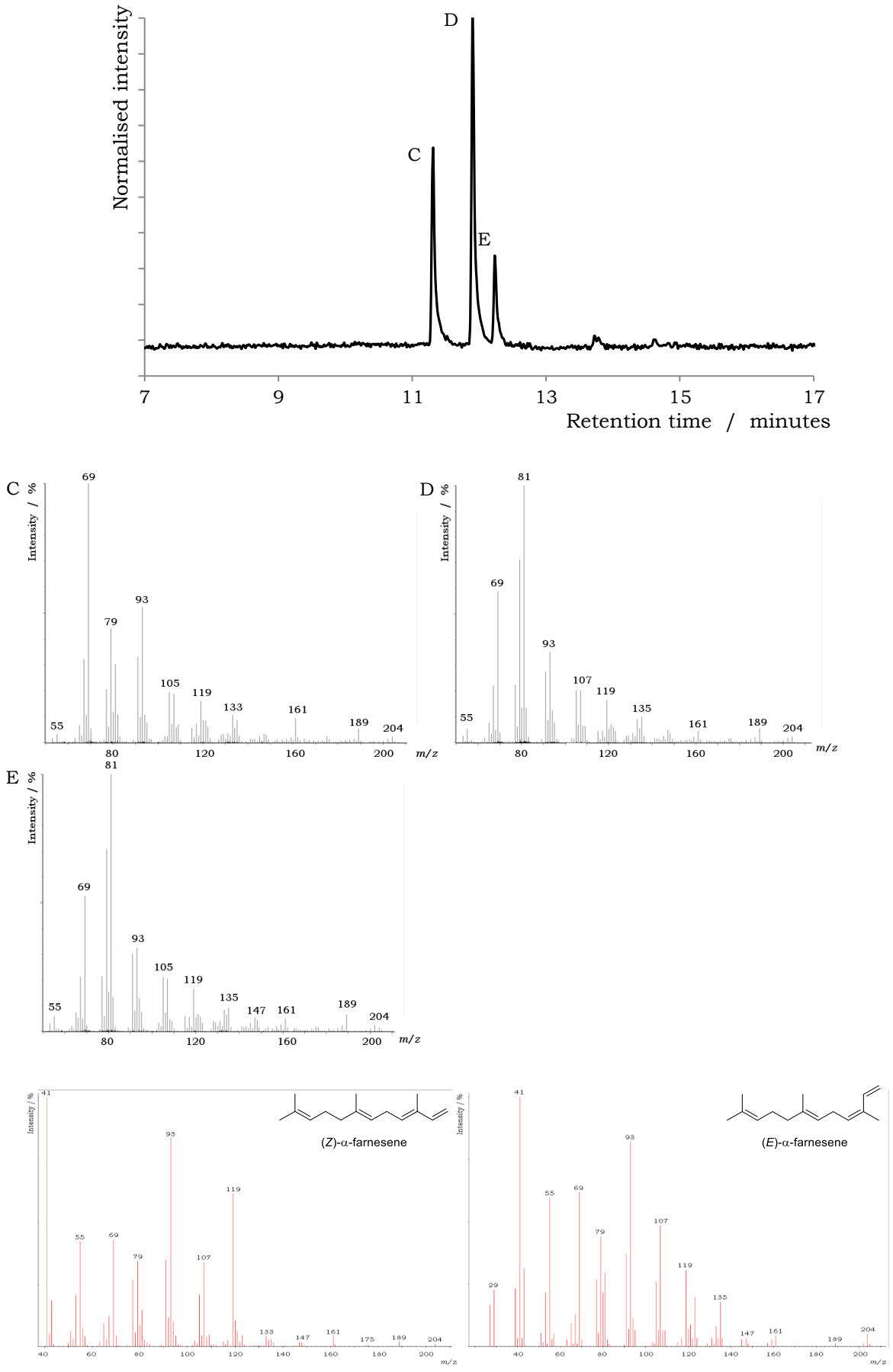


Figure 4.15 GC-MS analysis of the pentane extractable products arising from incubation of **180** with EBFS. Top, gas chromatogram; middle, mass spectra of compounds labelled corresponding to product; Bottom, mass spectra of (*Z*)-α-farnesene and (*E*)-α-farnesene.

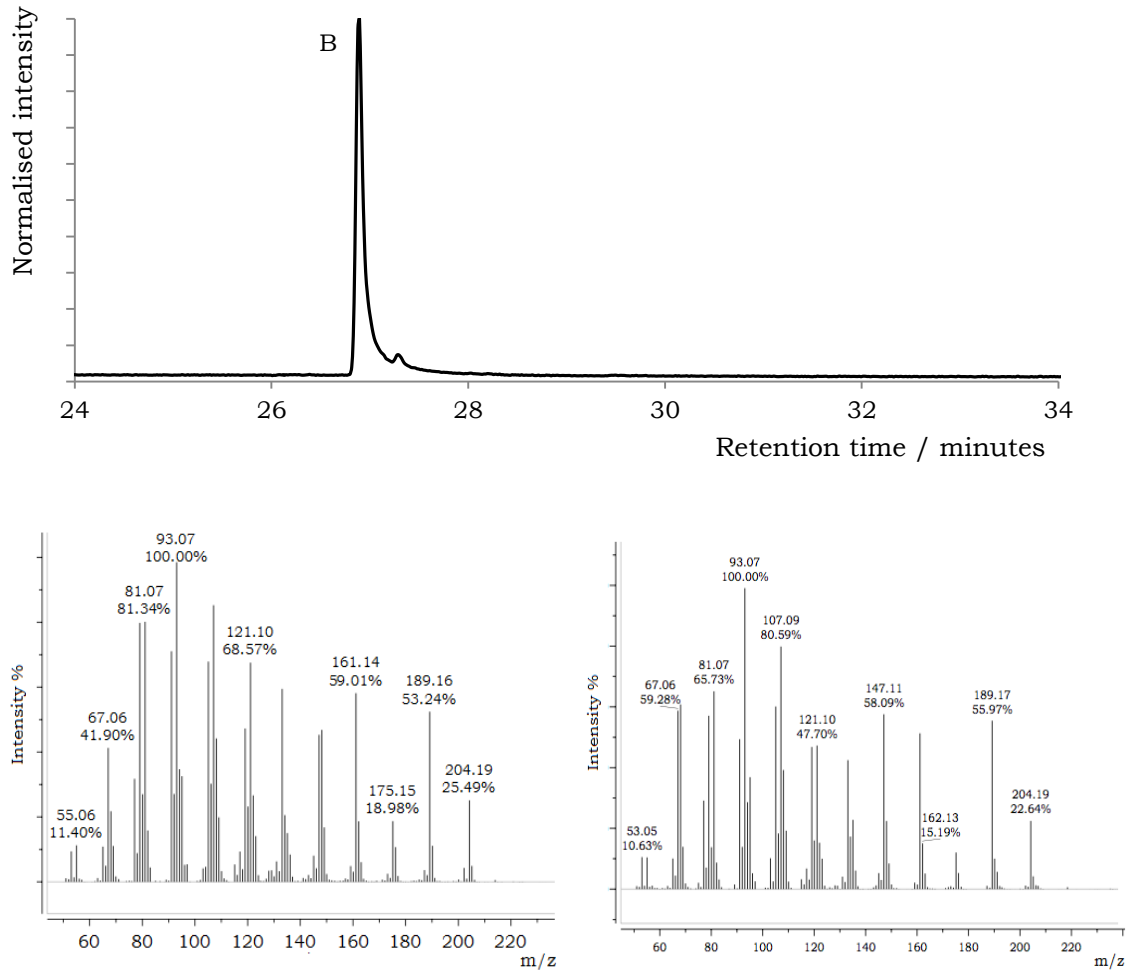
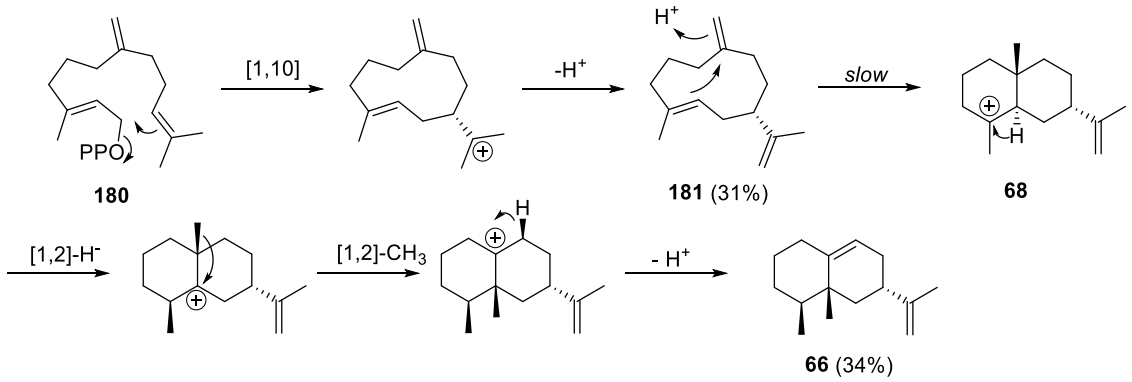


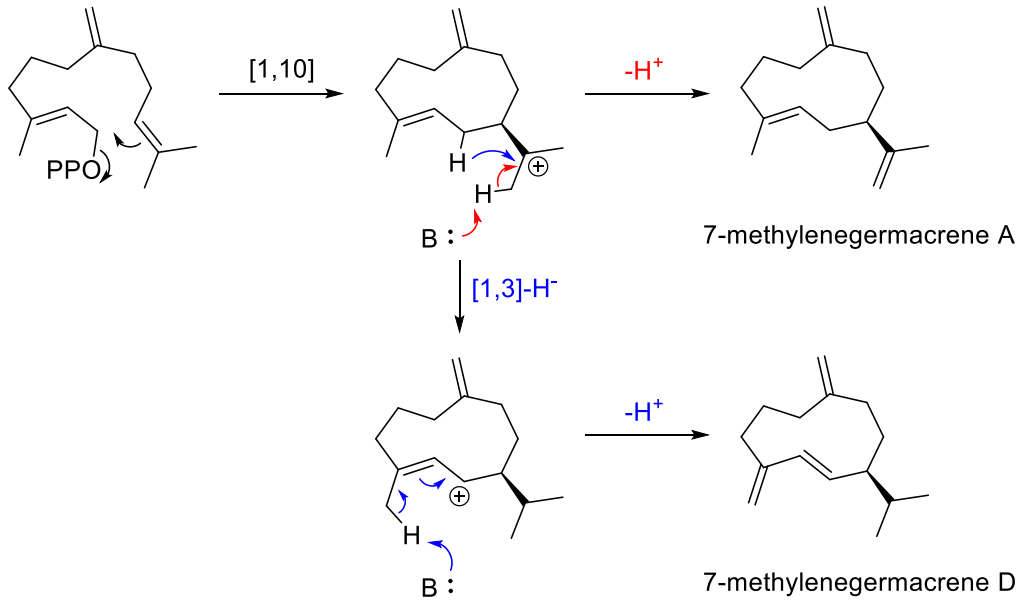
Figure 4.16 Normalised gas chromatogram of the pentane extractable products arising from incubation of **180** with GAS (top) and mass spectra of the product at 26.88 minutes (bottom right) and germacrene A **34** (bottom left)

This would be consistent with the observation that this compound is produced as the second major product by AS; naturally AS produces some germacrene A (**34** 7.5 %) from incubation with FDP (**19**), so it could be concluded that incubation with **180** would also produce some **181**. Due to the less productive positioning of the double bond in **181**, protonation of the double bond to drive the second ring closure and formation of the eudesmane cation **68** is slower leading to greater proportions of **181** released from the active site, Scheme 4.21.



Scheme 4.21 Cyclisation of 7-methylene-FDP by PR-AS

In addition to compound B (**181**), the major product from incubation with GDS appears to be a cyclic product from analysis of the mass-spectrum, with a molecular ion with m/z = 204. The fragment at m/z = 161 is 100% intensity, indicating loss of the isopropyl group, this might suggest that GDS is able to turn over this product to a germacrene-D analogue in addition to **181** depending on the deprotonation step, Scheme 4.22.



Scheme 4.22 Cyclisation of 7-methylene-FDP by GDS

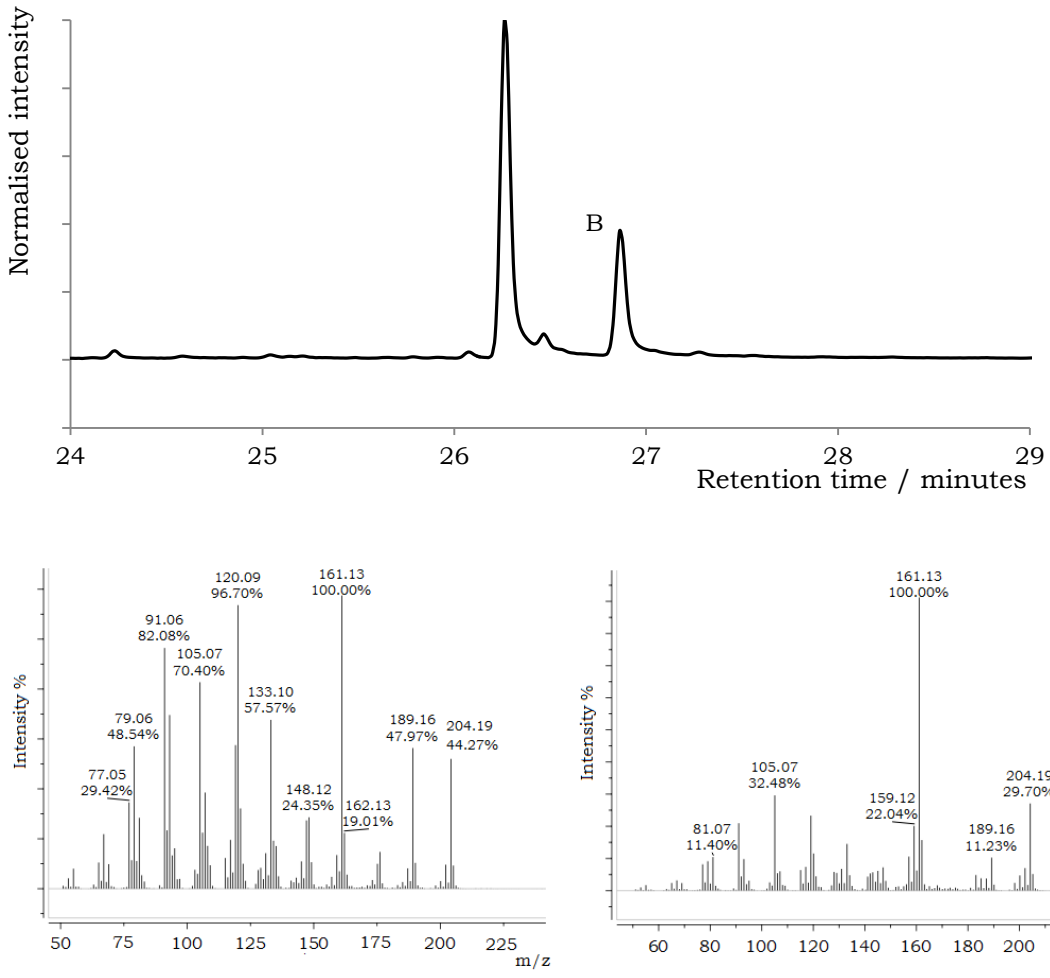


Figure 4.17 Normalised gas chromatogram of the pentane extractable products arising from incubation of 7-methylene-FDP with GDS (top) and mass spectrum of the major product (bottom left) and of germacrene D (bottom right)

GdolS is able to turn over **180** to generate a number of products. In addition to **181** already discussed (B, retention time 26.88 minutes, 24 %), the major product has a retention time of 30.09 minutes (43 %) and the mass spectrum shows a molecular ion with m/z = 222. The two minor products with retention times of 31-32 minutes (9 %) also show molecular ions with m/z = 222 and the mass spectra of these three peaks (30.09, 31.01 and 31.61 minutes) are all very similar, and similar to that obtained for germacradien-4-ol (**79**). This suggests that GdolS is able to generate the germacradien-4-ol analogue from **180**. Incubation with this enzyme also generates 4 other minor products (24 %) each possessing molecular ions with m/z = 204.

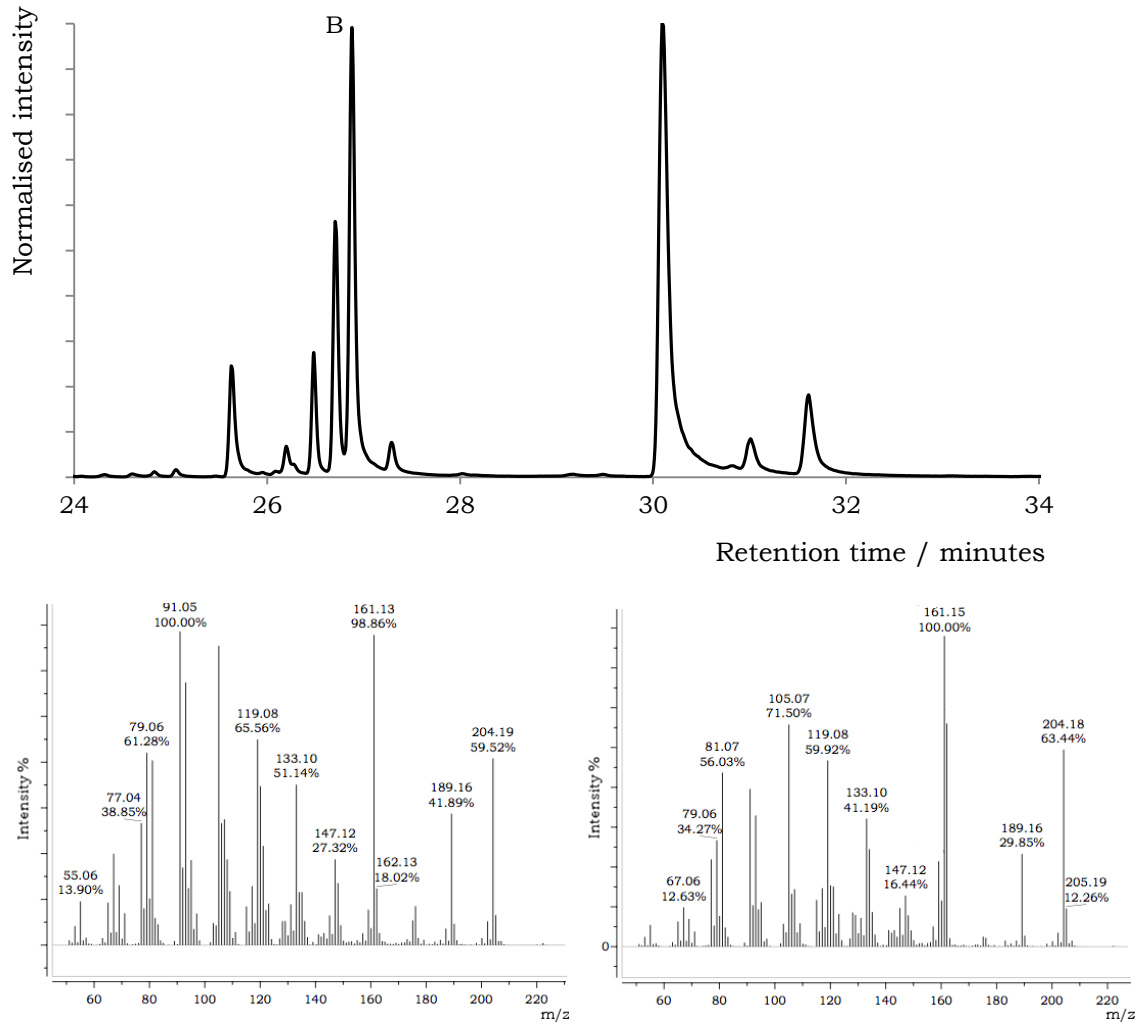


Figure 4.18 Normalised gas chromatogram of pentane extractable products arising from incubation of **180** with GdolS (Top) and mass spectra of the product at 30.1 minutes (bottom left) and of germacradien-4-ol, **79** (bottom right).

Incubation with DCS produces **181** (retention time 26.88 minutes, 70 %) and the same terpene alcohol as with GdolS (retention time 30.1 minutes, 21 %) in addition to minor traces of other products (4 %), Figure 4.19.

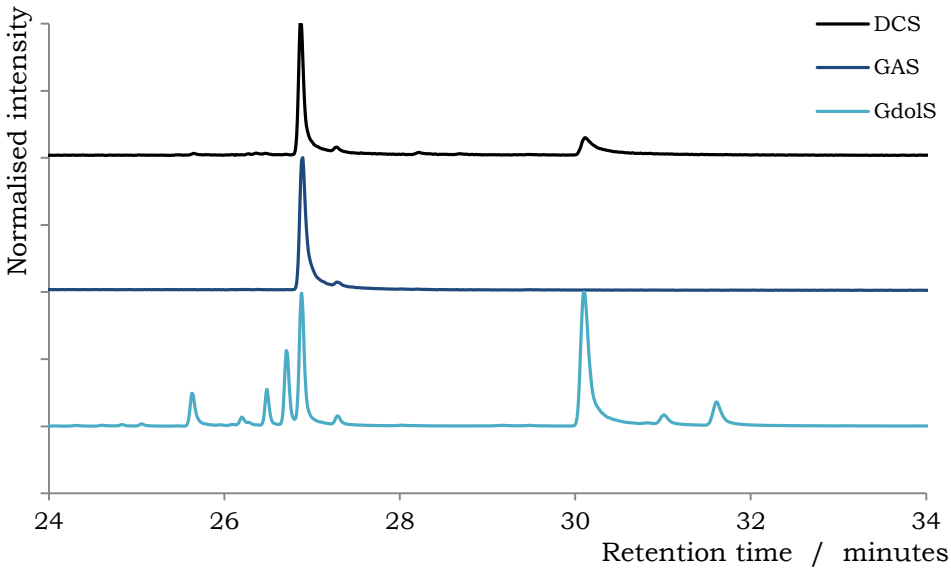


Figure 4.19 Normalised overlaid gas chromatogram of pentane extractable products arising from incubation of **180** with DCS, GAS and GdolS

While incubation with ADS does produce **181**, this is a minor product in this incubation (~5 %) with the major product having a retention time of 26.07 minutes (51 %) and 3 other minor products (39 %). It is not possible for ADS to generate amorpha-4,11-diene from **181** due to the necessity of the 6,7-double bond in the initial 1,6-ring closure. However, if the isomerisation to 7-methylene-NDP, **196**, were to occur followed by 1,10-ring closure, the newly formed double bond in the 10-membered macrocycle would have *(Z)*-stereochemistry, yielding the 7-methylene derivative of helminthogermacrene A, **197**, Scheme 4.23.

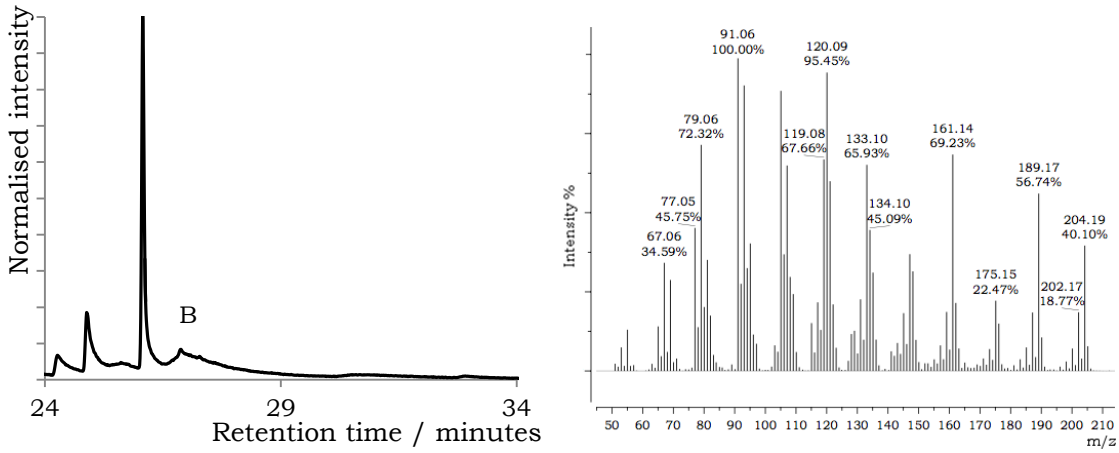
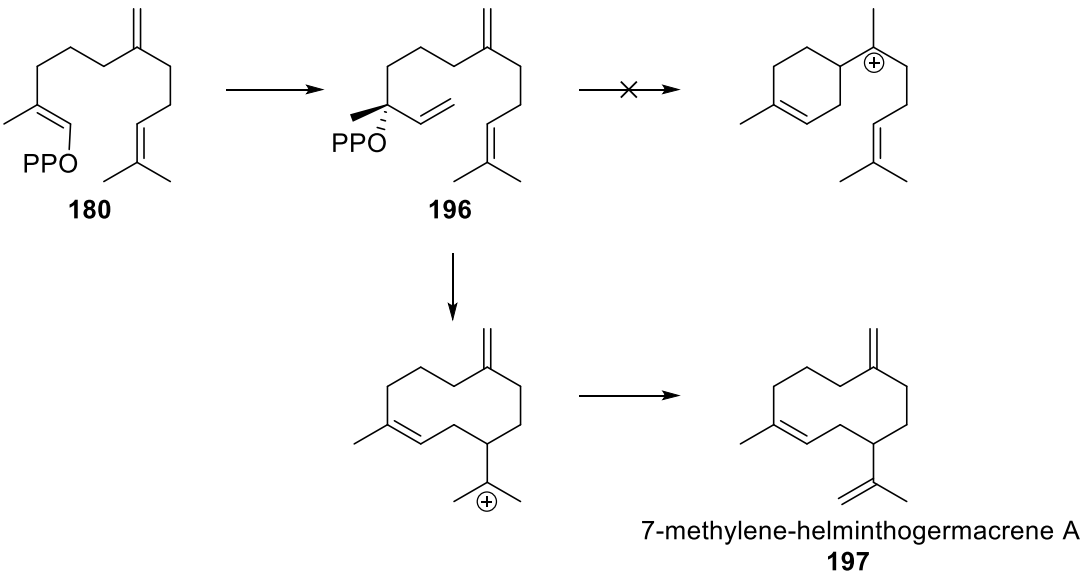


Figure 4.20 Normalised gas chromatogram of the pentane extractable products arising from incubation of **180** with ADS (left) and mass spectrum of the major product (right)

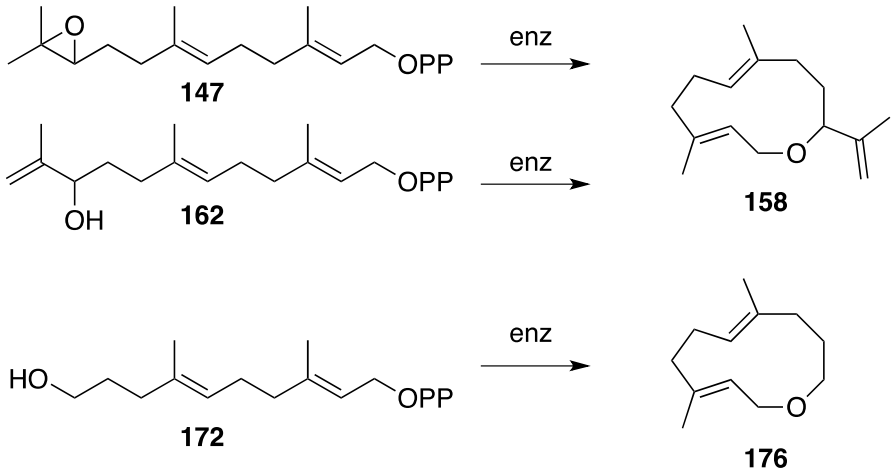


Scheme 4.23 Cyclisation of **180** *via* 7-methylene-NDP,(**196**)

4.3 Summary

A number of functionalised-FDP analogues have been synthesised and incubated with a range of sesquiterpene synthases to investigate the scope of these enzymes; how they are able to turnover unnatural substrates and if they are able to generate novel sesquiterpenoids.

The 10,11-epoxy-FDP **147**, was found to be turned over by a number of the enzymes and the products identified as a 11-membered macrocyclic ether **158**, proposed to result from the direct quenching of the farnesyl cation followed by a deprotonation. A second analogue was designed and synthesised to test this theory, the 10-hydroxy-11-ene-FDP **162**, which could also directly trap the farnesyl cation, but without the need for further deprotonation, pleasingly this analogue generated the same macrocyclic ether **158**. Due to the enforced folding of the substrate in the active site both enantiomers of the hydroxyl group are in close proximity to the generated farnesyl cation leading to no stereochemical preference in the generation of the ether. The generation of **158** from both **147** and **162** demonstrate the first method in the literature of trapping the proposed farnesyl cation to date. Extending this principle, an analogue without the isopropyl tail **172**, was generated and demonstrated that this analogue was still turned over by all enzymes, producing what was proposed to be the cyclic ether **176**, due to the high reactivity of the allylic cation and the confinement of the active site.



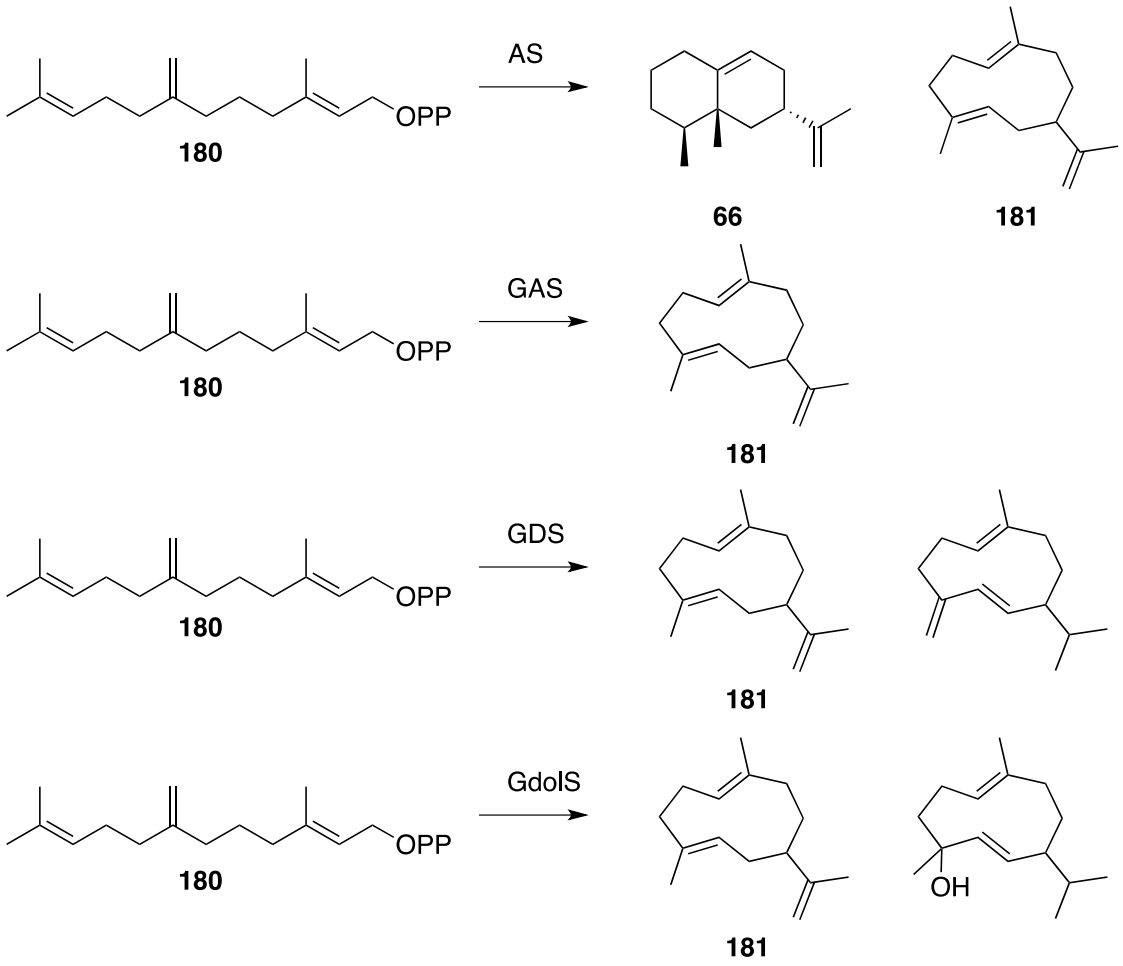
Scheme 4.24 Key products from incubations of sesquiterpene synthases with FDP-analogues

While modifications to the nature of the farnesyl chain were reasonably well tolerated around the distal double bond, with cyclisation's efficiently producing specific products, epoxidation to the C6,C7 double bond was less successful with only two of the tested cyclases able to use **146** as a substrate (PR-AS and GdolS), neither of which produced a single overriding product. Epoxidation of the 2,3-double bond produced an FDP analogue, **145**, that was not a substrate for any of the enzymes tested, and the analogue with epoxides in place of both the C6,C7- and C10,C11-double bonds, **148**, was only turned over by EBFS to generate farnesene analogues.

The 7-methylene-FDP, **180**, transposed the trans 6,7-double bond out of the main chain, to investigate the versatility of PR-AS and the ability of sequiterpene synthases to accept more minor changes to the centre of the farnesyl chain. This analogue was accepted by all the tested enzymes to varying extents and in the case of PR-AS, was turned over to the natural product **66**, albeit with lower specificity than the natural substrate. All the other enzymes tested are proposed to be able to turnover this substrate into monocyclic products.

These results indicate a level of flexibility in the purely hydrocarbon half of FDP (C7 – C10), which enables many sesquiterpene synthases to accept analogues of FDP with modifications in this region. While the 6,7-epoxy-FDP **146**, was turned over by PR-AS and GdolS, the effect of changing configuration and imposing stereochemistry at two centres proved too much for the other enzymes, and even in the cases of PR-AS and GdolS turnover was lower, and these normally high-fidelity enzymes produced numerous products. The more minor change in the 7-methylene-FDP **180**, affecting only the single centre on

the farnesyl chain, was well accepted by all enzymes and turned over to primarily cyclic products.



Scheme 4.25 Summary of products generated by incubation of 180 with sesquiterpene synthases

The generation of **147** and **162** were turned over in varying degrees, but produced the same macrocyclic ether in all cases, demonstrating the first reported trapping of the farnesyl cation. With the lack of stereochemical bias in the cyclisation of the 10-hydroxy-11-ene-FDP demonstrating a degree of freedom in the isopropyl tail and pointing to this region of FDP as an ideal platform for further modifications.

5 Conclusions

This thesis investigated means of generating simple terpenoids using terpene synthases. Two distinct and complementary approaches were employed. These involved a mechanistic investigation of germacradien-4-ol synthase, a sesquiterpene synthase which incorporates water in the cyclisation cascade to generate a single terpene alcohol, and the synthesis of a range of FDP-analogues designed to offer means of trapping carbocation intermediates in the catalysis by a range of sesquiterpene synthases.

5.1 Mechanistic investigation of germacradien-4-ol synthase

Fluorinated FDP analogues 2-F-FDP and 15-CF_3-FDP proved to be potent inhibitors of GdolS, suggesting a stepwise mechanism for the cyclisation to the germacryl cation through the farnesyl cation. The intermediacy of the farnesyl cation was further evidenced by the turnover of the 12,13-difluoro-FDP to the 12,13-difluorofarnesene. While GdolS also turned-over the 10-F-FDP to the 10-F-farnesene, due to destabilisation of the germacryl cation, it was also able to direct a 1,11-cyclisation to form the fluorinated-α-humulene derivative.

Incubation of GdolS with 1-$^{2}H_2$-FDP demonstrated that the enzyme catalyses a 1,3-hydride shift subsequent to the ring closure, and the stereochemistry of this shift was investigated using 1-(*R*-$^{2}H_1$)-FDP. Suprisingly, we saw no stereochemical preference in this shift counterintuitive to the highly stereodefined methyl- and hydride-shift reported in other sesquiterpene synthases. This has lead us to propose a longlived, stabilised allylic-cation in rapid equilibrium with the germacryl cation, causing a loss of sterochemical information. This allylic cation is then proposed to be quenched from the bulk solvent due to the opening of the active site in the product-release step.

Incubations of GdolS with FDP in buffers prepared with $^{2}H_2O$ and $H_2{}^{18}O$ are congruent with this proposal; indicating that both the proton and oxygen in the 4-hydroxyl group originate from the bulk solvent. This simple mechanism of carbocation quenching, incorporating the hydroxyl group is unusual in catalysis by terpene synthases due to the highly reactive nature of the carbocation intermediates. Terpene synthases typically function by desolvating the active site upon formation of the Michaelis complex to prevent water ingress and carbocation quenching.

Incubations with NDP, both racemic and the 3-(*R*)-stereoisomer, suggested that the cyclisation may go *via* the formation of the tertiary diphosphate NDP, a mechanism further supported by the ability of GdolS to turn over the shorter

10-carbon isoprenoid precursor, GDP, to cyclic products. This may suggest an insight into the evolutionary past of this enzyme; with the isomerisation to NDP a palimpsest for a previous mechanism that generated a product bearing a (*Z*)-double bond. Serendipitous mutations have lead to an enzyme which allows water to enter the active site and quench the intermediate prior to a second cyclisation.

Incubation with (2*Z*, 6*E*)-FDP however, which would be set up for direct cyclisation without the need for conversion to NDP if this was indeed the necessary active site conformation, produced only minor amounts of germacradien-4-ol amongst numerous other, unidentified, products. The proposed active-site conformation of 3-(*R*)-NDP potentially shows a favourable alignment of π*-anti-bonding orbitals. This could allow concerted ring-closure and allylic elimination of the diphosphate in a single step. The initial ionisation of FDP is widely accepted as being the chemical rate determining step, and so direct generation of the germacryl cation in a concerted manner could be postulated to accelerate the mechanism.

5.2 Structure–function investigation into germacradien-4-ol synthase (GdolS) by site-directed mutagenesis

Point mutations to the metal-binding domains proved detrimental to the catalytic efficiency of GdolS, highlighting the importance of these regions to catalysis with particular emphasis on the first aspartate in the DDxxD motif and to the whole of the NSE-motif. Interestingly, the arginine of the NSE-motif (arginine-218) was also implicated in the quenching of the intermediate carbocation by water to yield the product. Mutation of the arginine to glutamine caused the production of 45 % germacrene A, however, replacement of this residue with a lipophilic leucine reverted to the product distribution of wildtype. The NSE-arginine has been identified in X-ray crystal structures of AT-AS as part of a coordination sphere of a water molecule bound in the active site. A further arginine residue (ATAS arginine-299) and a serine (ATAS serine-303) were also identified in coordination to this water molecule and the corresponding residues in GdolS (tyrosine-303 and glutamate-307) were also targeted by site-directed mutagenesis. Mutations to these residues had only a minor effect on product distribution but were vital to the catalytic efficiency of the enzyme. The aromaticity of the tyrosine residue was vital, possibly in the stabilisation of carbocation intermediates, while the carbonyl functionality of

the glutamate side-chain is necessary. Mutation of the glutamate to a glutamine shows a minor reduction in catalytic activity while the generation of the methionine mutant was severely hampered in terms of activity.

To further investigate interactions between arginine-218 and proximal residues a helix-switch approach was adopted; changing the terminal loop-helix region of GdolS with the same, slightly longer region, from DCS; a δ-cadinene synthase available in our lab. Unfortunately this generated hybrid showed only sparing solubility due to incomplete folding and was completely inactive.

Following a report on the effect of the H-1 α loop on product distribution in fungal sesquiterpene synthases, this region was analysed and mutated in GdolS. Two mutants were generated to disrupt ionic interactions between the H-1 α loop and the active site, resulting in one wildtype-like enzyme (R229E) and an inactive mutant (D231L); indicating the importance of the aspartate residue in closing the active site to form the Michaelis complex.

5.3 Synthesis and evaluation of farnesyl diphosphate analogues

A range of epoxy-FDP analogues were synthesised and incubated with a number of sesquiterpene synthases. Epoxidation of the 2,3-double bond lead to an FDP analogue which was not a substrate for any of the sesquiterpene synthases tested. The 6,7-10,11-epoxy-FDP was also not a substrate for any of the enzymes, while the 6,7-epoxy-FDP was poorly turned over by AS and GdolS. The 10,11-epoxy-FDP was found to be turned over by a number of the enzymes and the product identified as a macrocyclic ether, resulting from the direct quenching of the farnesyl cation followed by a deprotonation. A second analogue was designed and synthesised to test this theory, the 10-hydroxy-11-ene-FDP, which also directly trapped the farnesyl cation, directly generating the same macrocyclic ether without an intermediate carbocation. The 10-hydroxy-11-ene-FDP was prepared racemically and the generation of the macrocyclic ether showed no stereochemical bias, attributed to the close proximity of the hydroxyl enforced in the substrate folding in the active site.

The ability of the 10,11-epoxy- and 10-hydroxy-11-ene-FDP's to quench a carbocation and generate the macrocyclic ether demonstrate the first reported trapping of the proposed farnesyl cation to date.

To further investigate the trapping of the farnesyl cation an analogue lacking the isopropyl tail was generated, and demonstrated that this analogue was still turned over by all cyclases, producing what was proposed to be the cyclic ether **176**. This would suggest that the isopropyllidine tail of the farnesyl chain is not necessary for enforcing the correct folding of the substrate in the active site; with the reactivity of the allylic cation and the confinement of the active site overriding the lack of this directing influence.

The 7-methylene-FDP was prepared from elongation of γ-geraniol with the transposition of the 6,7-double bond. This analogue was primarily synthesised to test the versatility of AS, but was also used to investigate if the other sesquiterpene synthases were able to more minor changes in configuration at the centre of the farnesyl chain. Pleasingly, AS was able to turnover this analogue to aristolochene through protonation of the 7-methylene double bond to drive the second cyclisation to the eudesmane cation. The other enzymes tested were also able to turn over this analogue to monocyclic products.

These results indicate a level of flexibility in the purely hydrocarbon half of FDP (C7 – C10), which enables many sesquiterpene synthases to accept analogues of FDP with modifications in this region. While the 6,7-epoxy-FDP was turned over by AS and GdolS, the effect of changing configuration and imposing stereochemistry at two centres proved too much for the other enzymes, and even in the cases of AS and GdolS turnover was lower, and these normally high-fidelity enzymes produced numerous products. The more minor change in the 7-methylene-FDP, affecting only the single centre on the farnesyl chain, was well accepted by all enzymes and turned over to primarily cyclic products.

5.4 Future approaches

This investigation has characterised germacradien-4-ol synthases and determined that cyclisation of FDP by this enzyme occurs by a stepwise loss of diphosphate followed by 1,10-ring closure to the germacryl cation. We have postulated that the subsequent 1,3-hydride shift generates a long-lived stabilised allylic cation, which is then quenched by water from the bulk solvent upon the opening of the active site. Despite a targeted approach using the apo-crystal structure and homology models to identify residues for site-directed mutagenesis we were unable to conclusively identify any residues solely responsible for the incorporation of water into the product. It would be of great

interest to obtain a crystal structure of GdolS with an inhibitor bound in the active site, since we have established 15-CF_3-FDP as a potent competitive inhibitor, this is an ideal choice. This would allow detailed analysis of the Michaelis complex and could reveal details of the folded conformation of FDP in the active site giving further detail to the mechanism both of cyclisation and of water incorporation.

The synthesis of a range of FDP-analogues has demonstrated the first trapping of the farnesyl cation, as well as highlighting the importance of the C10-isopropyl tail of FDP in enforcing cyclisation. Further analogues of interest are highlighted in Figure 5.1, including those designed to trap the cisoid allylic cation arising after isomerisation to NDP and further minor modifications to probe the scope of sesquiterpene synthases in generating novel terpenoids.

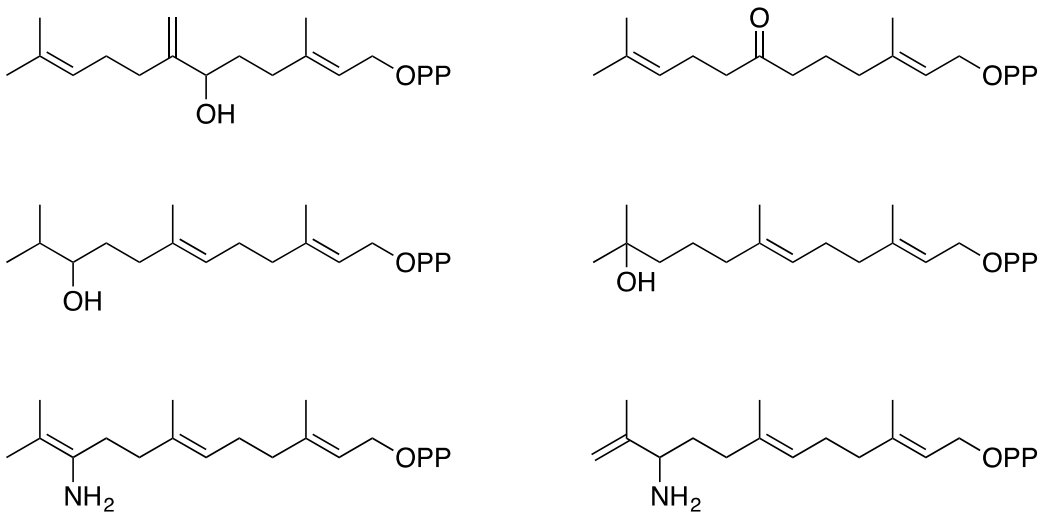


Figure 5.1 Prospective future analogues

6 Materials and Methods

6.1 Biological Methods

6.1.1 Materials

Oligonucleotide primers for site-directed mutagenesis were purchased from Eurofins or Sigma-Aldrich, the pET16b-SC1 plasmid construct was a generous gift from Prof. Yasuo Ohnishi, University of Tokyo; Restriction enzymes were purchased from NEB UK or Thermofisher Scientific, Pfu DNA polymerase and Unstained protein marker were purchased from Thermofisher scientific, dNTP's were purchased from Promega, [1-^{3}H]-FDP was purchased from American Radiolabelled Chemicals, 'Ecoscint™ O' scintillation fluid was purchased from National Diagnostics. All other chemicals were purchased from Sigma-Aldrich, Fisher Scientific or Melford. All the mutated and ligated constructs were confirmed by DNA sequence analysis using the internal Walesbiogrid facilities (School of Bioscience, Cardiff University, UK).

6.1.2 Bacterial strains and Preparation

A number of strains of *E. coli* were used in the course of this investigation; XL1-Blue cloning strain was used for the production of plasmid DNA and for the transformation of SDM reactions; Bl21(DE3) and BL21-Star™-(DE3) expression strains were used for test expressions of GdolS-WT and ultimately BL21-(DE3) was used for the expression of all GdolS proteins and mutants. Stratagene competent cells (purchased from Agilent as part of the Strataclone Blunt PCR cloning kit) were used for the production of the DCS-fragment cloning vector in the preparation of the GdolS-DCS hybrid.

6.1.3 Competent cells

Calcium chloride Buffer I

Calcium chloride (100mM) was dissolved in deionised water. The final volume was then made up and sterilised in an autoclave (121 °C, 20 minutes).

Calcium chloride Buffer II

Calcium chloride (100 mM) was dissolved in deionised water and glycerol (15% v/v) was added. The final volume made up with deionised water and in an autoclave (121 °C, 20 minutes).

Competent Cell Preparation

A 50 µL aliquot of the desired cell strain was incubated in LB-media overnight at 37 ºC, in the case of XL1-Blue cells, tetracycline was added to the media (15

mg L^{-1}). 1 mL of the overnight culture was used to inoculate 100 mL of LB-media (15 mg L^{-1} of tetracycline in the case of XL1-Blue) and grown at 37 °C, to an optical density of 0.6 at 600 nm (OD_{600}). The cells were put on ice for 15 minutes, then harvested by centrifugation (3400 g, 5 minutes), and the supernatant discarded. The pellet was gently re-suspended in calcium chloride buffer I (40 mL) and incubated on ice for 15 minutes. Centrifugation was repeated and the pellet gently re-suspended in calcium chloride buffer II and incubated on ice for 15 minutes. The cells were divided into 50 µL aliquots, frozen immediately in liquid N_2 and stored at -80 °C.

6.1.4 Supercompetent cells

Rubidium chloride buffer I

Potassium acetate (30 mM), rubidium chloride (100 mM) and calcium chloride (10 mM) were dissolved in deionised water, glycerol (15%) was added and the pH of the solution adjusted to 5.8 before addition of manganese chloride (50 mM). The final volume was made up with deionised water and sterilised by filtering through a sterile 0.2 µM syringe filter under aseptic conditions, and stored at 4 °C.

Rubidium chloride buffer II

3-(*N*-morpholino)propanesulfonic acid (MOPS, 10 mM), calcium chloride (75 mM) and rubidium chloride (10 mM) were dissolved in deionised water, glycerol (15) was added and the pH adjusted to 6.5. The final volume made up with deionised water and sterilised by filtering through a sterile 0.2 µM syringe filter under aseptic conditions, and stored at 4 °C.

Super-competent cell preparation

Supercompetent cells were prepared in the same manner as competent cells (6.1.3) but with rubidium chloride buffers I and II (0 and0) in place of calcium chloride buffers I and II (0 and 0) respectively.

6.1.5 Growth media and antibiotic solutions

Luria-Bertani (LB) medium

Tryptone (10 g L^{-1}), yeast extract (5 g L^{-1}) and sodium chloride (10 g L^{-1}) were dissolved in deionised water, the resulting solution was sterilised in an autoclave (121 °C, 20 minutes) and cooled to room temperature before use. Any required antibiotic was added immediately before use under aseptic conditions.

LB Agar plates

Tryptone (10 g L^{-1}), yeast extract (5 g L^{-1}), sodium chloride (10 g L^{-1}) and agar (15 g L^{-1}) were dissolved in deionised water, the resulting solution was sterilised in an autoclave (121 °C, 20 minutes). The agar was allowed to cool to approximately 40 °C before the required antibiotic was added and the LB agar poured into plates under aseptic conditions. Plates were allowed to cool and set before storage at 4 °C.

Ampicillin

Ampicillin was dissolved in deionised water to give a concentration of 100 mg mL^{-1}. The solution was filtered through a sterile 0.2 µM syringe filter under aseptic conditions and stored at 4 °C.

Kanamycin

Kanamycin was dissolved in deionised water to give a concentration of 50 mg mL^{-1}. The solution was filtered through a sterile 0.2 µM syringe filter under aseptic conditions and stored at 4 °C.

Tetracycline

Tetracycline was dissolved in deionised water to give a concentration of 15 mg mL^{-1}. The solution was filtered through a sterile 0.2 µM syringe filter under aseptic conditions and stored at 4 °C.

6.1.6 Cloning

Site-directed mutagenesis

Primers for SDM

Oligonucleotide primers were designed using PrimerX (www.bioinformatics.org/primerx) with the consideration given to the following criteria:

1. Both forward and reverse primer must contain the desired mutation and anneal to the same sequence on each strand of the template plasmid.
2. The primers should be between 25 and 45 bases long with the mutation in the centre of the primer.
3. The melting temperature, T_m, should be equal to or greater than 78 °C
4. The primers should begin and terminate with a G or a C and have 40 and 60 % content of GC bases. Due to the GC rich nature of *Streptomycetes* this was not always possible and an upper limit of 75 - 80 % was applied where necessary.

The primers are as follows in Table 6.1.

Table 6.1 Oligonucleotide primers for site-directed mutagenesis (5' to 3') with the changed codon underlined and the mutated base(s) in italics

Name	5'-Sequence-3'
GdolS-Y303F Fwd	CTGGCTCGCGGGTT*TT*CTCCACTGGGAGTC
GdolS-Y303F Rev	GACTCCCAGTGGAG*AAA*ACCCGCGAGCCAG
GdolS-Y303I Fwd	CTCTGGCTCGCGGGT*ATT*CTCCACTGGGAGTC
GdolS-Y303I Rev	GACTCCCAGTGGAG*AAT*ACCCGCGAGCCAGAG
GdolS-W306A Fwd	GTTACCTCCAC*GC*GGAGTCCCACAC
GdolS-W306A Rev	GTGTGGGACTCC*GC*GTGGAGGTAAC
GdolS-E307Q Fwd	GTTACCTCCACTGG*CA*GTCCCACACCCG
GdolS-E307Q Rev	CGGGTGTGGGAC*TG*CCAGTGGAGGTAAC
GdolS-E307M Fwd	GTTACCTCCACTGG*AT*GTCCCACACCCGC
GdolS-E307M Rev	GCGGGTGTGGGAC*AT*CCAGTGGAGGTAAC
GdolS-Y303T Fwd	CTCTGGCTCGCGGGT*AC*CTCCACTGGGAGTC
GdolS-Y303T Rev	GACTCCCAGTGGAG*GT*ACCCGCGAGCCAGAG
GdolS-R172K Fwd	CTATCTGACGCTCA*AA*CGCGGTACCGCC
GdolS-R172K Rev	GGCGGTACCGCGT*TT*GAGCGTCAGATAG
GdolS.SmaIns F	GACCAACTCCCGG*GC*GCCTGCCGCTC
GdolS.SmaIns R	GAGCGGCAGGCG*CC*CGGGAGTTGGTC
GdolS.R172Q Fwd	CTATCTGACGCTC*CAG*CGCGGTACCGCCG
GdolS.R172Q Rev	CGGCGGTACCGCG*CTG*GAGCGTCAGATAG
GdolS-D80E Fwd	CTACTTCCTCTTCG*AA*GACCAGTTCGACAG
GdolS-D80E Rev	CTGTCGAACTGGTC*TT*CGAAGAGGAAGTAG
GdolS-D81E Fwd	CCTCTTCGACG*AA*CAGTTCGACAGCC
GdolS-D81E Rev	GGCTGTCGAACTG*TT*CGTCGAAGAGG
GdolS-D84E Fwd	CGACCAGTTCG*AA*AGCCCGCTCGGG
GdolS-D84E Rev	CCCGAGCGGGCT*TT*CGAACTGGTCG
GdolS-N218Q Fwd	CATCCCGTCGTTCACC*CAG*GACGTGCGCTCCTTC
GdolS-N218Q Rev	GAAGGAGCGCACGTC*CTG*GGTGAACGACGGGATG
GdolS-S222A Fwd	CCAATGACGTGCGC*GCG*TTCGCACAGGAGTC
GdolS-S222A Rev	GACTCCTGTGCGAA*CGC*GCGCACGTCATTGG
GdolS-E226D Fwd	CTTCGCACAGG*AT*TCCGAGCGCGGC
GdolS-E226D Rev	GCCGCGCTCGGA*AT*CCTGTGCGAAG
GdolS-N218L Fwd	CATCCCGTCGTTCACC*CTG*GACGTGCGCTCCTTC
GdolS-N218L Rev	GAAGGAGCGCACGTC*CAG*GGTGAACGACGGGATG
GdolS-D80N Fwd	GGTTCTACTTCCTCTTC*AAT*GACCAGTTCGACAGCC
GdolS-D80N Rev	GGCTGTCGAACTGGTC*ATT*GAAGAGGAAGTAGAACC

GdolS-D81N Fwd	CTACTTCCTCTTCGAC*AAT*CAGTTCGACAGCCCGC
GdolS-D81N Rev	GCGGGCTGTCGAACTG*ATT*GTCGAAGAGGAAGTAG
GdolS-D84N Fwd	GACGACCAGTTC*AAT*AGCCCGCTCGG
GdolS-D84N Rev	CCGAGCGGGCT*ATT*GAACTGGTCGTC
GdolS-R134E Fwd	CTTGGCATGACGGACGAATGGCGCGCACGGGC
GdolS-R134E Rev	GCCCGTGCGCGCCATTCGTCCGTCATGCCAAG
GdolS-N218E Fwd	CATCCCGTCGTTCACC*GAA*GACGTGCGCTCCTTC
GdolS-N218E Rev	GAAGGAGCGCACGTC*TTC*GGTGAACGACGGGATG
GdolS-N218T Fwd	CGTCGTTCACCA*CC*GACGTGCGCTCC
GdolS-N218T Rev	GGAGCGCACGTC*GG*TGGTGAACGACG
GdolS-R134A Fwd	GCATGACGGAC*GCG*TGGCGCGCACGG
GdolS-R134A Rev	CCGTGCGCGCCA*CGC*GTCCGTCATGC
GdolS-D231L Fwd	GTCCGAGCGCGGC*CTG*GTGGCCAACCTG
GdolS-D231L Rev	CAGGTTGGCCAC*CAG*GCCGCGCTCGGAC
GdolS-R229E Fwd	CACAGGAGTCCGAG*GAA*GGCGACGTGGCCAAC
GdolS-R229E Rev	GTTGGCCACGTCGCC*TTC*CTCGGACTCCTGTG

Site-directed mutagenesis protocol

Site-directed mutagenesis was carried out using the polymerase chain reaction (PCR), The reaction mixture was made up according to Table 6.2 and PCR carried out in a thermocycler programmed according to Table 6.3.

After the SDM program was completed the reaction mixture was treated with 1 μL of *Dpn*I endonuclease and incubated at 37 ºC for 1 hour to cleave the parental template plasmid. *Dpn*I is specific to *N*6-methyladenine within the recognition sequence (G mA^TC), found in DNA isolated from almost all *E.coli*.

Supercompetent *E.coli* XL1-blue were then transformed with 5 μL of the digested SDM mixture, Section 6.1.9.

Table 6.2 SDM reaction mixture

Component (Stock concentration)	Volume (μL)
Forward primer (10 μM)	2
Reverse primer (10 μM)	2
dNTP's (10 mM)	1
Template plasmid	0.5
Pfu DNA polymerase	0.5
Pfu buffer + $MgSO_4$	5
d. H_2O	39

Table 6.3 SDM program conditions

Step	Standard Program		Long Program	
	Temperature (°C)	Time (min)	Temperature (°C)	Time (min)
Initial Denaturation	95	3	95	3
Denaturation	95	1	95	1
Annealing	55	2	65	2
Elongation*	72	10	72	12
Final Elongation	72	10	75	10
	*Denaturation, annealing, elongation steps cycled 15 times		*Denaturation, annealing, elongation steps cycled 15 times	

Stocks of separate dNTP's were purchased at 100 mM, these were diluted to a working concentration of 10 mM of each dNTP (40 mM total dNTP concentration), aliquoted and stored at -20°C.

Polymerase chain reaction (PCR)

Polymerase chain reaction was used to amplify the desired section of the cDNA for DCS while also inserting an SmaI restriction site prior to this section.

Primers for PCR

The primers were designed manually. For the reverse primer, starting from the 3′ end of the DCS cDNA, the sequence had no mismatched basepairs and it was ensured there was a suitable overlap after the BamHI site for efficient operation of the endonuclease. The forward primer contained a mismatched section containing the SmaI restriction site and a small section identical to GdolS prior to the loop-helix. The DCS complementary section was of equal length to the non-matched section and contained no mismatched basepairs, Table 6.4

Table 6.4 PCR Primers

	5′-Sequence-3′
DCS.PCR Fwd	5'- CTCCCGGGCGCCTGCCGCTCGTTTTTGAAACCAACAGAA -3' -SmaI- |-F--L--K--P--T--E- GdolS helix end|DCS section start
DCS.PCR Rev	5′-ATTCGGGATCCTCAAAGTGCAATTGGTTCAATGAGC-3′

PCR Protocol

PCR was carried out using Pfu DNA polymerase and a protocol designed based on the recommended conditions. The reaction mixture was made up according

to Table 6.2 and PCR carried out using the designed protocol, Table 6.5. After PCR 5 µL of the product was analysed by electrophoresis on 2 % agarose and upon successful amplification of the segment, a sample was cloned into Strataclone holding vector and confirmed by DNA sequence analysis.

Table 6.5 PCR protocol

	Temperature (ºC)	Time (min)
Initial Denaturation	95	2
Denaturation	95	1
Annealing	65	0.5
Extension*	74	1
Final Extension	74	5
	*Denaturation, annealing, extension steps cycled 25 times	

Restriction digest of DNA

Manufacturer's guidelines regarding temperature, buffers and denaturation were followed in all restriction enzyme digests. Restriction digests were performed sequentially with *Sma*I (3.5 h, 25 °C) followed by *BamH*I (3 h, 37 °C). *Sma*I was heat inactivated at 65 °C for 15 minutes before addition of the second enzyme. After digestion DNA fragments were isolated by agarose gel purification (0).

Ligation of DNA into expression vectors

Following restriction digest and agarose gel purification, the required gene and expression plasmid were ligated to form new expression vector constructs. Plasmid and insert were incubated in a 1:5 molar ratio, respectively, with T4 ligase buffer (2 µL) and sterile dH_2O in a volume of 20 µL. T4 ligase 1 (µL) was added and the ligation reaction incubated for 12 hours at 16 °C. The reaction was frozen for later use or immediately used to transform super-competent cells.

Cloning into StrataClone PCR cloning vector

StrataClone Blunt Cloning Buffer (3 µL), PCR product (2 µL) and StrataClone Blunt Vector Mix amp/kan (1 µL) were incubated at room temperature for 5 minutes and 2 µL of the reaction used to transform Stratagene ultra-competent cells.

6.1.7 DNA visualisation

TAE buffer

Tris base (242 g), glacial acetic acid (57.1 mL) and 100 mL EDTA (0.5 M, pH 8.0) were dissolved in deionised water and the total volume taken to 1 L with

deionised water. Immediately prior to use the solution was diluted in a 1:49 ratio with deionised water.

DNA loading buffer

Glycerol (3 mL) and 0.5% bromophenol blue solution (2 µL) were added to 7 mL deionised water.

Agarose gel electrophoresis

Agarose gels were prepared using dilute TAE buffer (0) in a 1-2 % w/v ratio dependent on the size of fragment being analysed. 5 µL DNA loading buffer (0) was added to the DNA sample which was then loaded onto the gel and run at a constant voltage of 100 V for 1 hour before being placed in a dilute ethidium bromide solution for 10 minutes whilst shaking. The DNA bands were visualised using a Syngene GeneFlash UV light box (Syngene, Cambridge, UK) and compared to a molecular weight ladder run on the same gel.

6.1.8 DNA purification

Miniprep of DNA

A single colony from an agar plate, or a spike from a glycerol stock, harbouring the required transformed cells was used to inoculate 10 mL of LB medium (containing the appropriate selective antibiotic) and incubated overnight at 37 °C whilst shaking. Cells were harvested by centrifugation (3400 g, 5 minutes). The QIAprep spin miniprep kit (QIAGEN, Crawley, UK) was used with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) to purify the plasmid following the manufacturer's protocol.

Agarose gel DNA purification

The required samples underwent agarose gel electrophoresis (0) and the DNA visualised by staining in dilute ethidium bromide solution. DNA bands were visualised using a Syngene GeneFlash UV light box (Syngene, Cambridge, UK) and the appropriate DNA bands cut from the agarose gel using a razor blade. DNA was extracted from the agarose gel using a QIAquick gel extraction kit (QIAGEN, Crawley, UK) with EconoSpin All-in-1 mini spin columns (Epoch Biolabs, Inc, TX, USA) following the manufacturer's protocol.

6.1.9 Transformation of competent cells

The appropriate competent cells and plasmid DNA were allowed to thaw slowly on ice. 0.5 – 1 µL plasmid DNA solution (5-10 µl in the case of ligation or SDM reactions) was transferred to the competent cells and mixed gently. The

cell/DNA mixture was incubated on ice for at least 20 minutes, heat shocked at 42 °C for 45 seconds and returned to ice for at least 2 minutes. Non-selective LB medium (1 mL) was added to the cell/DNA mixture and incubated at 37 °C for 1 hour whilst shaking. Cells were harvested by centrifugation (3300 g, 1 minute, Eppendorf centrifuge 5415R) and the supernatant solution discarded. The cell pellet was re-suspended in 100 µL of LB medium, plated on agar plates containing the appropriate antibiotic and incubated at 37 °C overnight.

6.1.10 Expression

A single transformed colony was used to inoculate 100 mL of LB-media containing ampicillin (100 mg L^{-1}); this was grown overnight at 37 °C with shaking. This starter culture was used to inoculate LB-media containing antibiotic (5 mL of culture per 500 mL of LB medium). The cultures were incubated at 37°C until an optical density of 0.6 at 600 nm was reached, at which time gene expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG) and incubated for the required time and temperature as judged by test expression, Table 6.6. Cells were harvested by centrifugation (3400 g, 10 minutes) and cell pellets stored at -20 °C.

Test expression

Test expressions were used to optimise the over-expression for production of proteins, investigating factors such as expression cells, temperature of expression, concentration of IPTG and length of expression. They were typically carried out on a 1 L scale, and a 1 mL sample was removed immediately prior to induction and at hourly intervals after induction. The cells in these samples were harvested by centrifugation, (3300 g, 1 minute, Eppendorf centrifuge 5415R) the pellets resuspended in SDS-sample buffer and analysed by SDS-PAGE.

Table 6.6 Expression conditions for GdolS and mutants

Protein	Temperature °C	[IPTG] mM	Time h
GdolS-WT	37	0.5 / 0.2	3 – 4
D80E	37	0.2	4
D80N	37	0.2	4
D81E	37	0.2	4
D81N	37	0.2	4
D84E	37	0.2	4
D84N	37	0.2	4
E226D	37	0.2	4
S222A	37	0.2	4

N218Q	37	0.2	4
N218L	37	0.2	4
N218T	16	0.2	21
N218E	16	0.2	21
R172K	37	0.2	4
R172E	16	0.2	21
Y303F	37	0.2	4
Y303I	37	0.2	4
Y303T	16	0.1	21
W306A	37	0.2	4
E307Q	37	0.2	4
E307M	37	0.2	4
GdolS-DCS hybrid	16	0.1	6 - 21
D274G	37	0.2	4
D231L	37	0.2	4
R229E	37	0.2	4
R134E	37	0.2	4
R134A	37	0.2	4

6.1.11 Purification

Cell lysis

Induced cells harvested after expression were re-suspended in buffer (dependent on purification technique) and lysed by sonication (5 second sonication period followed by 10 second rest period for a total of 15 minutes). The solution was clarified by centrifugation at 4 °C, (38000 g, 30 minutes).

Basic extraction

Basic extraction was used to denature and refold protein found as insoluble inclusion bodies in order to solubilise the protein. The cell pellet after clarification (0) was re-suspended in buffer (20 mL, 20 mM Tris, pH 8.0) and stirred at 4 °C for 30 minutes before the pH was adjusted to 12 with sodium hydroxide at which point the solution becomes clearer. After stirring for a further 30 minutes at 4 °C, 2-mercaptoethanol is added to a concentration of 10 mM and the pH gradually reduced to 8.0 by dropwise addition of 1 M hydrochloric acid. The solution at pH 8.0 is stirred at 4 °C for a further 30 minutes at 4 °C before centrifugation (38000 g, 30 minutes), and the resulting supernatant and pellet analysed by SDS-PAGE.

Ni-NTA affinity chromatography

Ni-NTA Buffer I

Tris-base (50 mM, 6.2 g), Sodium chloride (100 mM, 5.9 g) and Imidazole (10 mM, 0.78 g) were dissolved in 800 mL deionised water and the pH adjusted to 8.0. The solution was made up to a final volume of 1 L with deionised water

Ni-NTA Buffer II

Tris-base (50 mM, 6.2 g), Sodium chloride (100 mM, 5.9 g) and Imidazole (250 mM, 17.2g) were dissolved in 800 mL deionised water and the pH adjusted to 8.0. The solution was made up to a final volume of 1 L with deionised water.

Ni-NTA Buffer III

Tris-base (50 mM, 3.1 g), Sodium chloride (100 mM, 2.9 g) and Imidazole (500 mM, 17.2g) were dissolved in 400 mL deionised water and the pH adjusted to 8.0. The solution was made up to a final volume of 500 mL with deionised water.

Ni-NTA affinity chromatography protocol

Harvested supernatant was applied to a column equilibrated with Ni-NTA buffer I and the flow-through collected. The column was then washed with 10 column volumes (CV) Ni-NTA buffer I prior to elution with 5 CV Ni-NTA buffer II, and 5 CV Ni-NTA buffer III. All fractions were retained and analysed by SDS-PAGE, those containing the target protein were then dialysed.

Dialysis

Fractions from purification found to contain the desired protein were dialysed to remove the residual imidazole. Fractions of equal purity are combined in dialysis tubing (MWCO 12 – 14 kDa) and stirred overnight (16 h) in buffer (5 L,10 mM Tris, pH 8.0). After dialysis the protein solution was concentrated by ultrafiltration (AMICON system, YM 30).

6.1.12 SDS-PAGE

Samples of protein were analysed by Sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were prepared, cast and run using the Mini-PROTEAN system from Bio-Rad.

SDS resolving buffer

Tris-Base (1.5 M, 27.23 g) was dissolved in deionised water and the pH adjusted to 8.8. The final volume was taken to 150 mL.

SDS stacking buffer

Tris-Base (0.5 M, 6.06 g) was dissolved in deionised water and the pH adjusted to 6.8. The final volume was then taken to 100 mL.

Electrode running buffer 10X

Tris-base (0.25 M, 30.3 g) and Glycine (2 M, 150.14 g) were dissolved in 600 mL deionised water, once dissolved SDS (10% w/v, 10 g) was added and left to

stir slowly to dissolve. Once all the SDS had dissolved the final volume was brought up to 1 L with deionised water. Electrode running buffer 10X was diluted 10-fold prior to use.

SDS sample buffer

Deionised water (3.55 mL), SDS stacking buffer (1.25 mL), glycerol (2.5 mL), bromophenol blue (0.2 mL of 0.6% w/v solution), SDS (2 mL of 10% w/v solution) and β-mercaptoethanol were mixed together and stored at room temperature.

Resolving gel (12 %)

30% acrylamide/bis-acrylamide (4.0 mL), resolving buffer (2.5 mL), 10% (w/v) SDS (0.1 mL) and deionised water (3.4 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

Stacking gel (4%)

30% acrylamide/bis-acrylamide (1.7 mL), stacking buffer (2.5 mL), 10% (w/v) SDS (0.1 mL) and deionised water (5.7 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

SDS gel stain

Coomasie brilliant blue R-250 (0.25 g) was dissolved in 1:1 Ethanol:Water (90 mL), once dissolved glacial acetic acid (10 mL) was added.

SDS gel destain

Deionised water (500 mL) and ethanol (400 mL) were mixed together, and then glacial acetic acid (100 mL) was added.

SDS-PAGE protocol

Monomer solutions were prepared as above and the resolving gel poured, covered with a layer of isopropanol and allowed to polymerise. Isopropanol was then removed and the stacking gel poured, a comb inserted to create loading wells and allowed to polymerise. Fraction samples (10 μL) were incubated with SDS sample buffer (10 μL, 0) for 10 minutes at 85 °C and loaded into the gel. Electrode running buffer (0) was added and the gel allowed to run at 150 V for 1 hour.

Visualising

Once the gel had run, it was removed from the plates and stained for 30 minutes with SDS gel stain, and then destained with gel destain in 30 minutes periods until bands were visible.

6.1.13 BN-PAGE

The non-denaturing polyacrylamide gel system was used for analysis of the GdolS-R134 mutants.

Gel buffer

BisTris (1.25 M, 26.13 g) was dissolved in deionised water and the pH adjusted to 7.5. The final volume was taken to 100 mL.

Anode running buffer 10X

BisTris (0.5 M) and Tricine (0.5 M) were dissolved in deionised water and the pH adjusted to 6.8. The final volume was brought up to 1 L with deionised water. Anode running buffer 10X was diluted 10-fold prior to use.

Cathode buffer 1X

To Anode running buffer was added Coomasie G250 brilliant blue dye (0.02% w/v) and stirred until dissolved.

Sample buffer

BisTris (50 mM, 105 mg) and sodium chloride (50 mM, 30 mg) were dissolved in 5 mL deionised water. Followed by the addition of glycerol (10% w/v, 1 mL) and Ponceau S dye (0.001% w/v, 0.1 mg). The pH was adjusted to 7.2 with HCl and the final volume made up to 10 mL.

Resolving gel (10 %)

30% acrylamide/bis-acrylamide (3.3 mL), gel buffer (2.9 mL) and deionised water (3.7 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

Stacking gel (4%)

30% acrylamide/bis-acrylamide (1.3 mL), stacking buffer (2.9 mL) and deionised water (5.7 mL) were mixed together. Immediately prior to pouring, freshly made 10% ammonium persulfate (APS, 50 μL) and N,N,N',N'-

tetramethylenediamine (TEMED, 20 μL) were added and the solution gently mixed to initiate polymerisation.

BN-PAGE protocol

BN-polyacrylamide gels were prepared as for SDS-PAGE following the above recipes. Anode running buffer was added to the tank and the cathode well filled with cathode running buffer. The gel was allowed to run at 150 V for 1h30.

Visualising

Once the gel had run, it was removed from the plates and stained and destained as for SDS-PAGE.

6.1.14 Bradford assay

The Bradford assay was used to determine the concentration of protein in solution

Bradford reagent

Brilliant blue G250 (20 mg) was dissolved in 2 mL of ethanol and H_3PO_4 added (80%, 20 mL). The final volume was brought up to 200 mL with dH_2O and stored in the dark at 4 °C.

Bradford assay

A 1 mg mL^{-1} sample of bovine serum albumin (BSA) in dH_2O was used to prepare a range of concentrations from 0 to 100 μg mL^{-1} in 200 μL deionised water. Bradford reagent (1 mL, 0) was added and the absorbance measured from 400 to 600 nm using a Jasco V-660 spectrophotometer. The ratio of absorbances at 590 nm and 450 nm was calculated and used to plot a standard curve. The procedure was repeated for the purified proteins and the ratio of absorbances compared to the BSA standard to calculate the concentration of each sample.

6.1.15 Circular dichroism spectroscopy

CD measurements

CD spectrum

Circular dichroism experiments were performed on an Applied PhotoPhysics Chirascan spectrometer. Spectra of 10 μM protein in dH_2O (or 20 mM Tris) were measured between 190 nm and 300 nm in a 10 mm quartz cuvettes under N_2 with a 50 nm/min scan speed, 0.5 nm data pitch, 1 nm bandwidth and 0.5 s response time.

CD melt

Circular dichorism melts were performed measuring at 222 nm between 4 °C and 96 °C with a temperature ramp of 2 °C min^{-1}, step of 0.5 °C and tolerance of 0.2 °C, with measurements taken for 12 s per point.

CD spectra melt

Circular dichroism spectra melts were performed with the measurement of full spectra between 190 nm and 300 nm between 5 °C and 90 °C with stepped temperature ramping enabled. Measurements were taken at 5 °C intervals with a 300 s settling time between measurements and a tolerance of 0.2 °C.

pH spectra

For the spectra measured at different pH values 10 μM protein in buffer was used (20 mM; pH 6, 4-morpholineethanesulfonic acid, MES; pH 7 - 9, Tris; pH 10, N-cyclohexyl-2-aminoethanesulfonic acid, CHES).

Mean residue ellipticity (MRE)

The spectra obtained from the CD spectrometer were converted to MRE using

Equation 1:

Equation 1

$$[\theta]_{MRE} = \frac{[\theta]}{\mathbf{10}.n.c.l}$$

Where $[\theta]$ = CD signal in millidegrees, n = number of peptide bonds, c = molar concentration of sample and l = pathlength of the cuvette in cm.

6.1.16 Size-exclusion chromatography

Size exclusion was carried out using a Superdex 200 10/300 GL prepacked gel filtration column (24 mL CV) with an AKTA FPLC system. Samples of 500 μL were loaded into the loop and washed through the column with 1.5 CV of buffer at a rate of 0.5 mL min^{-1} while monitoring UV absorbance at 280 nm.

Size-exclusion buffer I

Tris base (2.4 g, 20 mM), ethylenediaminetetraacetic acid (EDTA, 1.5 g, 5 mM) and 2-mercaptoethanol (350 μL, 5 mM) were dissolved in deionised water. The pH was adjusted to pH 8.0 and the final volume taken to 1 L. The solution was

filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

Size-exclusion buffer II

Tris base (2.4 g, 20 mM), ethylenediaminetetraacetic acid (EDTA, 1.5 g, 5 mM), sodium chloride (8.8 g, 150 mM) and 2-mercaptoethanol (350 µL, 5 mM) were dissolved in deionised water. The pH was adjusted to pH 8.0 and the final volume taken to 1 L. The solution was filtered and degassed via vacuum pump (Vacuubrand GmbH +CO KG, MD4C, Werthein, Germany) and stored at 4 °C.

6.1.17 GC-MS analysis of products

The pentane extracted products of small (250 µL) analytical incubations of enzymes with FDP were analysed by GC-MS.

Incubation buffer

Tris-base (1.2 g, 50 mM) and Magnesium chloride (200 mg, 5 mM) were dissolved in deionised water, 2-mercaptoethanol (70 µL, 5 mM) was added and the pH adjusted to 8 before the final volume was made up to 200 mL and stored at 4 °C.

Incubation protocol

The standard reaction mixture of 1 µM enzyme and 200 µM FPP in 250 µL incubation buffer was overlaid with 1 mL pentane. This mixture was incubated whilst shaking at room temperature for 4-6 hours before extraction of products into the pentane layer by vortexing for 10 seconds and analysis by GC-MS.

GC-MS method

GC-MS analyses were run on a Hewlett Packard 6890 GC fitted with a J&W scientific DB-5MS column (30 m x 0.25 mm internal diameter) and a Micromass GCT Premiere detecting in the range m/z 50-800 in EI+ mode with scanning once a second with a scan time of 0.9. A temperature gradient of 50ºC (1 minute hold) to 150ºC (4ºC min^{-1}, 13 minute hold) was used.

6.1.18 GC-FID analysis

GC-FID analysis was performed on an Agilent 7890A GC system fitted with a SUPELCO Aztec CHIRALDEX™ B-DM silica capillary column.

GC-FID method I

A 5 - 10 µL sample was injected with a 50:1 split, the oven temperature was held at 70 ºC for 3 minutes and then raised 5 ºC min^{-1} to 160 ºC and held for 30 minutes.

GC-FID method II

A 5 - 10 µL sample was injected with a 50:1 split, the oven temperature was held at 90 ºC for 2 minutes and then raised 5 ºC min^{-1} to 170 ºC and held for 40 minutes.

GC-FID method III

A 5 - 10 µL sample was injected with a 50:1 split, the oven temperature was held at 155 ºC for 60.

GC-FID method IV

A 5 - 10 µL sample was injected with a 50:1 split, the oven temperature was held at 100 ºC for 2 minutes and then raised 2 ºC min^{-1} to 150 ºC and held for 30 minutes.

6.1.19 Measurement of steady-state kinetic parameters

Kinetics buffer

4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 2.4 g, 50 mM) and magnesium chloride (100 mg, 2.5 mM) were dissolved in deionised water, 2-mercaptoethanol (70 µL, 5 mM) was added and the pH adjusted to 8 before the final volume was made up to 200 mL and stored at 4 °C.

For the determination of optimum Mg^{2+} concentration a second identical kinetics buffer was prepared without the magnesium chloride.

Kinetics protocol

Steady-state kinetics assays based on previous methods were carried out using [1-^{3}H]-(*E,E*)-FDP (240000 dpm $nmol^{-1}$) in buffer. Reactions (250 µL) were initiated by addition of enzyme (10 – 30 nm) to assay buffer solutions containing [1-^{3}H]-(*E,E*)-FDP (0-7 µM) at 0 °C overlaid with 1 mL of hexane. The resulting enzymatic preparations were incubated at 30 °C for 10 minutes, cooled on ice and immediately quenched by addition of EDTA (50 µL, 0.5 M) and vortexing for 30 s. The organic overlay was then passed through a small silica column (~500 mg) into 15 mL EcoScint™ fluid (National Diagnostics), the aqueous portion was further extracted with a further two 750 µL portions of 11:1 hexane/diethyl ether by vortexing for 10 s and the organic extracts

passed through the silica column into the EcoScint™ fluid. The silica column was washed with a 750 µL portion of 11:1 hexane/diethyl ether before analysis of the combined organic extracts in EcoScint™ fluid on a scintillation counter (Packard 2500 TR™) in ^{3}H mode for 4 min per sample. Background samples were run identically, across the range of FDP concentrations without the addition of enzyme. Background counts were plotted and subtracted from the enzyme reaction data. The K_M and k_{cat} were calculated by fitting the data to the Michaelis-Menten equation, Equation 2, using Systat Sigmaplot or KaleidaGraph:

Equation 2

$$V = \frac{V_{max}[S]}{K_M + [S]}$$

The turnover number, k_{cat}, was derived from the calculated V_{max} using Equation 3.

Equation 3

$$k_{cat} = \frac{V_{max}}{t.c.A}$$

Where V_{maz} is in counts per minute (cpm), t is the incubation time in seconds, c is the enzyme concentration in µM and A is the activity of the FDP in cpm µM^{-1}.

Kinetics on alternate substrates

For measurement of steady-state kinetic parameters with alternate substrates (*rac*-(^{3}H)-NDP, (^{3}H)-*(3R)*-NDP, (^{3}H)-10F-FDP) the same procedure as previously described was followed but the initial run was prepared with a wider range of substrate concentrations (1 – 15 µM) to identify a suitable range.

6.1.20 Inhibition studies

Two methods of studying inhibition were used, both methods relying on double-reciprocal plots for identification of the type of inhibition.

Lineweaver-Burk

The Lineweaver-Burk plot arises from taking the reciprocal of the Michaelis-Menten equation (Equation 2) to give Equation 4:

Equation 4

$$\frac{1}{V} = \frac{1}{V_{max}} + \frac{K_M}{V_{max}[S]}$$

A plot of $1/V$ against 1/[S] gives a straight line with the x-axis intercept at $-1/K_M$, the y-axis intercept $1/V_{max}$ and the gradient K_M/V_{max}.

Inhibition studies using this method of representation comprised of a series of incubations as above with increasing concentrations of inhibitor present, i.e. in each series the concentration of inhibitor is constant while concentration of substrate increases. The data was treated as described above to generate a Lineweaver-Burk plot for the calculation of K_i, where $K_{M(apparent)} = K_M\ (1 + [I]/K_I)$.

Dixon

In a Dixon plot the concentration of inhibitor is plotted against the reciprocal rate varying the substrate concentration each series, incubations were carried out with increasing [I] at constant [S], [S] was then increased to generate multiple series. The K_i can be read from the plot as the intersection of the lines. The data was then replotted in a Cornish-Bowden plot, S/v against [I], to determine the type of inhibition.

Equation 5

$$\frac{1}{v} = \frac{K_M}{V_{max}[S]K_i}[I] + \frac{1}{V_{max}}\left(1 + \frac{K_M}{[S]}\right)$$

6.1.21 Preparative-scale enzymatic incubation

Preparative-incubations were used for the production of milligram quantities of enzymatic products for the purposes of analysis. Incubations were scaled to a suitable size dependent on the quantity of substrate being used to a final substrate concentration of 0.3 mM in buffer (50 mM tris, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, pH 8.0) with a maximum enzyme concentration of 4 μM.

Incubation

Buffer was prepared and the enzyme added at room temperature while gently stirring, 50 % of the total FDP substrate was added and the incubation overlaid with pentane (5 – 10 % incubation volume) and stirred at room temperature for 24 h. The remaining FDP-substrate was then added at stirred for a further 12 – 24 h before it was cooled to 4 °C and stirred for a further 64 h.

Extraction

The pentane overlay was removed and the aqueous layer was further washed with pentane (10 x 5% V_{tot}) by gentle swirling and slow separation. The

combined pentane extracts were washed with brine and solvent removed under reduced pressure (500 mbar) and 25 °C to yield the terpene product as an oil.

6.1.22 Calculation of errors

Standard Errors of the Mean

The errors in this work are expressed as standard error of the mean (σ_M) that is defined by the following equations:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

Where n is the number of samples and σ is the standard deviation, defined as:

$$\sigma = \sqrt{\frac{\sum(X - M)^2}{(n-1)^2}}$$

Where X is each value in the sample and M is the mean of the sample.

Propagation of the Errors

Propagation of errors was performed where appropriate using the following equations:

If $Z = X + Y$ or $Z = X - Y$ then $\Delta Z = \sqrt{\Delta X^2 + \Delta Y^2}$

If $Z = X * Y$ or $Z = X/Y$ then$\Delta Z = \sqrt{\left(\frac{\Delta X}{X}\right)^2 + \left(\frac{\Delta Y}{Y}\right)^2}$

If $Z = \ln X$ then $\Delta Z = \frac{\Delta X}{X}$

If $Z = e^X$ then $\Delta Z = \Delta X e^X$

Where X and Y are the experimentally measured values, ΔX and ΔY are their respective errors, Z is the calculated value and ΔZ the propagated error.

6.1.23 Normalisation

Where data has been normalised it has been performed using unity-based normalisation using the following equation:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Where X' is the normalised value, X is the original value and $X_{\min}$ and $X_{\max}$ are the minimum and maximum values in the dataset, respectively.

6.2 Synthesis

6.2.1 General experimental

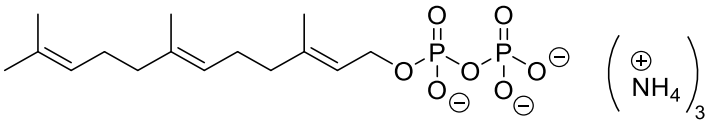
All chemicals were purchased from Sigma-Aldrich, Alfa-Aesar or Fisher Scientific and used without further purification unless otherwise stated. Anhydrous tetrahydrofuran (THF), diethyl ether, toluene and acetonitrile were obtained from a MBraun SPS800 solvent purification system. Dichloromethane, and triethylamine were distilled from calcium hydride and KOH under nitrogen respectively.

^{1}H and ^{13}C NMR spectra were measured on a Bruker Avance 500 NMR spectrometer or a Bruker Fourier300 NMR spectrometer and are reported as chemical shifts in parts per million downfield from tetramethylsilane, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling (to the nearest 0.5 Hz) and assignment, respectively. ^{1}H, ^{13}C and ^{31}P NMR spectra were measured on a Bruker Avance 500 NMR spectrometer and are reported as chemical shift downfield from tetramethylsilane, coupling constant where appropriate and assignment. Assignments are made to the limitations of COSY, DEPT 90/135, gradient HSQC and gradient HMBC spectra. ^{2}H, ^{19}F and ^{31}P NMR spectra were recorded on a Jeol Eclipse +300 NMR spectrometer.

Thin layer chromatography was performed on pre-coated aluminium plates of silica G/UV_{254}. TLC visualizations were performed with 4.2% ammonium molybdate and 0.2% ceric sulfate in 5% H_2SO_4, or 0.1% berberine hydrochloride in EtOH or UV light. Reverse phase HPLC was performed on a system comprising of a Dionex P680 pump and a Dionex UVD170U detector unit.

6.2.2 Experimental

(2E,6E)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl tris-ammonium diphosphate 19



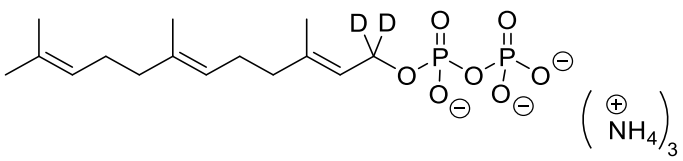
A stirring solution of farnesol (200 mg, 0.90 mmol) and 2,4,6-collidine (654 mg, 713 μL, 5.4 mmol) in anhydrous DMF (20 mL) was cooled to 0 °C and methanesulfonyl chloride (206 mg, 140 μL, 1.8 mmol) added. The resulting milky solution was stirred for 15 minutes before the addition of LiCl (152 mg,

3.6 mmol) and then left to stir at 0 °C for a further 2 h. The reaction was quenched with d.H_2O and washed with hexane (5 x 5 mL). The combined organic fractions were washed with $CuSO_4$, $NaHCO_3$ and brine, dried ($MgSO_4$) and the solvent removed under reduced pressure, and dried under high-vacuum for 1 h to yield the chloride, which was used without further purification.

The chloride intermediate was taken up in anhydrous acetonitrile and tris(tetrabutylammonium)hydrogen diphosphate (2.4 g, 2.7 mmol) added and stirred under N_2 for 16 h. The acetonitrile was removed under vacuum and the sticky brown residue dissolved in 2 mL of ion-exchange buffer (25 mM NH_4HCO_3, 2% iPrOH), and the resulting cloudy solution was passed slowly through a column of Dowex 50W-X8 cation-exchange resin pre-equilibriated in ion-exchange buffer. Fractions were eluted with the same buffer and analysed by TLC (3:1:1, iPrOH:NH_4OH:buffer) and the fractions found to contain product were combined and lyophilised. The white powder was dissolved in 2 mL ion-exchange buffer and purified by reverse-phase HPLC (C18 column, eluting with 10% A for 20 min, then a linear gradient to 60% A over 25 min and finally a linear gradient to 100% A over 5 min.; solvent A: CH_3CN, solvent B: 25 mM NH_4HCO_3, flow rate 5.0 cm3 min-1, detecting at 220 nm) to give the title compound as a white solid in 64% yield (250 mg, 0.576 mmol)

δ_H (400 MHz, D_2O) 5.44 (1 H, t, *J* 7.0, CC*H*CH_2O), 5.24 – 5.12 (2 H, m, 2 x C=C*H*), 4.46 (2 H, t, *J* 6.0, CCHC*H*$_2$O), 2.22 – 1.95 (8 H, m, 4 x C*H*$_2$), 1.70 (3 H, s, C*H*$_3$), 1.66 (3 H, s, C*H*$_3$), 1.60 (6 H, s, 2 x C*H*$_3$); δ_P (122 MHz, D_2O) -6.13 (d, *J* 22.0), -9.85 (d, *J* 22.0); HRMS (ES$^-$, [M - H]$^-$) found 381.1235, $C_{15}H_{27}O_7P_2$ requires 381.1232

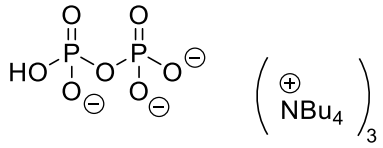
(2E,6E)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl-1,1-d2 trisammonium diphosphate 64



This compound was prepared from 1,1-$^2H_2$-farnesol (50 mg, 0.22 mmol), which was a gift from Dr Juan A. Faraldos. This compound was prepared in a manner identical to that for the diphosphate **19**, but without HPLC purification. Instead the crude pale yellow powder isolated after ion exchange was partially purified by trituration with anhydrous methanol (3 x 2 mL), and dried under N_2 flow to yield a white powder in 87% yield (83.5 mg, 1.9 mmol).

δ_H (500 MHz, D_2O) 5.47 (1 H, s, CC*H*CD_2O), 5.29 – 5.23 (1 H, m, C=C*H*), 5.23 – 5.18 (1 H, m, C=C*H*), 2.22 – 1.99 (8 H, m, 4 x CH_2), 1.74 (3 H, s, CH_3), 1.70 (3 H, s, CH_3), 1.64 (6 H, s, 2 x CH_3); δ_P (202 MHz, D_2O) -6.62; HRMS (ES^-, $[M - H]^-$) found 383.1369, $C_{15}{}^1H_{25}{}^2H_2O_7P_2$ requires 383.1358.

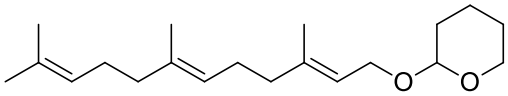
Tris(tetra-n-butyl ammonium)hydrogen diphosphate 106



A solution of disodium dihydrogen diphosphate (3.3 g, 15 mmol) in ammonium hydroxide (15 mL, 10% v/v) was passed through a column of Amberlyst® 131 cation exchange resin (Hydrogen-form) and the free acid eluted with deionised water (110 mL). The flowthrough (pH 1) was collected until the pH increased (~ pH 3-4) and immediately titrated with aqueous tetra-n-butyl ammonium hydroxide (40% w/v) to pH 7.3. The resulting clear solution (~150 mL) was lyophilised and the residue dried by azeotropic removal of water with acetonitrile (100 mL) and recrystallised from ethyl acetate to yield the title compound as hygroscopic white crystals in 48% yield (6.5 g, 7.2 mmol).

δ_H (300 MHz, $CDCl_3$) 3.44 – 3.27 (2 H, m), 1.66 (2 H, m), 1.57 – 1.37 (2 H, m), 0.98 (3 H, t, *J* 7.5); δ_C (75 MHz, $CDCl_3$) 58.88, 24.30, 19.84, 13.92; δ_P (202 MHz, $CDCl_3$) -5.95.

2-(((2E,6E)-3,7,11-Trimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran 126

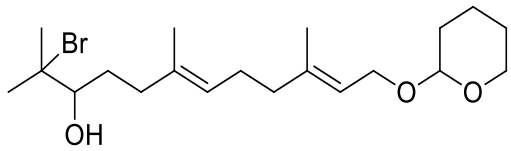


To a stirring solution of farnesol (3.93 g, 17.7 mmol) and 3,4-dihydropyran (3.7 g, 4.0 mL, 44 mmol) in DCM (100 mL) at room temperature was added pyridinium *p*-toluenesulfonate (PPTS, 150 mg, 0.6 mmol) and left stirring at room temperature for 3 hours. Water was added to the mixture and the DCM layer separated. The aqueous layer was further washed with DCM (3 x 30 mL) and the combined DCM portions washed with brine, dried ($MgSO_4$) and solvent removed in vaccuo. Purification by flash column chromatography on silica (7% EtOAc in hexane, R_f 0.34) gave the title compound as a clear free-flowing oil in 98% yield (12.4 mmol).

δ_H (300 MHz, $CDCl_3$) 5.36 (1 H, t, *J* 6.4, C*H*CH_2O), 5.18 – 5.02 (2 H, m, 2 x CC*H*CH_2CH_2), 4.65 – 4.60 (1 H, m, OC*H*O), 4.13 (2 H, ddd, *J* 19.3, 11.9, 6.9,

CHC*H*$_2$O), 3.95 – 3.45 [2 H, m, (CH$_2$)$_3$C*H*$_2$O], 2.15 – 1.92 (8 H, m, 2 x C*H*$_2$), 1.90 – 1.48 (6 H, m, (C*H*$_2$)$_3$CH$_2$O), 1.68 (6 H, s, 2 x CC*H*$_3$), 1.60 (6 H, s, 2 x CC*H*$_3$); δ_C (75 MHz, $CDCl_3$) 135.38 & 131.46 (CH$_3$*C*CH), 124.46 & 124.02 (CH$_3$C*C*H), 120.68 (*C*HCH$_2$O), 97.92 (O*C*HO), 63.79 (CH*C*H$_2$O), 62.43 [(CH$_2$)$_3$*C*H$_2$O), 39.84 & 39.78 (*C*H$_2$), 30.85 (*C*H$_2$), 26.86 & 26.42 (CH*C*H$_2$CH$_2$), 25.85 (CH$_3$), 25.64 & 19.77 (CH$_2$), 17.83, 16.57 & 16.16 (CH$_3$); HRMS (APCI$^+$, [M + H]$^+$) found 307.2622, $C_{20}H_{35}O_2$ requires 307.2637.

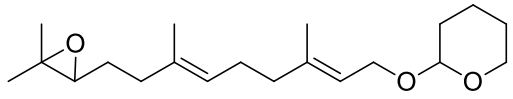
(6E,10E)-2-Bromo-2,6,10-trimethyl-12-((tetrahydro-2H-pyran-2-yl)oxy)dodeca-6,10-dien-3-ol 127



A stirring solution of **126** (5.33 g, 17.4 mmol) in THF:dH_2O (120 mL, 2:1) was cooled to 0 °C and N-bromosuccinimide (3.5 g, 19.7 mmol) added in portions over 30 minutes. The reaction was stirred at 0 °C for 2 h before it was allowed to warm to room temperature for a further 2 h. The reaction was diluted with hexane and separated with brine, the aqueous layer was further washed with hexane (3 x 30 mL), dried (Na_2SO_4) and solvent removed under reduced pressure. Purification by flash column chromatography on silica (20% EtOAc in hexane, R_f 0.35) yielded the title compound as a pale yellow oil in 70% yield (4.97 g, 12.3 mmol).

δ_H (300 MHz, $CDCl_3$) 5.36 (1 H, t, *J* 6.9, C*H*CH$_2$O), 5.20 (1 H, t, *J* 6.4, C*H*CH$_2$CH$_2$), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.30 – 4.01 (2 H, m, CHC*H*$_2$O), 4.01 – 3.93 [1 H, m, C*H*(OH)], 3.93 – 3.42 (2 H, m, [(CH$_2$)$_3$C*H*$_2$O], 2.40 – 1.70 [8 H, m, 2 x CH(C*H*$_2$)$_2$], 1.68 (3 H, s, CH$_3$), 1.59 (3 H, s, CH$_3$), 1.65 – 1.44 [6 H, m, (C*H*$_2$)$_3$CH$_2$O], 1.34 (3 H, s, CH$_3$), 1.33 (3 H, s, CH$_3$); δ_C (75 MHz, $CDCl_3$) 140.14 (CH$_3$*C*CH), 133.43 (CH$_3$*C*CH), 125.79 (*C*HCH$_2$CH$_2$), 120.80 (*C*HCH$_2$O), 97.88 (O*C*HO), 70.82 (*C*H(OH)), 63.75 (CH*C*H$_2$O), 62.41 (O*C*H$_2$(CH$_2$)$_3$), 39.58 (CHCH$_2$*C*H$_2$), 38.22 (CHCH$_2$*C*H$_2$), 32.09 (CH$_2$), 30.80 (CH$_2$), 26.62 (CH$_2$), 26.33 (*C*H$_3$CBr), 25.96 (*C*H$_3$CBr), 25.58 (CH$_2$), 19.72 (CH$_2$), 16.57 (CH$_3$), 15.96 (CH$_3$); HRMS (ES$^+$, [M + Na]$^+$) found 425.1652, $C_{20}H_{35}O_3BrNa$ requires 425.1667.

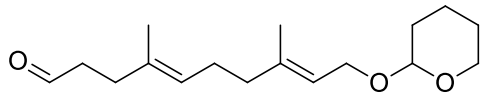
2-(((2E,6E)-9-(3,3-Dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-yl)oxy)tetrahydro-2H-pyran 128



To a solution of **127** (4.97 g, 12.3 mmol) stirring in MeOH (100 mL) at room temperature was added K_2CO_3 (3.4 g, 25 mmol) and the resulting slurry stirred for 1 h. The mixture was concentrated under reduced pressure and then diluted with hexane (50 mL) and dH_2O (50 mL). The aqueous layer separated and washed with hexane (3 x 15 mL). The combined organic fractions were washed with HCl (1 M, 15 mL) and brine, dried (Na_2SO_4) and solvent removed under reduced pressure. Purification by flash column chromatography on silica (20% EtOAc in hexane, R_f 0.5) yielded the title compound as a clear oil in 97% yield (3.94 g, 12.0 mmol).

δ_H (300 MHz, $CDCl_3$) 5.36 (1 H, t, *J* 6.3, C*H*CH_2O), 5.15 (1 H, t, *J* 6.1, C*H*CH_2CH_2), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.13 (2 H, ddd, *J* 19.2, 11.8, 6.9, CHC*H*$_2$O), 3.95 – 3.44 (2 H, m, $(CH_2)_3CH_2O$), 2.70 (1 H, t, *J* 6.3. C(O)C*H*CH_2), 2.23 – 1.97 (6 H, m, 3 x allylic C*H*$_2$), 1.67 (3 H, s, C*H*$_3$CCH), 1.92 – 1.38 (8 H, m, $(CH_2)_3CH_2O$ and C(O)CHC*H*$_2$), 1.61 (3 H, s, C*H*$_3$CCH), 1.30 (3 H, s, *C*H$_3$C(O)CH), 1.25 (3 H, s, *C*H$_3$C(O)CH); δ_C (75 MHz, $CDCl_3$) 140.18 (CH_3CCH), 134.37 (CH_3CCH), 124.55 (C*C*HCH_2), 120.62 (*C*HCH_2O), 97.86 (O*C*HO), 64.23 (C(O)*C*H), 63.67 (CH*C*H$_2$O), 62.43 (O*C*H$_2$(CH$_2$)$_3$), 39.57 (CH_2), 36.32 (CH_2), 30.72 (CH_2), 27.43 (CH_2), 26.28 (CH_2), 25.50 (CH_2), 24.93 (*C*H$_3$C(O)CH), 19.65 (CH_2), 18.78 (*C*H$_3$C(O)CH), 16.47 (CH_3), 16.14 (CH_3); HRMS (ES^+, $[M + Na]^+$) found 345.2400, $C_{20}H_{34}O_3Na$ requires 345.2406.

(4E,8E)-4,8-Dimethyl-10-((tetrahydro-2H-pyran-2-yl)oxy)deca-4,8-dienal 129

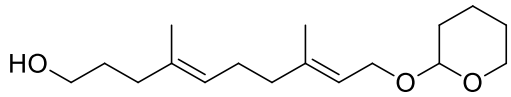


A stirring solution of **128** (3.94 g, 12.2 mmol) in THF:dH_2O (100 mL, 3:1) was cooled to 0 °C and a solution of periodic acid (3.05 g, 13.4 mmol) and sodium periodate (1.3 g, 6.1 mmol) in d.H_2O added dropwise over 15 minutes. The reaction was stirred at 0 °C for a further 30 minutes before it was allowed to warm to room temperature. After stirring at room temperature for a further 2 h the reaction was quenched with sat. $NaHCO_3$ (20 mL) and separated with hexane. The aqueous layer was washed with hexane (3 x 10 mL) and the combined organic fractions washed with HCl (1 M, 10 mL), brine, dried (Na_2SO_4) and solvent removed under reduced pressure. Purification by flash

column chromatography on silica (10% EtOAc in hexane) yielded the title compound as a pale yellow oil in 79% yield (2.69 g, 9.6 mmol).

δ_H (300 MHz, $CDCl_3$) 9.75 (1 H, t, *J* 1.9, C*H*O), 5.35 (1 H, t, *J* 6.4, C*H*CH_2O), 5.14 (1 H, t, *J* 6.1, C*H*CH_2CH_2), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.13 (2 H, ddd, *J* 19.0, 11.8, 6.6, CHC*H*$_2$O), 3.94 – 3.45 (2 H, m, $(CH_2)_3CH_2O$), 2.41 (2 H, dt, *J* 61.0, 7.7, CHOC*H*$_2$), 2.19 – 1.97 (6 H, m, 3 x CH_2), 1.91 – 1.46 (6 H, m, $(CH_2)_3CH_2O$), 1.66 (3 H, s, C*H*$_3$), 1.61 (3 H, s, C*H*$_3$); δ_C (75 MHz, $CDCl_3$) 202.46 (*C*HO), 139.72 (CH_3CCH), 133.20 (CH_3CCH), 124.93 ($CCHCH_2$), 120.81 ($CHCH_2O$), 97.73 (O*C*HO), 63.54 (CH*C*H$_2$O), 62.19 ($OCH_2(CH_2)_3$), 42.08 (CHO*C*H$_2$), 39.33 (CH_2), 31.75 (CH_2), 30.66 (CH_2), 26.08 (CH_2), 25.46 (CH_2), 19.58 (CH_2), 16.31 (CH_3), 16.05 (CH_3); HRMS (ES^+, $[M + Na]^+$) found 303.1938, $C_{17}H_{28}O_3Na$ requires 303.1936.

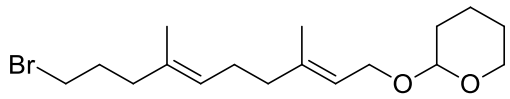
(4E,8E)-4,8-Dimethyl-10-((tetrahydro-2H-pyran-2-yl)oxy)deca-4,8-dien-1-ol 130



A stirring solution of **129** (3.28 g, 11.7 mmol) in methanol (70 mL) was cooled to 0 ºC and $NaBH_4$ (660 mg, 17.5 mmol) added and stirred for 2 h. The volume of methanol was reduced by 80% in vaccuo and then diluted with water and separated with hexane. The aqueous layer was further washed with hexane (3 x 20 mL) and the combined organic fractions washed with brine, dried (Na_2SO_4) and solvent removed under reduced pressure. Purification by flash column chromatography on silica (3% EtOAc in hexane, R_f 0.3) yielded the title compound as a clear oil in 69% yield (2.28 g, 8.07 mmol).

δ_H (300 MHz, $CDCl_3$) 5.35 (1 H, t, *J* 6.2, C*H*CH_2O), 5.13 (1 H, t, *J* 6.8, C*H*CH_2CH_2), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.13 (2 H, ddd, *J* 19.2, 11.5, 6.5, CHC*H*$_2$O), 3.94 – 3.45 (2 H, m, $(CH_2)_3CH_2O$), 3.60 (2 H, t, *J* 6.5, C*H*$_2$OH), 2.21 – 2.00 (6 H, m, 3 x CH_2), 1.90 – 1.46 (8 H, m, $(CH_2)_3CH_2O$ and C*H*$_2CH_2$OH), 1.66 (3 H, s, C*H*$_3$), 1.60 (3 H, s, C*H*$_3$); δ_C (75 MHz, $CDCl_3$) 140.05 (CH_3CCH), 134.91 (CH_3CCH), 124.64 ($CCHCH_2$), 120.92 ($CHCH_2O$), 98.09 (O*C*HO), 63.90 (CH*C*H$_2$O), 62.56 (*C*H$_2$OH), 62.51 ($OCH_2(CH_2)_3$), 39.62 (CH_2), 35.98 (CH_2), 30.81 (CH_2), 30.52 (CH_2), 26.12 (CH_2), 25.59 (CH_2), 19.78 (CH_2), 16.42 (CH_3), 15.96 (CH_3); HRMS (ES^+, $[M + Na]^+$) found 305.2098, $C_{17}H_{30}O_3Na$ requires 305.2093.

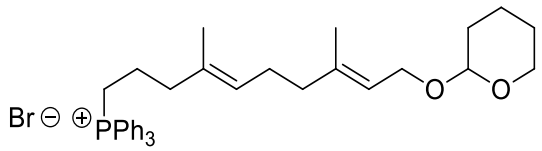
2-(((2E,6E)-10-Bromo-3,7-dimethyldeca-2,6-dien-1-yl)oxy)tetrahydro-2H-pyran 131



A stirring solution of **130** (508 mg, 1.8 mmol) and triethylamine (1.09 g, 1.5 mL, 10.8 mmol) in anhydrous THF (35 mL) was cooled to -45 ºC and methanesulfonyl chloride (412 mg, 280 μL, 3.6 mmol) added and stirred at -45 ºC for 30 minutes. The reaction was then allowed to warm to room temperature, LiBr (1.56 g, 18 mmol) was added and the reaction stirred overnight. The reaction was quenched with water and separated with hexane. The aqueous layer was further washed with hexane (3 x 10 mL) and the combined organic extracts washed with brine, dried ($MgSO_4$) and the solvent removed under reduced pressure. Purification by column chromatography on silica (7.5% EtOAc in hexane, R_f 0.28) yielded the title compound in 79% yield (483 mg, 1.4 mmol).

δ_H (300 MHz, $CDCl_3$) 5.35 (1 H, t, *J* 6.9, C*H*CH_2O), 5.15 (1 H, t, *J* 6.3, C*H*CH_2CH_2), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.13 (2 H, ddd, *J* 19.3, 11.9, 6.9, OC*H*O), 3.95 – 3.46 (2 H, m, $(CH_2)_3CH_2O$), 3.36 (2 H, t, *J* 6.7, CH_2Br), 2.21 – 2.00 (6 H, m, 3 x CH_2), 1.90 – 1.44 (8 H, m, $(CH_2)_3CH_2O$ and CH_2CH_2Br), 1.67 (3 H, s, CH_3), 1.59 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 140.16 (CH_3CCH), 133.42 (CH_3CCH), 125.49 ($CCHCH_2$), 120.86 ($CHCH_2O$), 97.97 (OCHO), 63.78 ($CHCH_2O$), 62.44 ($OCH_2(CH_2)_3$), 39.62 (CH_2), 37.93 (CH_2), 33.61 (CH_2Br), 30.96 (CH_2), 30.84 (CH_2), 26.31 (CH_2), 25.62 (CH_2), 19.76 (CH_3), 16.52 (CH_3), 15.98(CH_2); HRMS (APCI+, [M + Na]+) found 367.1231, $C_{17}H_{29}O_2BrNa$ requires 367.1249.

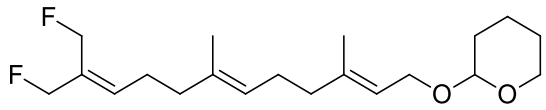
((4E,8E)-4,8-Dimethyl-10-((tetrahydro-2H-pyran-2-yl)oxy)deca-4,8-dien-1-yl)triphenylphosphonium bromide 125



A solution of **131** (409 mg, 1.19 mmol) and triphenylphosphine (340 mg, 1.3 mmol) in anhydrous acetonitrile (30 mL) was heated under reflux under an atmosphere of argon for 36 h. The mixture was allowed to cool and the acetonitrile removed under vacuum to yield a sticky brown residue. Partial purification by trituration with anhydrous diethyl ether yielded the title compound as an off-white hygroscopic solid.

δ_H (250 MHz, $CDCl_3$) 7.97 – 7.59 (15 H, m, Ar*H*), 5.45 – 5.24 (1 H, m, C=C*H*), 5.11 (1 H, t, *J* 7.0, C=C*H*), 4.58 (1 H, t, *J* 3.5, OC*H*O), 4.32 – 3.81 (2 H, m, CCHCH_2O), 3.81 – 3.42 (2 H, m, $(CH_2)_3CH_2O$), 2.29 (2 H, t, *J* 7.0, $Ph_3P^+CH_2$), 2.18 – 1.92 (6 H, m, 3 x CH_2), 1.89 – 1.38 (8 H, m, 4 x CH_2), 1.64 (3 H, s, CH_3), 1.43 (3 H, s, CH_3); δ_P (122 MHz, $CDCl_3$) 29.85; HRMS (ES+, [M]+) found 527.3071, $C_{35}H_{44}O_2P$ requires 527.3079.

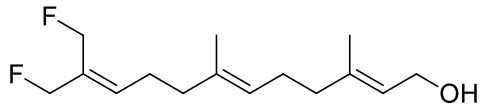
2-(((2E,6E)-12-Fluoro-11-(fluoromethyl)-3,7-dimethyldodeca-2,6,10-trien-1-yl)oxy)tetrahydro-2H-pyran 132



A stirring solution of **125** (140 mg, 0.23 mmol) in anhydrous THF (20 mL) was cooled to -78 ºC and n-BuLi (100 μL 2.5 M, 0.25 mmol) added, developing a deep yellow colour as the ylide is formed. Difluoroacetone (23 mg, 19 μL, 0.25 mmol) was added dropwise and stirred at -78 ºC for 2 h before being allowed to warm to -20 ºC and quenched with water and Et_2O (10 mL, 1:1). The aqueous layer was separated and further washed with Et_2O (3 x 5 mL) and the combined ethereal extracts washed with brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification by column chromatography on silica (20% EtOAc in hexane, R_f 0.5) yielded the title compound in 47% yield (38 mg, 0.11 mmol).

δ_H (300 MHz, $CDCl_3$) 5.89 – 5.79 (1 H, m, CH_2FCCH), 5.40 – 5.31 (1 H, m, $CCHCH_2OH$), 5.16 – 5.09 (1 H, m, $CCHCH_2$), 4.99 (2 H, d, *J* 47.6, $CHCCH_2F$), 4.87 (2 H, d, *J* 47.6, $CHCCH_2F$), 4.62 (1 H, t, *J* 3.5, OC*H*O), 4.13 (2 H, ddd, *J* 68.0, 11.9, 7.2, CHCH_2O), 3.95 – 3.46 (2 H, m, $(CH_2)_3CH_2O$), 2.42 – 1.96 (6 H, m, 2 x CH_2CH_2), 1.96 – 1.46 (6 H, m, $(CH_2)_3CH_2O$), 1.69 (3 H, s, CH_3), 1.60 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 140.19 (CH_3CCH), 133.89 (CH_3CCH), 125.23 ($CCHCH_2$), 120.80 ($CHCH_2O$), 98.00 (OCHO), 84.91 (d, *J* 168.0, CH_2F), 77.63 (d, *J* 156.8, CH_2F), 63.80 ($CHCH_2O$), 62.46 ($OCH_2(CH_2)_3$), 39.62 (CH_2), 30.84 (CH_2), 26.34 (CH_2), 25.62 (CH_2), 19.77 (CH_2), 16.56 (CH_3), 16.07 (CH_3); δ_F (283 MHz, $CDCl_3$) -211.84 (t, *J* 47.6), -216.85 (t, *J* 47.6); HRMS (ES^+, $[M + Na]^+$) found 365.2257, $C_{20}H_{32}O_2F_2Na$ requires 365.2268.

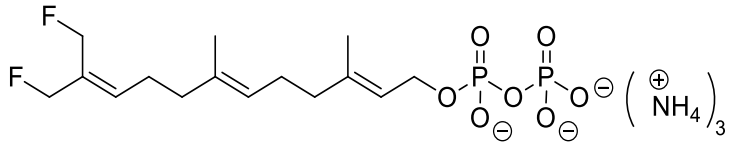
(2E,6E)-12-Fluoro-11-(fluoromethyl)-3,7-dimethyldodeca-2,6,10-trien-1-ol 112



To a stirring solution of **132** (38 mg, 0.11 mmol) in methanol (10 mL) was added *p*-toluenesulfonic acid (cat.) and stirred at room temperature for 2 h. The volume of methanol was reduced by 80 % in vaccuo, diluted with sat. $NaHCO_3$ and separated with hexane. The aqueous layer was further washed with hexane (3 x 5 mL) and the combined organic fractions washed with brine, dried (anhydrous $MgSO_4$) and solvent removed under reduced pressure. Purification by column chromatography on silica (20% EtOAc in hexane, R_f 0.29) yielded the title compound in 91% yield (28 mg, 0.10 mmol).

δ_H (300 MHz, $CDCl_3$) 5.89 – 5.80 (1 H, m, CH_2FCCH), 5.45 – 5.37 (1 H, m, $CCHCH_2OH$), 5.19 – 5.09 (1 H, m, $CCHCH_2CH_2$), 4.99 (2 H, d, *J* 47.6, CH_2F), 4.87 (2 H, d, *J* 47.6, CH_2F), 4.16 (2 H, d, *J* 6.9, CH_2OH), 2.33 – 1.96 (8 H, m, 2 x CH_2CH_2), 1.68 (3 H, s, CH_3), 1.60 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 139.69 (CH_3CCH), 133.99 (CH_3CCH), 125.11 ($CCHCH_2CH_2$), 123.55 ($CCHCH_2OH$), 84.84 (d *J* 164.3, CH_2F), 77.63 (d, *J* 168.0, CH_2F), 59.53 (CH_2OH), 39.51 (CH_2), 38.96 (CH_2), 26.31 (CH_2), 16.41 (CH_3), 16.06 (CH_3); δ_F (283 MHz, $CDCl_3$) -214.24 (t, *J* 47.4), -219.22 (t, *J* 47.4); HRMS ($APCI^+$, $[M + Na]^+$) found 281.1683, $C_{15}H_{24}OF_2Na$ requires 281.1693.

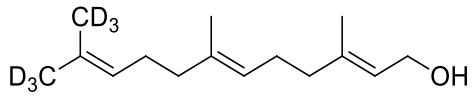
(2E,6E)-12-Fluoro-11-(fluoromethyl)-3,7-dimethyldodeca-2,6,10-trien-1-yl tris-ammonium diphosphate 73



This compound was prepared and purified from alcohol **112** in the same manner as diphosphate **19**, to give the title compound as a white solid in 59% yield (27.7 mg, 0.059 mmol).

δ_H (500 MHz, D_2O) 6.08 – 5.98 (1 H, m, CH_2FCCH), 5.46 (1H, t, *J* = 7.3 Hz, $CCHCH_2OH$), 5.23 (1 H, t, J = 6.9, $CCHCH_2CH_2$), 5.10 (2 H, d, *J* 47.6, CH_2F), 4.95 (2 H, d, *J* 47.6, CH_2F), 4.47 (2 H, t, *J* 6.6, CH_2OH), 2.40 – 2.03 (8 H, m, 2 x CH_2CH_2), 1.72 (3 H, s, CH_3), 1.62 (3 H, s, CH_3); δ_F (283 MHz, D_2O) -207.11 (t, *J* 47.4), -212.67 (t, *J* 47.5); δ_P (202 MHz, D2O) -6.17 (d, *J* 19.0), -10.18 (d, *J* 21). HRMS (ES^-, $[M - H]^-$) found 417.1037, $C_{15}H_{25}O_7F_2P_2$ requires 417.1044.

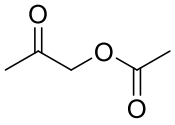
(2E,6E)-3,7-Dimethyl-11-(methyl-d3)- 2,6,10-Dodecatrien-12,12,12-d3-1-ol 133



A stirring solution of **125** (420 mg, 0.69 mmol) in anhydrous THF (10 mL) was cooled to -78 °C and n-BuLi (280 μL, 0.69 mmol) was added dropwise, a bright orange colour developed. The solution was allowed to warm to -10 °C and d6-acetone (51 μL, 0.69 mmol) was added dropwise, the reaction was maintained at -10 °C for 1 h and the colour of the ylid was slowly lost. The reaction was allowed to stir for a further hour at room temperature before quenching with d.H_2O (5 mL) and diluting with methanol (10 mL). A catalytic amount of para-toluenesulfonic acid was added and stirred vigorously for 30 minutes before quenching with $NaHCO_3$ (5 mL). The mixture was separated with Et_2O (10 mL) and the aqueous layer further washed with Et_2O (3 x 5 mL), the combined ethereal extracts were washed with water, brine, dried (MgSO4) and solvent removed under reduced pressure. Purification by column chromatography on silica (20% EtOAc in hexane) yielded the title compound in 41% yield (88 mg, 0.28 mmol).

δ_H (300 MHz, $CDCl_3$) 5.42 (1 H, t, *J* 5.5 Hz, C*H*CH_2O), 5.18 – 5.03 (2 H, m 2 x C*H*CH_2CH_2), 4.15 (2 H, d, *J* 7.0 Hz, CHC*H*$_2$O), 2.26 – 1.91 (8 H, m, 2 x CH(C*H*$_2$)$_2$), 1.68 (3 H, s, C*H*$_3$), 1.60 (3H, s, C*H*$_3$); δ_C (151 MHz, CDCl3) 140.02, 135.54 & 128.38 (CH_3*C*CH), 124.50 & 123.92 (CH_3C*C*H), 123.50 (*C*HCH_2O), 59.59 (CH*C*H_2O), 39.86, 39.70, 26.86 & 26.46(CH*C*H_2CH_2), 16.44 & 16.16 (CH_3); HRMS (EI+, [M]+) found 228.2354, $C_{15}{}^1H_{20}{}^2H_6O$ requires 228.2360.

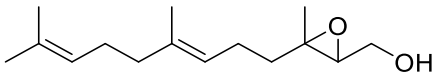
Acetoxyacetone 136



A solution of hydroxyacetone (500 mg, 6.75 mmol) and acetic anhydride (1.5 mL, 1.62 g, 15.9 mmol) in pyridine (3 mL) was stirred for 16 h, before it was diluted with dichloromethane and acetone (20 mL, 1:1), washed with water (20 mL) and separated. The aqueous layer was further washed with DCM (3 x 10 mL) and the combined organic extracts washed with sat. $CuSO_4$ (20 mL), $NaHCO_3$ (3 x 10 mL), brine, dried (Na_2SO_4) and solvent removed under reduced pressure. Purification by column chromatography on silica (30% EtOAc in hexane, R_f 0.34) yielded the title compound in 89% yield (700 mg, 6.03 mmol).

δ_H (300 MHz, $CDCl_3$) 4.65 (2 H, s, C(O)CH_2O), 2.17 (3 H, s, CH_3), 2.16 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 201.70 ($CH_3C(O)O$), 170.39 (*C*(O)CH_2), 68.44 (C(O)*C*H_2O), 26.22 (*C*H_3), 20.64 (*C*H_3); HRMS (EI^+, $[M]^+$) found 116.0472, $C_5H_8O_3$ requires 116.0473.

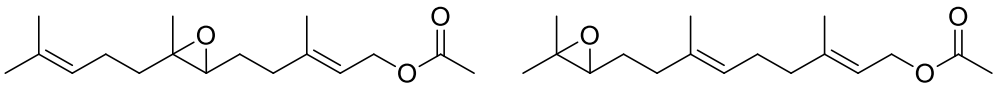
(E)-(3-(4,8-Dimethylnona-3,7-dien-1-yl)-3-methyloxiran-2-yl)methanol 137



A stirring solution of farnesol (253 mg, 1.14 mmol) in anhydrous toluene (10 mL) was cooled to 10 ºC and a catalytic amount of vanadyl acetoacetate added, and the solution turned a pale brown. *Tert*-butyl hydroperoxide (230 μL, 1.25 mmol, ~5.5 M in decanes) was added dropwise over 10 minutes, at which the solution turned a deep red. The reaction was stirred at 10 ºC for 1 h 45, over which time the colour faded to a pale yellow, and was quenched with NaOH/Brine (10 mL, 1:1) stirred for 10 minutes and separated. The aqueous layer was washed with EtOAc (3 x 10 mL) and the combined organic extracts washed with brine, dried ($MgSO_4$) and solvent removed under vacuum. Purification on a silica plug (20% EtOAc in hexane, R_f 0.2) yielded the title compound in 95% yield (257 mg, 1.08 mmol).

δ_H (300 MHz, $CDCl_3$) 5.09 (2 H, dt, *J* 6.7, 5.7), 3.89 – 3.63 (2 H, m), 2.98 (1 H, dd, *J* 6.6, 4.3), 2.21 – 1.89 (6 H, m), 1.67 (3 H, s), 1.60 (6 H, s), 1.80 – 1.39 (2 H, m, *J* 16.5, 13.7, 8.0), 1.30 (3 H, s); δ_C (75 MHz, $CDCl_3$) 135.97, 131.63, 124.33, 123.33, 62.99, 61.62, 61.33, 39.79, 38.65, 26.76, 25.84, 23.74, 17.84, 16.93, 16.13. HRMS (ES^-, $[M - H_2O - H]^-$) found 219.1752, $C_{15}H_{23}O$ requires 219.1749.

(E)-3-Methyl-5-(3-methyl-3-(4-methylpent-3-en-1-yl)oxiran-2-yl)pent-2-en-1-yl acetate 143 & (2E,6E)-9-(3,3-dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-yl acetate 142



A stirring solution of farnesyl acetate (608 mg, 2.3 mmol) in DCM (40 mL) was cooled to 0 ºC and 3-chloroperbenzoic acid (430 mg, 2.5 mmol) added. The solution was allowed to warm to room temperature and stirred for 1 h, before quenching with NaOH (20 mL, 2.5 M) and diluting with d.H_2O. The DCM was separated and the aqueous layer further washed with DCM (3 x 5 mL), and the combined organic extracts washed with brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification by flash column

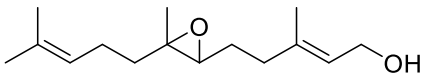
chromatography on silica (5% EtOAc in hexane) yielded the mixed mono-epoxides in 75% yield (477 mg, 1.7 mmol) along with recovered starting material (13%, 80 mg, 0.3 mmol) and diepoxide (<5%).

δ_H (300 MHz, $CDCl_3$) 5.47 – 5.28 (1 H, m, CC*H*CH_2), 5.14 (0.58 H, t, *J* 6.3, CC*H*CH_2), 5.07 (0.42 H, t, *J* 7.2, $(CH_3)_2$CC*H*), 4.58 (2 H, d, *J* 7.1, CCHC*H*$_2$O), 2.70 (1 H, t, *J* 6.3, C(O)C*H*), 2.05 (3 H, s, *J* 0.6, OC(O)C*H*$_3$).

(E)-3-Methyl-5-(3-methyl-3-(4-methylpent-3-en-1-yl)oxiran-2-yl)pent-2-en-1-ol 138 & (2E,6E)-9-(3,3-dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-ol 139

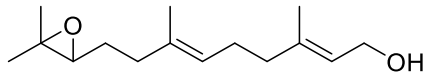
To a stirring slurry of K_2CO_3 (470 mg, 3.4 mmol) in methanol (10 mL) was added mixed epoxides (DG021, 1.7 mmol) and stirred for 1 h at room temperature. The reaction was concentrated by 80 % under reduced pressure, diluted with d.H_2O (10 mL) and separated with hexane (10 mL). The aqueous layer was further washed with hexane (3 x 5 mL) and the combined organic extracts washed with brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification by automated chromatography (Biotage® Isolera Four) on silica (100 g Biotage® SNAP-Ultra cartridge) using Et_2O in hexane (10 – 50% over 10 CV, 50 – 80% over 10 CV) yielded the purified title compounds in 21% (0.37 mmol) and 33% (0.55 mmol) yields respectively as well as recovering 20% mixed epoxides (80 mg, 0.34 mmol).

(E)-3-Methyl-5-(3-methyl-3-(4-methylpent-3-en-1-yl)oxiran-2-yl)pent-2-en-1-ol) 138



δ_H (300 MHz, $CDCl_3$) 5.49 – 5.40 (1 H, m, CC*H*CH_2OH), 5.11 – 5.00 (1 H, m, $(CH_3)_2$CC*H*), 4.15 (2 H, d, *J* 6.8, CCHC*H*$_2$OH), 2.70 (1 H, t, *J* 6.2, C(O)C*H*), 2.30 – 1.99 (4 H, m, 2 x CH_2), 1.69 (3 H, s, CH_3), 1.67 (3 H, s, CH_3), 1.60 (3 H, s, CH_3), 1.67 – 1.28 (4 H, m, 2 x CH_2), 1.25 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 138.87 (CH_3*C*CH), 132.05 (CH_3*C*CH), 123.96 (CC*H*CH_2OH), 123.74 ($(CH_3)_2$CC*H*), 63.37 (C(O)*C*H), 60.95 (CH_3*C*(O)CH), 59.41 (CCH*C*H_2OH), 38.89 (CH_2), 36.32 (CH_2), 27.14 (CH_2), 25.84 (CH_3), 23.99 (CH_2), 17.80 (CH_3), 16.62 (CH_3), 16.41 (CH_3); HRMS ($APCI^+$, $[M + Na]^+$) found 261.1833, $C_{15}H_{26}O_2Na$ requires 261.1831.

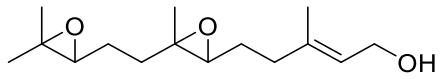
(2E,6E)-9-(3,3-Dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-ol 139



In addition to the method described above, this compound was also prepared by treatment of bromohydrin **170** (327 mg, 0.91 mmol) with K_2CO_3 (248 mg, 1.8 mmol) in methanol in the same manner as for **128**. Purification by flash column chromatography on a silica plug (30% EtOAc in hexane) yielded the pure product in 91% yield (198 mg, 0.83 mmol)

δ_H (300 MHz, $CDCl_3$) 5.45 – 5.35 (1 H, m, CC*H*CH_2OH), 5.23 – 5.09 (1 H, m, CC*H*CH_2CH_2), 4.14 (2 H, d, *J* 6.8, CCHC*H*$_2$OH), 2.70 (1 H, t, *J* 6.2, C(O)C*H*), 2.20 – 1.98 (6 H, m, 3 x CH_2), 1.67 (3 H, s, CH_3), 1.72 – 1.57 (2 H, m, CH_2), 1.61 (3 H, s, CH_3), 1.30 (3 H, s, CH_3), 1.25 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 139.51(CH_3*C*CH), 134.47 (CH_3*C*CH), 124.64 (C*C*HCH_2CH_2), 123.70 (C*C*HCH_2OH), 64.31 (C(O)*C*H), 59.47 (CCH*C*H$_2$OH), 58.57 (*C*(O)CH), 39.51 (CH_2), 36.45 (CH_2), 27.36 (CH_2), 26.23 (CH_2), 25.00 (CH_3), 18.91 (CH_3), 16.36 (CH_3), 16.10 (CH_3); HRMS (APCI$^+$, [M + Na]$^+$) found 261.1828, $C_{15}H_{26}O_2Na$ requires 261.1831.

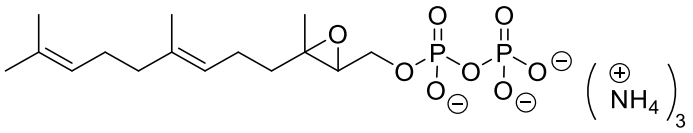
(E)-5-(3-(2-(3,3-Dimethyloxiran-2-yl)ethyl)-3-methyloxiran-2-yl)-3-methylpent-2-en-1-ol 137



To stirring solution of farnesyl acetate (200 mg, 0.9 mmol) in dichloromethane (20 mL) at 0 °C was added mCPBA (520 mg, 2.25 mmol) and stirred for 1 h. The reaction was quenched with NaOH (10 mL, 1 M) and stirred for 5 minutes before the aqueous layer was separated and washed with DCM (3 x 5 mL). The combined organic fractions were washed with brine, dried ($MgSO_4$) and solvent removed under reduced pressure. The resulting residue was taken up in MeOH (20 mL), K_2CO_3 (250 mg, 0.8 mmol) added and the resulting slurry stirred for 1 h. The volume of MeOH was reduced by 80 % and the remaining solution diluted with H_2O/EtOAc (20 mL, 1:1) and separated. The aqueous layer was washed with EtOAc (3 x 10 mL) and the combined organic fractions washed with HCl (10 mL, 1 M), brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification on a silica plug (40% EtOAc in hexane, R_f 0.3, 50% EtOAc in hexane) yielded the title compound in 65% yield (148 mg, 0.58 mmol).

δ_H (300 MHz, $CDCl_3$) 5.53 – 5.39 (1 H, m, CC*H*CH_2OH), 4.16 (2 H, d, *J* 7.0, CCHC*H*$_2$OH), 2.78 – 2.67 (2 H, m, 2 x C(O)C*H*), 2.31 – 2.07 (2 H, m, C*H*$_2$), 1.69 (3 H, s, C*H*$_3$), 1.83 – 1.54 (6 H, m, 3 x C*H*$_2$), 1.31 (3 H, s, C*H*$_3$), 1.28 (3 H, s, C*H*$_3$), 1.26 (3 H, s, C*H*$_3$); δ_C (75 MHz, $CDCl_3$) 138.78 (*C*CHCH$_2$OH), 124.21 (C*C*HCH$_2$OH), 64.16 & 62.84 (C(O)*C*H), 59.44 (CCH*C*H$_2$OH), 35.33, 27.02, 24.98, 24.91 (*C*H$_2$), 24.70, 18.79, 16.89 & 16.40 (*C*H$_3$); HRMS (APCI$^+$, [M + Na]$^+$) found 319.1883, $C_{17}H_{28}O_4Na$ requires 319.1885.

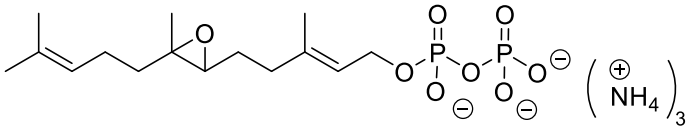
(E)-(3-(4,8-Dimethylnona-3,7-dien-1-yl)-3-methyloxiran-2-yl)methyl trisammonium diphosphate 145



Bis-triethylammonium phosphate (TEAP) was prepared immediately prior to use by the dropwise addition of a solution of phosphoric acid in acetonitrile to a stirring solution of triethylamine in acetonitrile. TEAP was then added in three portions (3 x 1 mL) at 5 minute intervals to a solution of **137** (150 mg, 0.63 mmol) in trichloroacetonitrile (1 mL) at 37 °C. The reaction was incubated at 37 °C for a further 5 minutes after the final addition and the entire reaction mixture applied to a silica column and washed onto the column with isopropanol. The mobile phase (isopropanol : conc. NH_4OH : H_2O, 6 : 2.5 : 0.5) was then begun and fractions collected and analysed by TLC (isopropanol : conc. NH_4OH : H_2O, 6 : 3 : 1, visualised with basic $KMnO_4$), and those containing the diphosphate (R_f 0.26) were combined, ammonia and isopropanol removed under vacuum and the resulting aqueous solution diluted to 60 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound as a light fluffy powder in 46 % yield (129 mg, 0.29 mmol).

δ_H (500 MHz, D_2O) 5.15 – 5.07 (2 H, m, 2 x C=C*H*), 4.25 – 3.78 (2 H, m, C(O)CHC*H*$_2$O), 3.21 (1 H, dd, *J* 7.5, 3.5, C(O)C*H*CH_2O), 2.10 – 1.89 (6 H, m, 3 x CH_2), 1.59 (3 H, s, C*H*$_3$), 1.54 (3 H, s, C*H*$_3$), 1.53 (3 H, s, C*H*$_3$), 1.27 (3 H, s, C*H*$_3$); δ_C (126 MHz, D_2O) 136.90 (C=C), 133.53 (C=C), 124.50 (C=*C*H), 123.71 (C=*C*H), 64.47 (C(O)CH*C*H$_2$O), 63.77 (C(O)*C*HCH$_2$O), 38.84 (*C*H$_2$), 37.67 (*C*H$_2$), 25.84 (*C*H$_2$), 24.89 (*C*H$_3$), 23.02 (*C*H$_2$), 17.02 (*C*H$_3$), 15.72 (*C*H$_3$), 15.23 (*C*H$_3$); δ_P (202 MHz, D_2O) -6.24, -10.32; HRMS (ES$^-$, [M – H]$^-$) found 397.1179, $C_{15}H_{27}O_8P_2$ requires 397.1181.

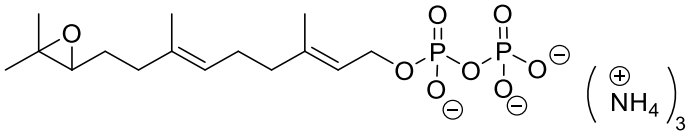
(E)-3-Methyl-5-(3-methyl-3-(4-methylpent-3-en-1-yl)oxiran-2-yl)pent-2-en-1-yl trisammonium diphosphate 146



This compound was prepared and purified from alcohol **138** (42 mg, 0.18 mmol) in the same manner as diphosphate **19**, to give the title compound as a white solid in 58% yield (46 mg, 0.10 mmol).

δ_H (300 MHz, D_2O) 5.49 – 5.40 (1 H, m, CC*H*CH_2OH), 5.15 – 5.08 (1 H, m, CC*H*CH_2CH_2), 4.46 – 4.36 (2 H, m, CCHC*H*$_2$OH), 3.00 – 2.93 (1 H, m, C(O)C*H*), 2.21 – 2.00 (6 H, m, 3 x CH_2), 1.67 (3 H, s, CH_3), 1.63 (3 H, s, CH_3), 1.57 (3 H, s, CH_3), 1.25 (3 H, s, CH_3); δ_P (202 MHz, D_2O) -6.48 (d, *J* 20.8), -10.27 (d, *J* 17.0); HRMS (ES^-, $[M - H]^-$) found 397.1199, $C_{15}H_{27}O_8P_2$ requires 397.1181.

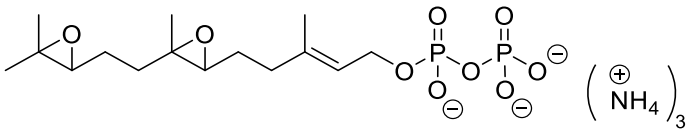
(2E,6E)-9-(3,3-Dimethyloxiran-2-yl)-3,7-dimethylnona-2,6-dien-1-yl trisammonium diphosphate 147



This compound was prepared and purified from alcohol **139** (65 mg, 0.28 mmol) in the same manner as diphosphate **19**, to give the title compound as a white solid in 66% yield (81 mg, 0.18 mmol).

δ_H (300 MHz, D_2O) 5.29 (1 H, t, *J* 7.0, CC*H*CH_2OH), 5.10 (1 H, t, *J* 6.0, CC*H*CH_2CH_2), 4.29 (2 H, t, *J* 6.5, CCHC*H*$_2$OH), 2.84 (1 H, t, *J* 6.5, C(O)C*H*), 2.06 – 1.88 (6 H, m, 3 x CH_2), 1.60 – 1.49 (2 H, m, CH_2), 1.54 (3 H, s, CH_3), 1.48 (3 H, s, CH_3), 1.14 (3 H, s, CH_3), 1.12 (3 H, s, CH_3). δ_P (122 MHz, D_2O) -6.14, -9.83. HRMS (ES-, [M – H]-) found 397.1174, $C_{15}H_{27}O_8P_2$ requires 397.1181.

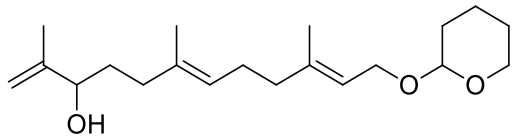
(E)-5-(3-(2-(3,3-Dimethyloxiran-2-yl)ethyl)-3-methyloxiran-2-yl)-3-methylpent-2-en-1-yl trisammonium diphosphate 148



This compound was prepared and purified from alcohol **140** (100 mg, 0.4 mmol) in the same manner as diphosphate **137**, to give the title compound as a white solid in 30% yield (55 mg, 0.12 mmol).

δ_H (500 MHz, D_2O) 5.42 (1 H, t, *J* 6.5, CC*H*CH_2O), 4.40 (2 H, t, *J* 6.0, CCHC*H*$_2$O), 3.02 – 2.89 (2 H, m, 2 x C(O)C*H*), 2.14 (1 H, t, *J* 7.5, C*H*$_2$), 1.65 (3 H, s, C*H*$_3$), 1.77 – 1.53 (6 H, m, 3 x C*H*$_2$), 1.25 (3 H, s, C*H*$_3$), 1.24 (6 H, s, 2 x C*H*$_3$); δ_P (202 MHz, D_2O) -10.55. HRMS (ES^-, $[M - H]^-$) found 413.1126, $C_{15}H_{27}O_9P_2$ requires 413.1130.

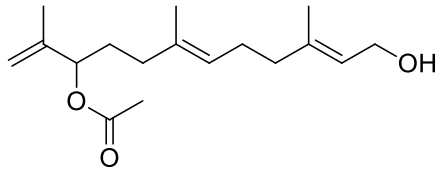
(6E,10E)-2,6,10-Trimethyl-12-((tetrahydro-2H-pyran-2-yl)oxy)dodeca-1,6,10-trien-3-ol
165



A solution of **128** (390 mg, 1.2 mmol) and aluminium isopropoxide (450 mg, 2.4 mmol) in anhydrous toluene (10 mL) was heated to 140 °C in a sealed tube for 16 h. After being allowed to cool HCl:EtOAc (5 mL, 1:1, 1 M HCl) was added and stirred for 1 h. The mixture was separated and the aqueous layer washed with EtOAc (3 x 5 mL) and the combined organic extracts washed with brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification by column chromatography on silica (20% EtOAc in hexane, R_f 0.32) yielded the title compound in 65% yield (253 mg, 0.78 mmol).

δ_H (300 MHz, $CDCl_3$) 5.42 – 5.28 (1 H, m, CC*H*CH_2O), 5.15 (1 H, t, *J* 6.5, CC*H*$(CH_2)_2$), 4.94 (1 H, dd, *J* 1.5, 1.0, C*H*H=C(CH_3)CHOH), 4.82 (1 H, dd, *J* 3.0, 1.5, C*H*H=C(CH_3)CHOH), 4.67 – 4.57 (1 H, m, OC*H*O), 4.31 – 3.95 (3 H, m, CCHC*H*$_2$O & C=CC*H*OH), 3.95 – 3.44 (2 H, m, $(CH_2)_3$C*H*$_2$O), 2.08 (6 H, m, 3 x CH_2), 1.72 (3 H, s, CH_3), 1.89 – 1.45 (9 H, m,), 1.67 (3 H, s, CH_3), 1.61 (3 H, s, CH_3); δ_C 147.62, 139.89 (C=C), 134.88 (C=C), 124.52 (C*C*HCH_2), 120.86 (C*C*HCH_2), 97.84 (C=C*C*HOH), 63.71 (CCH*C*H$_2$O), 62.30 (O*C*HO), 39.53, 35.66, 33.09, 30.66, 26.12, 26.07, 25.50 & 19.58 (CH_2), 17.61, 16.30 & 15.95 (CH_3); HRMS ($APCI^+$, $[M + Na]^+$) found 345.2401, $C_{20}H_{34}O_3Na$ requires 345.2406.

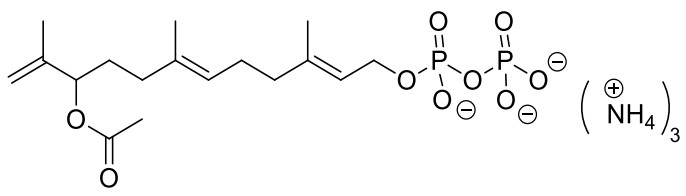
(6E,10E)-12-Hydroxy-2,6,10-trimethyldodeca-1,6,10-trien-3-yl acetate 167



To a stirring solution of **128** (253 mg, 0.78 mmol) in DCM (5 mL) was added acetic anhydride (0.408 g, 0.38 mL, 4 mmol) and pyridine (0.316 g, 0.32 mL, 4 mmol) and stirred at room temperature for 3 days until judged complete by TLC. The reaction was diluted with hexane and separated with water. The aqueous layer was further washed with hexane and the combined organic fractions dried (anhydrous $MgSO_4$) and solvent removed under reduced pressure. The resulting oil was taken up in methanol and a catalytic amount of *p*-toluenesulfonic acid added and stirred at room temperature for 1h. The reaction was quenched with $NaHCO_3$ and separated with hexane, the aqueous layer was further washed with hexane and the combined organic extracts washed with brine, dried (anhydrous $MgSO_4$) and solvent removed under reduced pressure. Purification by flash column chromatography on silica (30% EtOAc in hexane, R_f 0.3) yielded the title compound as a yellow oil in a 65% yield (142 mg, 0.51 mmol) over two steps.

δ_H (300 MHz, $CDCl_3$) 5.46 – 5.35 (1 H, m, CC*H*CH_2O), 5.18 – 5.04 (2 H, m, CC*H*$(CH_2)_2$ & C=CC*H*OAc), 4.99 – 4.91 (1 H, m, C*H*H=C(CH_3)CHOH), 4.91 – 4.86 (1 H, m, C*H*H=C(CH_3)CHOH), 4.15 (2 H, d, *J* 7.0, CCHC*H*$_2$O), 2.06 (3 H, s, C*H*$_3$C(O)O), 2.18 – 1.87 (6 H, m, 3 x CH_2), 1.82 – 1.62 (2 H, m, CH_2), 1.71 (3 H, s, CH_3), 1.67 (3 H, s, CH_3), 1.60 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 170.64 (CH_3*C*(O)), 143.23 (C=C), 139.41 (C=C), 134.29 (C=C), 124.69 (C*C*H(CH_2)), 123.75 (C*C*HCH_2OH), 112.98 (*C*HH=C(CH_3)CHOH), 77.10 (C=C*C*HOAc), 59.55 (CCH*C*H_2OH), 39.45, 35.34, 30.88 & 26.18 (*C*H_2), 21.40 (*C*H_3C(O)), 18.19, 16.38 & 16.13 (*C*H_3); HRMS (APCI$^+$, $[M + Na]^+$) found 303.1925, $C_{17}H_{28}O_3Na$ requires 303.1936.

(2E,6E)-10-Acetoxy-3,7,11-trimethyldodeca-2,6,11-trien-1-yl trisammonium diphosphate 169

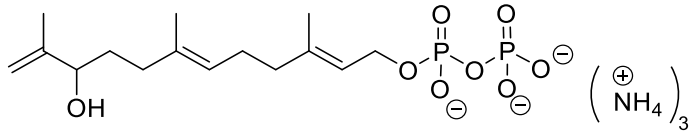


This compound was prepared and purified from alcohol **167** (79 mg, 0.26

mmol) in the same manner as diphosphate **19**, to give the title compound as a white solid in 74% yield (93 mg, 0.19 mmol).

δ_H (300 MHz, D_2O) 5.31 (1 H, t, *J* 7.5, CC*H*CH_2O), 5.05 (1 H, t, *J* 7.0, CC*H*$(CH_2)_2$), 4.89 (1 H, t, *J* 6.5, C*H*OAc), 4.79 (2 H, bs, (CH3)C=C*HH*), 4.31 (2 H, t, *J* 6.5, CCHC*H*$_2$O), 1.97 (3 H, s, C*H*$_3$), 2.09 – 1.84 (8 H, m, 4 x C*H*$_2$), 1.71 – 1.59 (2 H, m, C*H*$_2$), 1.56 (6 H, s, 2 x C*H*$_3$), 1.47 (3 H, s, C*H*$_3$); δ_P (202 MHz, D_2O) -6.53 (d, *J* 22.0), -10.25 (d, *J* 22.0).; HRMS (ES^-, $[M - H]^-$) found 439.1298, $C_{17}H_{29}O_9P_2$ requires 439.1287.

(2E,6E)-10-Hydroxy-3,7,11-trimethyldodeca-2,6,11-trien-1-yl trisammonium diphosphate 162



To a stirring solution of **163** (108 mg, 0.45 mmol) and trimethylamine (367 mg, 500 μL, 0.9 mmol) in THF (10 mL) was added *para*-Toluenesulfonyl chloride (86 mg, 0.45 mmol) and the reaction left to stir in darkness for 16 h before it was quenched with d.H_2O and separated with Et_2O. The aqueous layer was further washed with Et_2O and the combined ethereal extracts washed with $CuSO_4$, $NaHCO_3$, and brine, dried ($MgSO_4$) and solvent removed under reduced pressure. The tosylate was dried under high-vaccuum for 2 h before being used without further purification

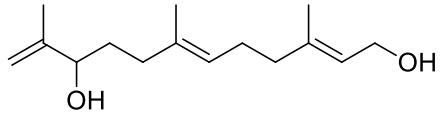
The tosylate intermediate was treated with tris(tetrabutylammonium)hydrogen diphosphate (1.53 mmol, 1.38 g) in anhydrous acetonitrile, and further treated and purified in the same manner as diphosphate **19** to yield the title compound as a white powder in 54% yield (95 mg, 0.24 mmol).

This compound was also prepared by treatment of **169** (50 mg, 0.10 mmol) with NH_4OH (1 mM) in D_2O (3 mL), and followed by 1H-NMR of a 500 μL sample at 12 hourly intervals. Upon complete disappearance of the acetate signal (2.1 ppm, 36 h) the reaction was diluted with d.H_2O and lyophilised to yield the title compound as a white powder in 90% yield (41 mg, 0.09 mmol).

δ_H (300 MHz, D_2O) 5.32 (1 H, t, *J* 6.5, CC*H*CH_2O), 5.09 (1 H, t, *J* 6.0, CC*H*$(CH_2)_2$), 4.79 (1 H, s, C*H*H=C(CH_3)CHOH), 4.76 (1 H, s, C*H*H=C(CH_3)CHOH), 4.33 (2 H, t, *J* 6.5, CCHC*H*$_2$O), 3.93 (1 H, t, *J* 7.0, C=CC*H*OH), 2.09 – 1.80 (6 H, m, 3 x CH_2), 1.60 – 1.48 (2 H, m, CH_2), 1.58 (3 H,

s, CH_3), 1.55 (3 H, s, CH_3), 1.49 (3 H, s, CH_3); δ_P (202 MHz, D_2O) -8.93, -10.35. HRMS (ES^-, $[M - H]^-$) found 397.1193, $C_{15}H_{27}O_8P_2$ requires 397.1181.

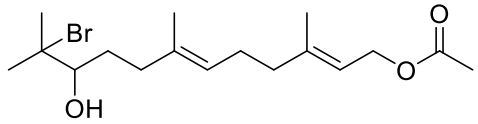
(2E,6E)-3,7,11-Trimethyldodeca-2,6,11-triene-1,10-diol 163



This compound was prepared from alcohol **139** (113 mg, 0.48 mmol) in the same manner as **165**. Purification by column chromatography on silica (40% EtOAc in hexane) yielded the title compound in 65% yield (74 mg, 0.31 mmol).

δ_H (300 MHz, $CDCl_3$) 5.49 – 5.33 (1 H, m, CC*H*CH_2O), 5.24 – 5.10 (1 H, m, CC*H*$(CH_2)_2$), 4.97 – 4.88 (1 H, m, C*H*H=C(CH_3)CHOH), 4.89 – 4.79 (1 H, m, C*H*H=C(CH_3)CHOH), 4.13 (2 H, d, *J* 7.0, CCHC*H*$_2$O), 4.04 (1 H, dd, *J* 7.5, 5.0, C=CC*H*OH), 2.24 – 2.00 (6 H, m, 3 x CH_2), 1.72 (3 H, s, CH_3), 1.66 (3 H, s, CH_3), 1.69 – 1.60 (2 H, m, CH_2), 1.63 (3 H, s, CH_3); δ_C (151 MHz, $CDCl_3$) 147.59 (C=C), 139.24 (C=C), 135.53 (C=C), 124.92 (C*C*H(CH_2)), 124.16 (C*C*HCH_2OH), 110.96 (*C*HH=C(CH_3)CHOH), 75.71(C=C*C*HOH), 59.51 (CCH*C*H_2OH), 39.50, 36.08, 32.89 & 25.99 (*C*H_2), 17.97, 16.14 & 16.04 (*C*H_3; HRMS (ES-, [M – H2O - H]-) found 219.1747, $C_{15}H_{23}O$ requires 219.1749.

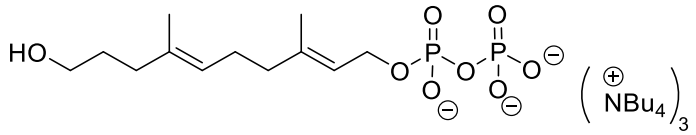
(2E,6E)-11-Bromo-10-hydroxy-3,7,11-trimethyldodeca-2,6-dien-1-yl acetate 170



This compound was prepared from farnesyl acetate (1.5 g, 5.7 mmol) in the same manner as **127**. Purification by column chromatography on silica (15% EtOAc in hexane, R_f 0.3 in 20%) yielded the title compound as a clear oil in 56% yield (1.15 g, 3.2 mmol).

δ_H (300 MHz, $CDCl_3$) 5.40 – 5.29 (1 H, m, CC*H*CH_2O), 5.19 (1 H, t, *J* 6.0, CC*H*$(CH_2)_2$), 4.59 (2 H, d, *J* 7.0, CCHC*H*$_2$O), 3.97 (1 H, dd, *J* 11.0, 2.0, CBrC*H*OH), 2.41 – 1.90 (8 H, m, 4 x CH_2), 2.06 (3 H, s, C*H*$_3$C(O)), 1.71 (3 H, s, C*H*$_3$), 1.59 (3 H, s, C*H*$_3$), 1.35 (3 H, s, CBrC*H*$_3$), 1.33 (3 H, s, CBrC*H*$_3$); δ_C (151 MHz, $CDCl_3$) 171.30 (*C*(O)CH_3), 142.10 (*C*=CH), 133.72 (*C*=CH), 125.52 (C=*C*H), 118.64 (C=*C*HCH_2O), 70.90 (*C*HOH), 61.56 (C=CH*C*H_2O), 39.52, 38.28, 32.20 & 26.74 (CH_2), 26.29 & 26.03 (CBr(*C*$H_3)_2$), 21.21 (C(O)*C*H_3), 16.61 & 16.00 (C=C*C*H_3); HRMS (ES^+, $[M + Na]^+$) found 383.1187, $C_{17}H_{29}O_3BrNa$ requires 383.1198.

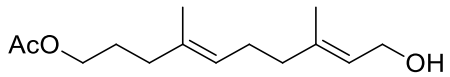
((2E,6E)-10-Hydroxy-3,7-dimethyldeca-2,6-dien-1-yl tris(tetrabutylammonium) diphosphate 172



This compound was prepared from diphosphate **174** (52 mg, 0.046 mmol) by treatment with NH_4OH (1 mM) in D_2O (2 mL), and followed by ^{1}H-NMR of a 500 μL sample at 12 hourly intervals. Upon complete disappearance of the acetate signal (~2.1 ppm, 24 h) the reaction was diluted with d.H_2O and lyophilised to yield the title compound as an off-white powder in 96% yield (48 mg, 0.044 mmol).

δ_H (500 MHz, D_2O) 5.40 (1 H, t, *J* 6.5, CC*H*CH_2O), 5.17 (1 H, t, *J* 7.0, CC*H*$(CH_2)_2$), 4.44 (2 H, s, CCHC*H*$_2O$), 3.51 (2 H, t, *J* 6.5, $(CH_2)_2$C*H*$_2$OH), 3.23 – 3.04 (24 H, m, 12 x N^+C*H*$_2(CH_2)_2CH_3$), 2.20 – 2.02 (4 H, m, 2 x CH_2), 1.99 (2 H, t, *J* 7.5, C*H*$_2(CH_2)_2$OH), 1.86 (3 H, s, C*H*$_3$), 1.66 (3 H, s, C*H*$_3$), 1.64 – 1.51 (26 H, m, 12 x N^+CH_2C*H*$_2CH_2CH_3$ & CH_2), 1.37 – 1.23 (24 H, m, 12 x $N^+(CH_2)_2$C*H*$_2CH_3$), 0.89 (36 H, t, *J* 7.5, 12 x $N^+(CH_2)_3$C*H*$_3$); δ_P (202 MHz, D_2O) -8.33, -8.64; HRMS (ES^-, $[M - H]^-$) found 357.0867, $C_{12}H_{23}O_8P_2$ requires 357.0868.

4E,8E)-10-Hydroxy-4,8-dimethyldeca-4,8-dien-1-yl acetate 173



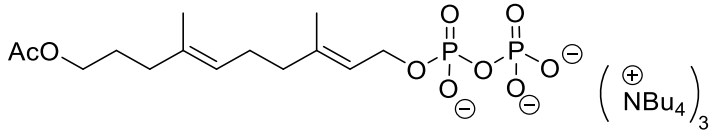
A solution of alcohol **130** (170 mg, 0.6 mmol) in pyridine (2 mL) and acetic anhydride (1 mL) was allowed to stir at room temperature for 16 h before being diluted with hexane/H_2O (10 mL, 1:1) and separated. The aqueous layer was further washed with hexane (5 x 5 mL) and the combined organic fractions washed with $CuSO_4$ (10 mL), $NaHCO_3$ (3x 10 mL), brine, dried ($MgSO_4$) and solvent removed under reduced pressure.

The resulting residue was then taken up in MeOH (10 mL) and a catalytic amount of *para*-toluenesulfonic acid was added, the reaction stirred for 1 h and then quenched with $NaHCO_3$ (1 mL). The MeOH was removed under reduced pressure, the aqueous layer diluted with H_2O/Et_2O (10 mL, 1:1) and separated. The aqueous layer was further washed with Et_2O (3 x 5 mL) and the combined ethereal extracts washed with brine, dried ($MgSO_4$) and the solvent removed. Purification by column chromatography on silica (20 % EtOAc in

hexane) yielded the title compound in 61 % yield over two steps (88 mg, 0.37 mmol).

δ_H (300 MHz, $CDCl_3$) 5.44 – 5.36 (1 H, m, CC*H*CH_2OH), 5.15 – 5.07 (1 H, m, CC*H*$(CH_2)_2$), 4.15 (2 H, bd, *J* 6.5, CCHC*H*$_2$OH), 4.02 (2 H, t, *J* 6.8, AcOC*H*$_2$), 2.18 – 1.99 (6 H, m, 3 x C*H*$_2$), 2.05 (3 H, s, C*H*$_3$C(O)), 1.78 – 1.67 (2 H, m, CH_2), 1.67 (3 H, s, C*H*$_3$), 1.60 (3 H, s, C*H*$_3$); δ_C (75 MHz, $CDCl_3$) 139.35 (C=C), 134.10 (C=C), 124.87 (C*C*H$(CH_2)_2$, 123.85 (C*C*HCH_2OH), 64.28 (AcO*C*H_2), 59.55 (CCH*C*H_2OH), 39.49, 35.79, 26.64 & 26.17 (CH_2), 21.17 (*C*H_3C(O)), 16.33 & 15.97 (CH_3); HRMS (ES^+, $[M]^+$) found 198.1619, $C_{12}H_{22}O_2$ requires 198.1620.

(2E,6E)-10-Acetoxy-3,7-dimethyldeca-2,6-dien-1-yl tris(tetrabutylammonium) diphosphate 174

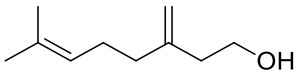


A solution of tetrabutyl dihydrogen phosphate (250 mg, 0.74 mmol) in acetonitrile (1 mL) was added in three portions (3 x 1 mL) at 5 minute intervals to a solution of **173** (78 mg, 0.32 mmol) in trichloroacetonitrile (0.5 mL) at 37 °C. The reaction was incubated at 37 °C for a further 5 minutes after the final addition and the entire reaction mixture applied to a silica column and washed onto the column with isopropanol. The mobile phase (isopropanol : conc. NH_4OH : H_2O, 6 : 2.5 : 0.5) was then begun and fractions collected and analysed by TLC (isopropanol : conc. NH_4OH : H_2O, 6 : 3 : 1, visualised with basic $KMnO_4$), and those containing the diphosphate (R_f 0.26) were combined, ammonia and isopropanol removed under vacuum and the resulting aqueous solution diluted to 60 mL with 25 mM NH_4HCO_3 and lyophilised to yield the title compound as an off-white powder in 31% yield (111 mg, 0.10 mmol).

δ_H (500 MHz, D_2O) 5.47 (1 H, bs, CC*H*CH_2O), 5.23 (1 H, t, *J* 6.5, CC*H*$(CH_2)_2$), 4.50 (2 H, s, CCHC*H*$_2$O), 4.07 (2 H, t, *J* 6.5, AcOCH_2), 3.25 – 3.13 (24 H, m, 12 x N^+C*H*$_2(CH_2)_2CH_3$), 2.25 – 2.02 (6 H, m, 3 x CH_2), 2.09 (3 H, s, C*H*$_3$C(O)), 1.80 – 1.73 (2 H, m, CH_2), 1.73 (3 H, s, CH_3), 1.71 – 1.60 (27 H, m, 12 x N^+CH_2C*H*$_2CH_2CH_3$ & CH_3), 1.41 – 1.31 (24 H, m, 12 x $N^+(CH_2)_2$C*H*$_2CH_3$), 0.94 (36 H, t, *J* 7.5, 12 x $N^+(CH_2)_3$C*H*$_3$); δ_C (126 MHz, D_2O) 124.92 (C*C*H$(CH_2)_2$), 124.91 (C*C*HCH_2OH), 65.13 (AcO*C*H_2), 62.64 (CCH*C*H_2OH), 58.19 (N^+*C*$H_2(CH_2)_2CH_3$), 35.15, 26.56 & 25.66 (CH_2), 23.19 (N^+CH_2*C*$H_2CH_2CH_3$), 20.47 (*C*H_3C(O)), 19.18

($N^+(CH_2)_2CH_2CH_3$), 12.86 ($N^+(CH_2)_3CH_3$); δ_P (202 MHz, D_2O) -6.00, -10.95; HRMS (ES^-, $[M - H]^-$) found 399.0988, $C_{14}H_{25}O_9P_2$ requires 399.0974.

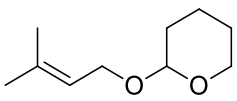
7-Methyl-3-methyleneoct-6-en-1-ol 183



A stirring solution of TMEDA (6.5 g, 8.4 mL, 56 mmol) in anhydrous Et_2O (8 mL) was cooled to -5 °C and n-BuLi (16 mL, 2.5 M in hexane, 40 mmol) was added dropwise and stirred for 1 hour at room temperature. The mixture was then cooled to 0 °C and 3-methylbut-3-en-1-ol (1.7 g, 2.0 mL, 20 mmol) added slowly over 30 minutes. After complete addition of the alcohol the mixture was allowed to warm to room temperature and stirred for a further 6 hours, before cooling to -78 °C and a room temperature solution of 3,3-dimethylallyl bromide (2.2 g, 1.7 mL, 15 mmol) in THF (5 mL) added by syringe driver over 50 minutes. The reaction was stirred at -78 °C and allowed to warm slowly to room temperature overnight before quenching with saturated NH_4Cl (20 mL). The organic layer was separated and the aqueous layer further washed with Et_2O (3 x 15 mL) and the combined ethereal extracts washed with brine and dried ($MgSO_4$) before the solvent was removed under reduced pressure. Purification by flash column chromatography on silica (10 % EtOAc in hexane) gave the title compound as a pale yellow oil in 63 % yield (1.46 g, 9.5 mmol).

δ_H (300 MHz, $CDCl_3$) 5.16 – 5.03 (1 H, m, $(CH_3)_2CCH$), 4.88 (1 H, s, C=C*H*H), 4.83 (1 H, s, C=C*H*H), 3.71 (2 H, t, *J* 6.3, CH_2OH), 2.30 (2 H, t, *J* 6.3, CH_2CH_2OH), 2.21 – 1.99 (4 H, m, 2 x CH_2), 1.68 (3 H, s, CH_3), 1.61 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 145.98, 132.09, 123.90 ($(CH_3)_2CCH$), 111.89 (C=*C*HH), 60.42 (CH_2OH), 39.41 (CH_2CH_2OH), 35.84 (CH_2), 26.45 (CH_2), 25.84 (CH_3), 17.87 (CH_3); HRMS ($APCI^+$, $[M + H]^+$) found 155.1435, $C_{10}H_{19}O$ requires 155.1436.

2-((3-Methylbut-2-en-1-yl)oxy)tetrahydro-2H-pyran 189

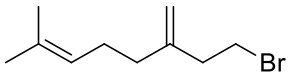


To a stirring solution of prenol (5 g, 58 mmol) and 3,4-dihydropyran (9.8 g, 11 mL, 116 mmol) in DCM (150 mL) was added a catalytic amount of PPTS (1 mol %) and stirred at room temperature for 1 h, before d.H_2O (50 mL) was added and the mixture separated. The aqueous layer was further washed with DCM (3 x 20 mL) and the combined organic extracts washed with brine, dried

($MgSO_4$) and sovent removed under reduced pressure. Purification by column chromatography on silica (10% EtOAc in hexane) yielded the title compound as a clear oil in 98% yield (56.7 mmol).

δ_H (300 MHz, $CDCl_3$) 5.43 – 5.27 (1 H, m, CC*H*CH_2), 4.71 – 4.55 (1 H, m, OC*H*O), 4.10 (2 H, ddd, *J* 73.1, 11.7, 7.2, CCHC*H*$_2$O), 3.95 – 3.42 (2 H, m, (C*H*$_2$)$_3$CH_2O), 1.90 – 1.44 (6 H, m, 3 x CH_2), 1.75 (3 H, s, CH_3), 1.69 (3 H, s, CH_3); δ_C (75 MHz, $CDCl_3$) 137.37 (CH_3*C*CH), 120.91 (C*C*HCH_2), 98.04 (O*C*HO), 63.81 (CCH*C*H_2O), 62.43 ((C*H*$_2$)$_3$CH_2O), 30.83 (CH_2), 26.03 (CH_3), 25.62 (CH_2), 19.76 (CH_2), 18.13 (CH_3); HRMS (ES^+, $[M]^+$) found 170.1301, $C_{10}H_{18}O_2$ requires 170.1307.

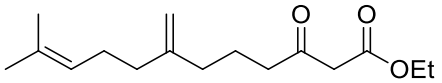
8-Bromo-2-methyl-6-methyleneoct-2-ene 187



A stirring solution of **183** (600 mg, 3.9 mmol) and Et_3N (2.36 g, 3.25 mL, 23 mmol) in anhydrous THF (50 mL) was cooled to -10 ºC and methanesulfonyl chloride (893 mg, 603 µL, 7.8 mmol) added dropwise over 5 minutes. The reaction was stirred for a further 30 minutes before lithium bromide (3.4 g, 39 mmol) was added and the reaction allowed to warm to room temperature and stirred for a further 20 h. The reaction was quenched with d.H_2O (15 mL) and separated with Et_2O (10 mL). The aqueous layer was further washed with Et_2O (3 x 10 mL) and the combined organic extracts washed with brine, dried ($MgSO_4$) and the solvent removed under reduced pressure. Purification by column chromatography on silica (5% EtOAc in hexane) yielded the title compound in 72% yield (605 mg, 2.79 mmol).

δ_H (300 MHz, $CDCl_3$) 5.23 – 5.04 (1 H, m, $(CH_3)_2$CC*H*), 4.91 – 4.85 (1 H, m, C=C*H*H), 4.85 – 4.78 (1 H, m, C=C*H*H), 3.47 (2 H, t, *J* 7.5, C*H*$_2$Br), 2.58 (2 H, t, *J* 7.5, C*H*$_2$CH_2Br), 2.19 – 1.99 (4 H, m, CCHC*H*$_2$C*H*$_2$), 1.69 (3 H, d, *J* 1.0, C*H*$_3$), 1.61 (3 H, s, C*H*$_3$); δ_C (75 MHz, $CDCl_3$) 146.48 (C=C), 132.18 (C=C), 123.77 (($CH_3)_2$C*C*H), 111.73 (C=*C*HH), 39.58, 35.80, 31.17 & 26.39 (*C*H_2), 25.85 & 17.89 (*C*H_3); HRMS ($APCI^+$, $[M + H]^+$) found 217.0593, $C_{10}H_{18}Br$ requires 217.0593.

Ethyl 11-methyl-7-methylene-3-oxododec-10-enoate 192

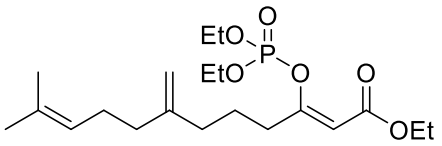


To an oven-dried RBF charged with a magnetic stirrer bar was added sodium hydride (8.4 mmol, 335 mg, 60% in mineral oil) and washed with anhydrous

hexane (3 x 20 mL) by stirring for 5 minutes and removal of the hexane by syringe. The washed sodium hydride was taken up in THF (40 mL) and cooled to 0 °C before the dropwise addition of ethyl acetoacetate (4.2 mmol, 531 μL) over 20 minutes and then stirred at 0 °C for a further 20 minutes. A solution of nBuLi (4.9 mmol, 1.96 mL, 2.5 M in hexane) was then added dropwise over 20 minutes at 0 °C and the solution stirred for a further 30 minutes. A room temperature solution of **187** (1.4 mmol, 300 mg) in THF (2 mL) was added dropwise over 20 minutes and stirred for 3 h allowing it to slowly warm to 10 °C. The reaction was quenched by dropwise addition of HCl (5 mL, 10% v/v) and stirred for 30 minutes before it was diluted with d.H_2O:EtOAc (1:1, 20 mL) separated and the aqueous layer washed further with EtOAc (3 x 20 mL). The combined organic extracts were washed with brine, dried ($MgSO_4$) and solvent removed under vaccum. Purification by column chromatography on silica (10% EtOAc in hexane, R_f 0.56 in 20%) gave the title compound in 65% yield (241 mg, 0.9 mmol).

δ_H (300 MHz, $CDCl_3$) 5.15 – 5.05 (1 H, m, $(CH_3)_2CCH$), 4.75 (1 H, bs, C=C*H*H), 4.72 (1 H, bs, C=C*H*H), 4.19 (2 H, q, *J* 7.0, $CH_3CH_2OC(O)$), 3.43 (2 H, s, $C(O)CH_2C(O)OEt$), 2.54 (2 H, t, *J* 7.5, $CH_2C(O)CH_2C(O)OEt$), 2.17 – 1.94 (6 H, m, 3 x CH_2), 1.81 – 1.69 (2 H, m, $CH_2CH_2C(O)$), 1.68 (3 H, d, *J* 1.0, CH_3), 1.60 (3 H, s, CH_3), 1.28 (3 H, t, *J* 7.0, CH_3CH_2O); δ_C (75 MHz, $CDCl_3$) 202.89 ($C(O)CH_2$), 167.38 ($C(O)OEt$), 148.67 (C=C), 131.84 (C=C), 124.14 (C=C), 109.84 (C=*C*HH), 49.55 ($C(O)CH_2C(O)OEt$, 35.88, 35.31, 26.48, 25.84, 21.38 (CH_3), 17.86 (CH_3), 14.25 (CH_3); HRMS (APCI$^+$, $[M + Na]^+$) found 289.1768, $C_{16}H_{26}O_3Na$ requires 289.1780.

Ethyl (Z)-3-((diethoxyphosphoryl)oxy)-11-methyl-7-methylenedodeca-2,10-dienoate194

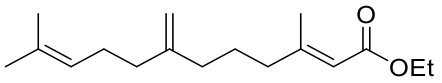


To an oven dried round-bottom flask charged with a magnetic stirrer bar and flushed with argon was added sodium hydride (256 mg, 6.4 mmol, 60 % dispersion in mineral oil) and washed twice with dry hexane. The washed NaH was taken up in dry Et_2O (30 mL) and cooled to 0 °C while stirring, a room temperature solution of **192** (340 mg, 1.28 mmol) in dry Et_2O (10 mL) was then added dropwise over 15 minutes before being allowed to warm to room temperature and stirred for a further 30 minutes. The reaction was then cooled to 0 °C and neat diethyl chlorophosphate (368 μL, 440 mg, 2.55 mmol) added

dropwise over 15 minutes. The reaction was stirred for a further 30 minutes at 0 °C before it was quenched with dropwise addition of sat. NH_4Cl (10 mL) and stirred for 10 minutes (0 °C). The mixture was diluted with H_2O (10 mL), separated, and the aqueous fraction further washed with Et_2O (3 x 10 mL). Combined ethereal extracts were then washed with brine, dried (Na_2SO_4) and solvent removed under reduced pressure, the residue was then stored at -20 °C and used the following day without further purification.

δ_H (300 MHz, $CDCl_3$) 5.36 (1 H, s, C(O)C*H*COP(O)$(OEt)_2$), 5.16 – 5.00 (1 H, m, $(CH_3)_2$CC*H*), 4.76 (1 H, bs, C=C*H*H), 4.73 (1 H, bs, C=C*H*H), 4.32 – 4.20 (4 H, m, P(O)(OC*H*$_2$CH$_3$)), 4.15 (2 H, q, *J* 7.0, $CH_3CH_2OC(O)$), 2.43 (2 H, t, *J* 7.5), 2.16 – 1.96 (6 H, m, 3 x C*H*$_2$), 1.78 – 1.69 (2 H, m, CH_2), 1.68 (3 H, d, *J* 1.0, C*H*$_3$), 1.60 (3 H, s, C*H*$_3$), 1.40 – 1.32 (6 H, m, P(O)(OCH$_2$C*H*$_3$)), 1.27 (7 H, t, *J* 7.0, C*H*$_3$CH$_2$OC(O)); δ_P (122 MHz, $CDCl_3$) -8.11 (s).

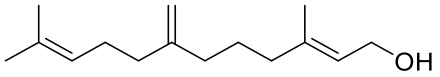
Ethyl (E)-5,13-dimethyl-9-methylene-3-oxotetradeca-4,12-dienoate 195



A stirring suspension of copper (I) iodide (495 mg, 2.6 mmol) and Et_2O (30 mL) in an oven-dried round-bottom flask charged with a magnetic stirrer bar and flushed with argon was cooled to 0 °C. Methyl lithium (3.25 mL, 5.2 mmol, 1.6 M in Et_2O) was added dropwise over 10 minutes and stirred at 0 °C for 30 minutes until all the CuI was dissolved to give a clear, almost colourless solution. The solutuion was then cooled to -78 °C and a solution of **194** (1.3 mmol assumed) in dry Et_2O (7 mL) added dropwise over 30 minutes, a yellow colour developed which then deepened while stirring at -78 °C for a further 2 h. The mixture was then allowed to warm to -45 °C and maintained at this temperature for a further 2 h, over which time the colour darkened and turned a deep purple. Neat methyl iodide (202 μL, 461 mg, 3.25 mmol) was then added and stirred at -45 °C for 20 minutes before the reaction mixture was poured into an ice-cold mixture of sat. NH_4Cl and conc. NH_4OH (100 mL, 4:1) and stirred for 30 minutes until all the copper salts were dissolved. The mixture was then separated, the aqueous layer further washed with Et_2O (5 x 15 mL) and the combined ethereal extracts washed with NH_4OH (2 x 20 mL, 10 % v/v), H_2O (2 x 20 mL), brine, dried ($MgSO_4$) and the solvent removed under reduced pressure. Purification by column chromatography on silica (10% EtOAc in hexane, R_f 0.7 in 20%) yielded the title compound in 31% yield over 2 steps (106 mg, 0.40 mmol).

δ_H (300 MHz, $CDCl_3$) 5.67 (1 H, dd, *J* 2.5, 1.0, CC*H*C(O)OEt), 5.15 – 5.07 (1 H, m, $(CH_3)_2CCH$), 4.74 (1 H, bs, C=C*H*H), 4.73 (1 H, bs, C=C*H*H), 4.14 (2 H, q, *J* 7.0, $CH_3CH_2OC(O)$), 2.16 (3 H, d, *J* 1.0, $CH_3CCHC(O)$), 2.18 – 1.97 (8 H, m, 4 x CH_2), 1.69 (3 H, d, *J* 1.0, CH_3), 1.61 (2 H, m, CH_2), 1.61 (3 H, s, CH_3), 1.28 (3 H, t, *J* 7.0, $CH_3CH_2OC(O)$); δ_C (75 MHz, $CDCl_3$) 149.10 (*C*(O)OEt), 131.84 (C=C), 124.18 ($(CH_3)_2CCH$), 115.75 (C*C*HC(O)), 110.19 (C=*C*HH), 59.63 ($CH_3CH_2OC(O)$) 40.64 (CH_2), 35.70 (CH_2), 26.52 (CH_2), 25.85 (CH_2), 25.53 (CH_3), 18.88 (CH_3), 17.86 (CH_3), 14.50 ($CH_3CH_2OC(O)$). HRMS (EI^+, $[M]^+$) found 264.2090, $C_{17}H_{28}O_2$ requires 264.2089.

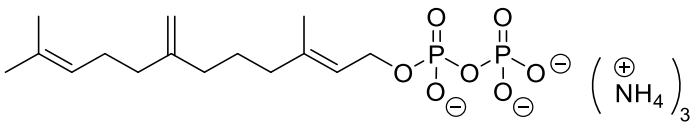
(E)-3,11-Dimethyl-7-methylenedodeca-2,10-dien-1-ol 182



A stirring solution of **195** (106 mg, 0.40 mmol) in anhydrous toluene (15 mL) was cooled to -78 °C for 15 minutes before the dropwise addition of DIBAL-H (1.2 mL, 1.2 mmol, 1 M in toluene) over 30 minutes. The reaction was stirred at -78 °C for a further 1 h before the dropwise addition of methanol (1 mL) and the mixture allowed to warm to room temperature. A mixture of sat. NH_4Cl and 1 M HCl (1:1, 10 mL) was added and the mixture stirred vigorously for 30 minutes before it was separated and the aqueous layer washed with Et_2O (5 x 5 mL). The combined organic extracts were washed with d.H_2O and brine, dried ($MgSO_4$) and solvent removed under reduced pressure. Purification by column chromatography on silica (10% EtOAc in hexane, R_f 0.26) yielded the title compound in 85% yield (76 mg, 0.34 mmol).

δ_H (300 MHz, $CDCl_3$) 5.48 – 5.36 (1 H, m, CC*H*CH_2OH), 5.16 – 5.06 (1 H, m, $(CH_3)_2CCH$), 4.73 (1 H, d, *C*=C*H*H), 4.72 (1 H, d, *C*=C*H*H), 4.16 (2 H, d, *J* 7.0, CCHC*H*$_2$OH), 2.19 – 1.93 (8 H, m, 4x CH_2), 1.69 (3 H, d, *J* 0.9), 1.67 (3 H, s, CH_3), 1.61 (3 H, s, CH_3), 1.59 – 1.48 (2 H, m, CH_2); δ_C (75 MHz, $CDCl_3$) 149.64, 140.05, 131.75, 124.29 ($(CH_3)_2CCH$), 123.49 (C*C*HCH_2OH), 109.07 $(CH_2)_2CCH_2(CH_2)_3$), 59.56 (CCH*C*H_2OH), 39.34 (CH_2), 36.16 (CH_2), 35.88 (CH_2), 26.56 (CH_2), 25.90 (CH_3), 25.85, 17.85 (CH_3), 16.34 (CH_3); HRMS (EI^+, $[M - H_2O]^+$) found 204.1879, $C_{15}H_{24}$ requires 204.1878.

(E)-3,11-Dimethyl-7-methylenedodeca-2,10-dien-1-yl trisammonium diphosphate 180



This compound was prepared and purified from alcohol **182** (76 mg, 0.34 mmol) in the same manner as diphosphate **137**, to give the title compound as a white solid in 45% yield (67 mg, 0.15 mmol).

δ_H (500 MHz, D_2O) 5.41 (1 H, bs, CC*H*CH_2OH), 5.15 (1 H, bs, $(CH_3)_2$CC*H*), 4.42 (2 H, bs, CCHC*H*$_2$O), 2.07 (8 H, bm, 4 x CH_2), 1.66 (3 H, bs, CH_3), 1.64 (3 H, bs, CH_3), 1.60 – 1.48 (2 H, m, CH_2), 1.58 (3 H, bs, CH_3); δ_P (202 MHz, D_2O) -6.41 (d, *J* 22.0), -10.21 (d, *J* 22.0).; HRMS (ES^-, $[M - H]^-$) found 381.1215, $C_{15}H_{27}O_7P_2$ requires 381.1232.

7 References

(1) Gershenzon, J.; Dudareva, N. *Nat. Chem. Biol.* **2007**, *3*, 408.

(2) Connolly, J. D.; Hill, R. A. *Dictionary of Terpenoids*; Chapman and Hall, London, 1991.

(3) Bick, J. A.; Lange, B. M. *Arch. Biochem. Biophys.* **2003**, *415*, 146.

(4) Rees, B. H. H.; Goad, L. J.; Goodwin, N. D. T. W. **1968**, 417.

(5) Hick, A. J.; Luszniak, M. C.; Pickett, J. a. *Nat. Prod. Rep.* **1999**, *16*, 39.

(6) Pickett, J. a; Allemann, R. K.; Birkett, M. a. *Nat. Prod. Rep.* **2013**, *30*, 1277.

(7) Manfredi, J.; Horwitz, S. *Pharmacol. Ther.* **1984**, *25*.

(8) White, N. J. *Science* **2008**, *320*, 330.

(9) Quin, M. B.; Flynn, C. M.; Schmidt-Dannert, C. *Nat. Prod. Rep.* **2014**, *31*, 1449.

(10) Cane, D. E.; Ikeda, H. *Acc. Chem. Res.* **2012**, *45*, 463.

(11) Rabe, P.; Citron, C.; Dickschat, J. *ChemBioChem* **2013**, 1.

(12) Tetzlaff, C. N.; You, Z.; Cane, D. E.; Takamatsu, S.; Omura, S.; Ikeda, H. *Biochemistry* **2006**, *45*, 6179.

(13) Lin, X.; Cane, D. E. *J. Am. Chem. Soc.* **2009**, *131*, 6332.

(14) Ruzicka, L. *Experientia* **1994**, *50*, 395.

(15) Clifford, K.; Cornforth, J. W.; Mallaby, R.; Phillips, G. T. *J. Chem. Soc. D Chem. Commun.* **1971**, 1599.

(16) Street, I. P.; Christensen, D. J.; Poulter, C. D. *J. Am. Chem. Soc.* **1990**, *112*, 8577.

(17) McCaskill, D.; Croteau, R. *Tetrahedron Lett.* **1999**, *40*, 653.

(18) Eisenreich, W.; Rohdich, F.; Bacher, A. *Trends Plant Sci.* **2001**, *6*, 78.

(19) Gräwert, T.; Groll, M.; Rohdich, F.; Bacher, A.; Eisenreich, W. *Cell. Mol. Life Sci.* **2011**, *68*, 3797.

(20) Dickschat, J. S. *Nat. Prod. Rep.* **2011**, *28*, 1917.

(21) Wang, W.; Wang, K.; Span, I. *J. Am. Chem. Soc.* **2012**.

(22) Tarshis, L.; Yan, M.; Poulter, C.; Sacchettini, J. *Biochemistry* **1994**, 10871.

(23) Lesburg, C. A. *Science (80-.).* **1997**, *277*, 1820.

(24) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.

(25) Katoh, S.; Hyatt, D.; Croteau, R. *Arch. Biochem. Biophys.* **2004**, *425*, 65.

(26) Back, K.; Chappell, J. *Proc. Natl. Acad. Sci.* **1996**, *93*, 6841.

(27) Li, A.; González, V.; Grundy, D. J.; Demiray, M.; Faraldos, J. A.; Allemann, R. K. *Unpubl. Work*.

(28) Gennadios, H. A.; Gonzalez, V.; Di Costanzo, L.; Li, A. A.; Yu, F. L.; Miller, D. J.; Allemann, R. K.; Christianson, D. W. *Biochemistry* **2009**, *48*, 6175.

(29) Prosser, I.; Altug, I. G.; Phillips, A. L.; König, W. A.; Bouwmeester, H. J.; Beale, M. H. *Arch. Biochem. Biophys.* **2004**, *432*, 136.

(30) Shishova, E.; Costanzo, L. Di; Cane, D. E.; Christianson, D. W. *Biochemistry* **2007**, *46*, 1941.

(31) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13543.

(32) Rynkiewicz, M.; Cane, D.; Christianson, D. *Biochemistry* **2002**, 1732.

(33) Caruthers, J. M.; Kang, I.; Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.* **2000**, *275*, 25533.

(34) Aaron, J. a; Lin, X.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2010**, *49*, 1787.

(35) Baer, P.; Rabe, P.; Citron, C. A.; de Oliveira Mann, C. C.; Kaufmann, N.; Groll, M.; Dickschat, J. S. *Chembiochem* **2014**, *15*, 213.

(36) Baer, P.; Rabe, P.; Fischer, K.; Citron, C. A.; Klapschinski, T. A.; Groll, M.; Dickschat, J. S. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7652.

(37) Shishova, E. Y.; Yu, F.; Miller, D. J.; Faraldos, J. a; Zhao, Y.; Coates, R. M.; Allemann, R. K.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.* **2008**, *283*, 15431.

(38) Chen, M.; Al-lami, N.; Janvier, M.; D'Antonio, E. L.; Faraldos, J. a; Cane, D. E.; Allemann, R. K.; Christianson, D. W. *Biochemistry* **2013**, *52*, 5441.

(39) van der Kamp, M. W.; Sirirak, J.; Zurek, J.; Allemann, R. K.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 8094.

(40) Faraldos, J. A.; Gonzalez, V.; Allemann, R. K. *Chem. Commun.* **2012**, *48*, 3230.

(41) Allinger, N.; Siefert, J. *J. Am. Chem. Soc.* **1975**.

(42) Croteau, R. B. *Chem. Rev.* **1987**, *87*, 929.

(43) Zhang, Q.; Tiefenbacher, K. *Nat. Chem.* **2015**, *7*, 197.

(44) Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem.* **1998**, *273*, 2078.

(45) Jones, C. G.; Keeling, C. I.; Ghisalberti, E. L.; Barbour, E. L.; Plummer, J. A.; Bohlmann, J. *Arch. Biochem. Biophys.* **2008**, *477*, 121.

(46) Agger, S. A.; Lopez-Gallego, F.; Hoye, T. R.; Schmidt-Dannert, C. *J. Bacteriol.* **2008**, *190*, 6084.

(47) Faraldos, J. A.; Kariuki, B.; Allemann, R. K. *J. Org. Chem.* **2010**, *75*, 1119.

(48) Faraldos, J. A.; Allemann, R. K. *Org. Lett.* **2011**, *13*, 1202.

(49) Cane, D. E. DE; Yang, G.; Coates, R. M. R.; Pyun, H. J.; Hohn, T. M. *J. Org. Chem.* **1992**, *57*, 3454.

(50) Steiger, A.; Pyun, H. J.; Coates, R. M. *J. Org. Chem.* **1992**, *57*, 3444.

(51) Peters, R. J.; Ravn, M. M.; Coates, R. M.; Croteau, R. B. *J. Am. Chem. Soc.* **2001**, *123*, 8974.

(52) Ravn, M. M.; Peters, R. J.; Coates, R. M.; Croteau, R. *J. Am. Chem. Soc.* **2002**, *124*, 6998.

(53) Roy, A.; Roberts, F. G.; Wilderman, P. R.; Zhou, K.; Peters, R. J.; Coates, R. M. *J. Am. Chem. Soc.* **2007**, *129*, 12453.

(54) Mann, F. M.; Prisic, S.; Hu, H.; Xu, M.; Coates, R. M.; Peters, R. J. *J. Biol. Chem.* **2009**, *284*, 23574.

(55) Koohang, A.; Coates, R. M.; Owen, D.; Poulter, C. D.; Avenue, S. M. *J. Org. Chem.* **1999**, *64*, 6.

(56) Sandifer, R. M.; Thompson, M. D.; Gaughan, R. G.; Poulter, C. D. *J. Am. Chem. Soc.* **1982**, *104*, 7376.

(57) Poulter, C. D. *Acc. Chem. Res.* **1990**, *23*, 70.

(58) Whittington, D. A.; Wise, M. L.; Urbansky, M.; Coates, R. M.; Croteau, R. B.; Christianson, D. W.; Douglas A. Whittington, Mitchell L. Wise, Marek Urbansky, Robert M. Coates, Rodney B. Croteau, and D. W. C. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15375.

(59) Vedula, L. S.; Rynkiewicz, M. J.; Pyun, H.-J.; Coates, R. M.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2005**, *44*, 6153.

(60) Cane, D. E.; Prabhakaran, P. C.; Salaski, E. J.; Harrison, P. H. M.; Noguchi, H.; Rawlings, B. J. *J. Am. Chem. Soc.* **1989**, *111*, 8914.

(61) Cane, D. E.; Swanson, S.; Murthy, P. P. N. *J. Am. Chem. Soc.* **1981**, *103*, 2136.

(62) Cane, D. E.; Ha, H.-J.; Pargellis, C.; Waldmeier, F.; Swanson, S.; Murthy, P. P. N. *Bioorg. Chem.* **1985**, *13*, 246.

(63) Steliopoulos, P.; Wüst, M.; Adam, K.-P.; Mosandl, A. *Phytochemistry* **2002**, *60*, 13.

(64) Schmidt, C. O. C.; Bouwmeester, H. J. H.; Franke, S.; K�nig, W. A.; König, W. *Chirality* **1999**, *362*, 353.

(65) Faraldos, J. A.; Miller, D. J.; Gonzalez, V.; Yoosuf-Aly, Z.; Cascon, O.; Li, A.; Allemann, R. K. *J. Am. Chem. Soc.* **2012**, *134*, 5900.

(66) Picaud, S.; Mercke, P.; He, X.; Sterner, O.; Brodelius, M.; Cane, D. E.; Brodelius, P. E. *Arch. Biochem. Biophys.* **2006**, *448*, 150.

(67) Yu, F.; Miller, D. J.; Allemann, R. K. *Chem. Commun.* **2007**, 4155.

(68) Miller, D. J.; Yu, F. L.; Allemann, R. K. *Chembiochem* **2007**, *8*, 1819.

(69) Miller, D. J.; Yu, F. L.; Knight, D. W.; Allemann, R. K. *Org. Biomol. Chem.* **2009**, *7*, 962.

(70) Cane, D. E.; Xue, Q.; Fitzsimons, B. C. *Biochemistry* **1996**, *35*, 12369.

(71) Felicetti, B.; Cane, D. *J. Am. Chem. Soc.* **2004**, 7212.

(72) Seemann, M.; Zhai, G. *J. Am. Chem. Soc.* **2002**, *124*, 7681.

(73) Calvert, M. J.; Taylor, S. E.; Allemann, R. K. *Chem. Commun.* **2002**, 2384.

(74) Deligeorgopoulou, A.; Taylor, S. E.; Forcat, S.; Allemann, R. K. *Chem. Commun.* **2003**, 2162.

(75) Seemann, M.; Zhai, G. Z.; Umezawa, K.; Cane, D. *J. Am. Chem. Soc.* **1999**, *121*, 591.

(76) Yoshikuni, Y.; Ferrin, T. E.; Keasling, J. D. *Nature* **2006**, *440*, 1078.

(77) López-Gallego, F.; Wawrzyn, G. T.; Schmidt-Dannert, C. *Appl. Environ. Microbiol.* **2010**, *76*, 7723.

(78) Prosser, I.; Phillips, A. L.; Gittings, S.; Lewis, M. J.; Hooper, A. M.; Pickett, J. A.; Beale, M. H. *Phytochemistry* **2002**, *60*, 691.

(79) Gonzalez, V.; Touchet, S.; Grundy, D. J.; Faraldos, J. A.; Allemann, R. K. *J. Am. Chem. Soc.* **2014**, *136*, 14505.

(80) Gonzalez, V.; Touchet, S.; Grundy, D. J.; Faraldos, J. a; Rudolf, K. **2014**.

(81) Schmidt, C. O.; Bouwmeester, H. J.; de Kraker, J.-W.; König, W. A. *Angew. Chemie Int. Ed.* **1998**, *37*, 1400.

(82) Benedict, C. R. *PLANT Physiol.* **2001**, *125*, 1754.

(83) Garms, S.; Köllner, T. G.; Boland, W. *J. Org. Chem.* **2010**, *75*, 5590.

(84) Mercke, P.; Bengtsson, M.; Bouwmeester, H. J.; Posthumus, M. A.; Brodelius, P. E. *Arch. Biochem. Biophys.* **2000**, *381*, 173.

(85) Picaud, S.; Olofsson, L.; Brodelius, M.; Brodelius, P. E. *Arch. Biochem. Biophys.* **2005**, *436*, 215.

(86) Kim, S.-H.; Heo, K.; Chang, Y.-J.; Park, S.-H.; Rhee, S.-K.; Kim, S.-U. *J. Nat. Prod.* **2006**, *69*, 758.

(87) Salmon, M.; Laurendon, C.; Vardakou, M.; Cheema, J.; Defernez, M.; Green, S.; Faraldos, J. a.; O'Maille, P. E. *Nat. Commun.* **2015**, *6*, 6143.

(88) Cane, D. E.; Kang, I. *Arch. Biochem. Biophys.* **2000**, *376*, 354.

(89) Proctor, R. H.; Hohn, T. M. *J. Biol. Chem.* **1993**, *268*, 4543.

(90) Calvert, M. J.; Ashton, P. R.; Allemann, R. K. *J. Am. Chem. Soc.* **2002**, *124*, 11636.

(91) Crock, J.; Wildung, M.; Croteau, R. *Proc. Natl. Acad. Sci.* **1997**, *94*, 12833.

(92) Beale, M. H.; Birkett, M. A.; Bruce, T. J. A.; Chamberlain, K.; Field, L. M.; Huttly, A. K.; Martin, J. L.; Parker, R.; Phillips, A. L.; Pickett, J. A.; Prosser, I. M.; Shewry, P. R.; Smart, L. E.; Wadhams, L. J.; Woodcock, C. M.; Zhang, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10509.

(93) Nakano, C.; Kudo, F.; Eguchi, T.; Ohnishi, Y. *Chembiochem* **2011**, *12*, 2271.

(94) Raldugin, V.; Salenko, V.; Gamov, N. *Chem. Nat. Compd.* **1980**, 154.

(95) Bohlmann, F.; Zdero, C.; King, R.; Robinson, H. *Phytochemistry* **1984**, *473*, 15.

(96) Cornwell, C. P.; Reddy, N.; Leach, D. N.; Wyllie, S. G. *Flavour Fragr. J.* **2001**, *16*, 263.

(97) Köpke, D.; Schröder, R.; Fischer, H. M.; Gershenzon, J.; Hilker, M.; Schmidt, A. *Planta* **2008**, *228*, 427.

(98) Yoshikuni, Y.; Martin, V. J. J.; Ferrin, T. E.; Keasling, J. D. *Chem. Biol.* **2006**, *13*, 91.

(99) Köpke, D.; Beyaert, I.; Gershenzon, J.; Hilker, M.; Schmidt, A. *Phytochemistry* **2010**, *71*, 909.

(100) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, a; Main, a; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Jiang, H.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. *Nature* **2013**, *496*, 528.

(101) Gilmore, K.; Kopetzki, D.; Lee, J. W.; Horváth, Z.; McQuade, D. T.; Seidel-Morgenstern, A.; Seeberger, P. *Chem. Commun.* **2014**.

(102) Karp, F.; Zhao, Y.; Santhamma, B.; Assink, B.; Coates, R. M.; Croteau, R. B. *Arch. Biochem. Biophys.* **2007**, *468*, 140.

(103) O'Maille, P. E.; Chappell, J.; Noel, J. P. *Anal. Biochem.* **2004**, *335*, 210.

(104) Vardakou, M.; Salmon, M.; Faraldos, J. A.; O'Maille, P. E. *MethodsX* **2014**, *1*, 187.

(105) Nordin, O.; Hedenström, E.; Högberg, H.-E. *Acta Chem. Scand.* **1999**, *53*, 124.

(106) Faraldos, J. A.; Touchet, S. *Unpubl. Work*.

(107) Stein, S.; Mirokhin, Y.; Tchekhovskoi, D.; Mallard, G.; Mikaia, A.; Neta, P.; Sparkman, D.; White, E.; Yang, X.; Zaikin, V.; Zhu, D. NIST Mass Spectral Search Program. Version 2.0 g. build May 19 2011.

(108) Cane, D. E.; Chiu, H. T.; Liang, P. H.; Anderson, K. S. *Biochemistry* **1997**, *36*, 8332.

(109) Mathis, J. R.; Back, K.; Starks, C.; Noel, J.; Poulter, C. D.; Chappell, J. *Biochemistry* **1997**, *36*, 8340.

(110) Bohlmann, J.; Crock, J.; Jetter, R.; Croteau, R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6756.

(111) Lopez-Gallego, F.; Agger, S. A.; Abate-Pella, D.; Distefano, M. D.; Schmidt-Dannert, C. *Chembiochem* **2010**, *11*, 1093.

(112) Faraldos, J. a; Wu, S.; Chappell, J.; Coates, R. M. *Tetrahedron* **2007**, *63*, 7733.

(113) Rabe, P.; Barra, L.; Rinkel, J.; Riclea, R.; Citron, C. a.; Klapschinski, T. a.; Janusko, A.; Dickschat, J. S. *Angew. Chemie Int. Ed.* **2015**, n/a.

(114) Nishimura, K.; Horibe, I.; Tom, K. *Tetrahedron* **1973**, *29*, 271.

(115) Turdybekov, K. M.; Lindeman, S. V.; Timofeeva, T. V.; Struchkov, Y. T. *Chem. Nat. Compd.* **1991**, *27*, 288.

(116) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195.

(117) Zhang, Y. *BMC Bioinformatics* **2008**, *9*, 40.

(118) Hu, Y.; Chou, W. K. W.; Hopson, R.; Cane, D. E. *Chem. Biol.* **2011**, *18*, 32.

(119) Cane, D. E.; Watt, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1547.

(120) Cane, D.; Sohng, J.; Lamberson, C. *Biochemistry* **1994**, 5846.

(121) Schägger, H.; von Jagow, G. *Anal. Biochem.* **1991**, *199*, 223.

(122) Wittig, I.; Braun, H.-P.; Schägger, H. *Nat. Protoc.* **2006**, *1*, 418.

(123) Davisson, V.; Woodside, A. *J. Org. Chem.* **1986**, *51*, 4768.

(124) Keller, R. K.; Thompson, R. *J. Chromatogr. A* **1993**, *645*, 161.

(125) Cramer, F.; Böhm, W. *Angew. Chemie* **1959**, *71*, 775.

(126) Cornforth, R. H.; Popják, G. *Steroids and Terpenoids*; Methods in Enzymology; Elsevier, 1969; Vol. 15.

(127) Cramer, F.; Rittersdorf, W.; Böhm, W. *Justus Liebigs Ann. Chem.* **1962**, *654*, 180.

(128) Popják, G.; Goodman, D. S.; Cornforth, R. H.; Ryhage, R.; Cornforth, J. W.; Cornforth, R. H.; Ryhage, R.; Goodman, D. S. *J. Biol. Chem.* **1962**, *237*, 56.

(129) Lira, L. M.; Vasilev, D.; Pilli, R. A.; Wessjohann, L. A. *Tetrahedron Lett.* **2013**, *54*, 1690.

(130) Turek, T. C.; Gaon, I.; Gamache, D.; Distefano, M. D. *Bioorganic Med.*

Chem. Lett. **1997**, *7*, 2125.

(131) Labadie, G. *J. Org. Chem.* **2007**, 30451.

(132) Ceruti, M.; Amisano, S.; Milla, P.; Viola, F.; Rocco, F.; Jung, M.; Cattel, L. *J. Chem. Soc. Perkin Trans. 1* **1995**, 889.

(133) Paizs, C.; To??a, M.; Majdik, C.; B??dai, V.; Nov??k, L.; Irimie, F.-D.; Poppe, L. *J. Chem. Soc. Perkin Trans. 1* **2002**, 2400.

(134) Yang, Y.; Mikeš, F.; Yang, L.; Liu, W.; Koike, Y.; Okamoto, Y. *J. Fluor. Chem.* **2006**, *127*, 277.

(135) Robine, S.; Grundy, D. J. *Unpubl. Work*.

(136) Lichtor, P. a; Miller, S. J.; Box, P. O.; Haven, N.; S-, P. *Nat. Chem.* **2012**, *4*, 990.

(137) Lichtor, P. a.; Miller, S. J. *J. Am. Chem. Soc.* **2014**, *136*, 5301.

(138) Sharpless, K. B.; Michaelson, R. C. *J. Am. Chem. Soc.* **1973**, *95*, 6136.

(139) Kigoshi, H.; Ojika, M.; Shizuri, Y.; Niwa, H.; Yamada, K. *Tetrahedron* **1986**, *42*, 3789.

(140) Cha, M.-R.; Choi, C.-W.; Lee, J.-Y.; Kim, Y.-S.; Yon, G.-H.; Choi, S.-U.; Ryu, S.-Y. *Bull. Korean Chem. Soc.* **2012**, *33*, 2378.

(141) Xia, D.; Du, Y.; Yi, Z.; Song, H.; Qin, Y. *Chem. - A Eur. J.* **2013**, *19*, 4423.

(142) Yong, K. H.; Lotoski, J. a.; Chong, J. M. *J. Org. Chem.* **2001**, *66*, 8248.

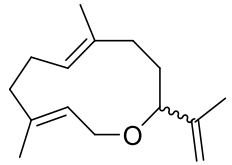
(143) Na, O.; Jin, Y.; Roberts, F. G.; Coates, R. M. *Org. Synth.* **2007**, *84*, 43.

(144) Sum, F. W.; Weiler, L. *Tetrahedron* **1981**, *37*, 303.

(145) Sum, F. W.; Weiler, L.; Cummins, C. H. *Org. Synth.* **1984**, *62*, 14.

8 Appendix

8.1 NMR spectroscopic analysis of macrocyclic ether 158



Compound **158** (12.2 mg, 0.055 mmol) was prepared by preparative-scale incubation of GdolS with compound **162** (60 mg, 0.134 mmol) in 41% yield as described in Sections 4.2.3 and 6.1.21.

δ_H (500 MHz, $CDCl_3$) 5.35 – 5.26 (1 H, m, C*H*CH_2O), 4.97 (1 H, t, *J* 7.0, C=C*H*CH_2CH_2), 4.87 – 4.84 (1 H, m, C=C*H*H), 4.83 – 4.80 (1 H, m, C=CH*H*), 3.92 – 3.64 (2 H, m, CHCH_2O), 3.54 – 3.42 (1 H, m, CH_2OC*H*C=C), 2.30 – 2.03 (6 H, m, 3 x CH_2), 1.70 (3 H, s, CH_3), 1.63 (3 H, s, CH_3), 1.68 – 1.52 (2 H, m, CH_2), 1.54 (3 H, s, CH_3). δ_C (126 MHz, $CDCl_3$) 124.46 (*C*HCH_2O), 110.82 (C=*C*H_2), 77.80 (CH_2O*C*HC=CH_2), 38.84 (CH_2), 31.49 (CH_2), 24.74 (CH_2), 17.55 (CH_3), 15.78 (CH_3), 14.67 (CH_3).

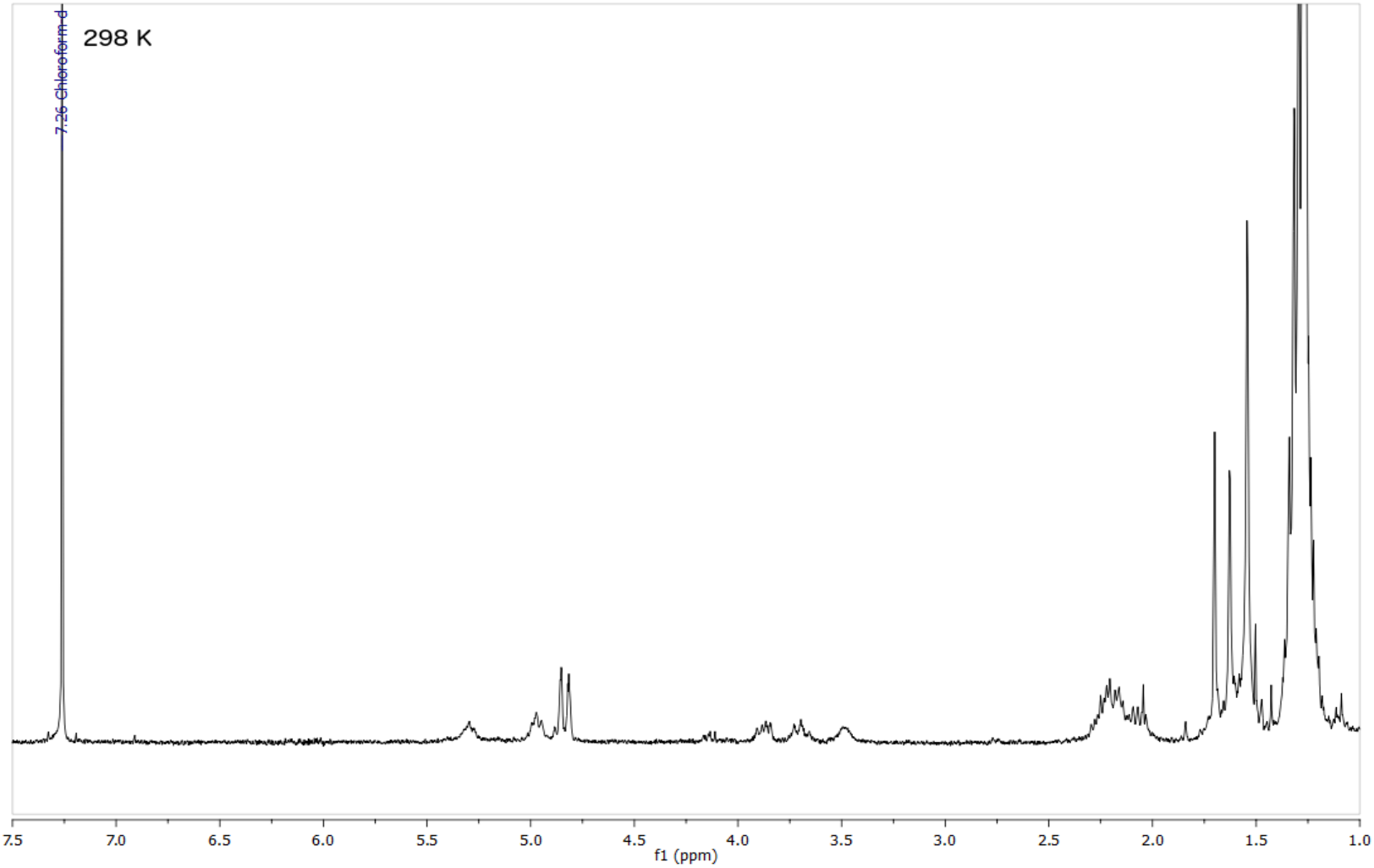


Figure 8.1 ^{1}H-NMR (300 MHz, $CDCl_3$, 298 K) of macrocyclic ether **158**

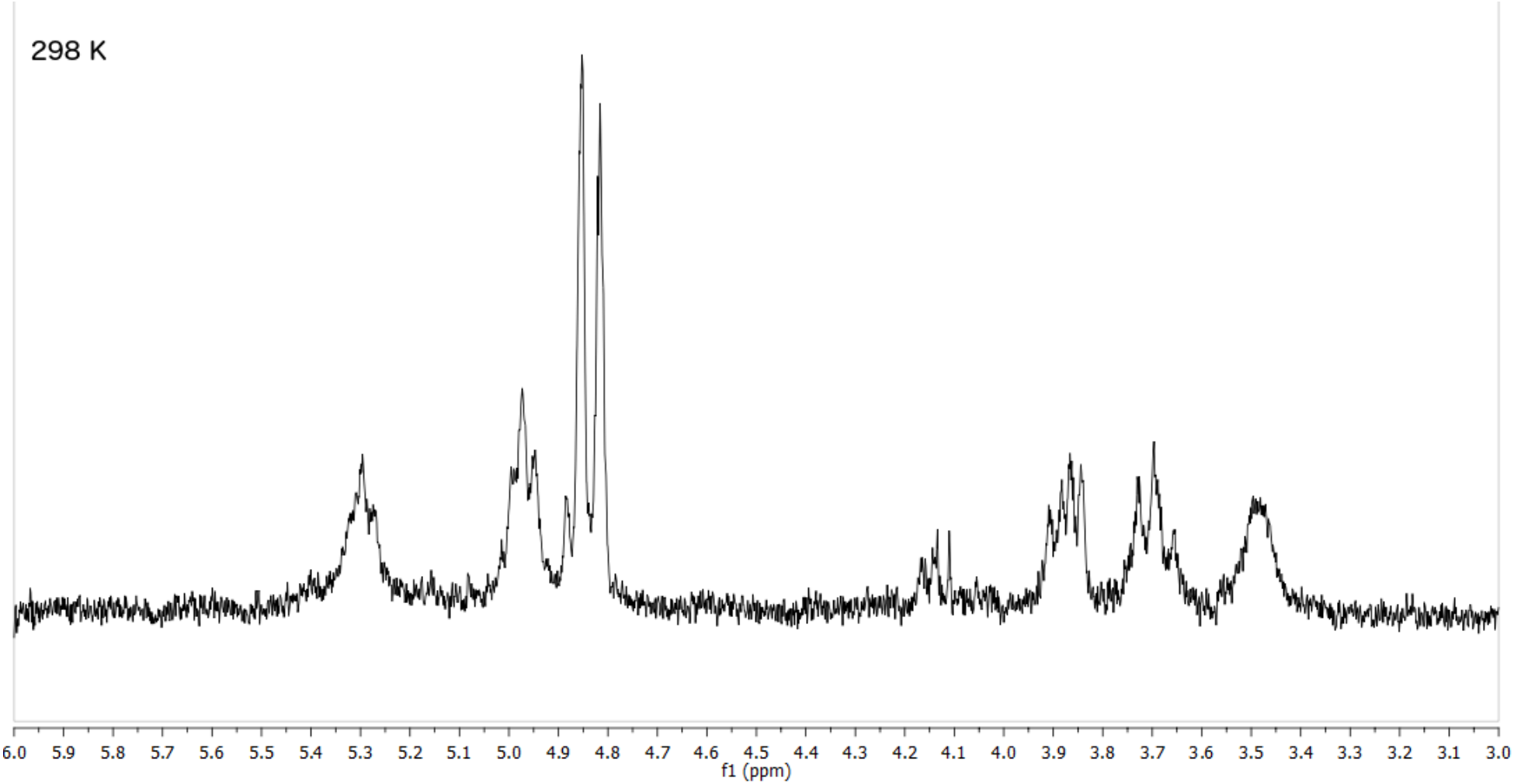


Figure 8.2 ^{1}H-NMR (300 MHz, $CDCl_3$, 298 K) of macrocyclic ether **158**. Expansion of the region from 3.0 to 6.0 ppm.

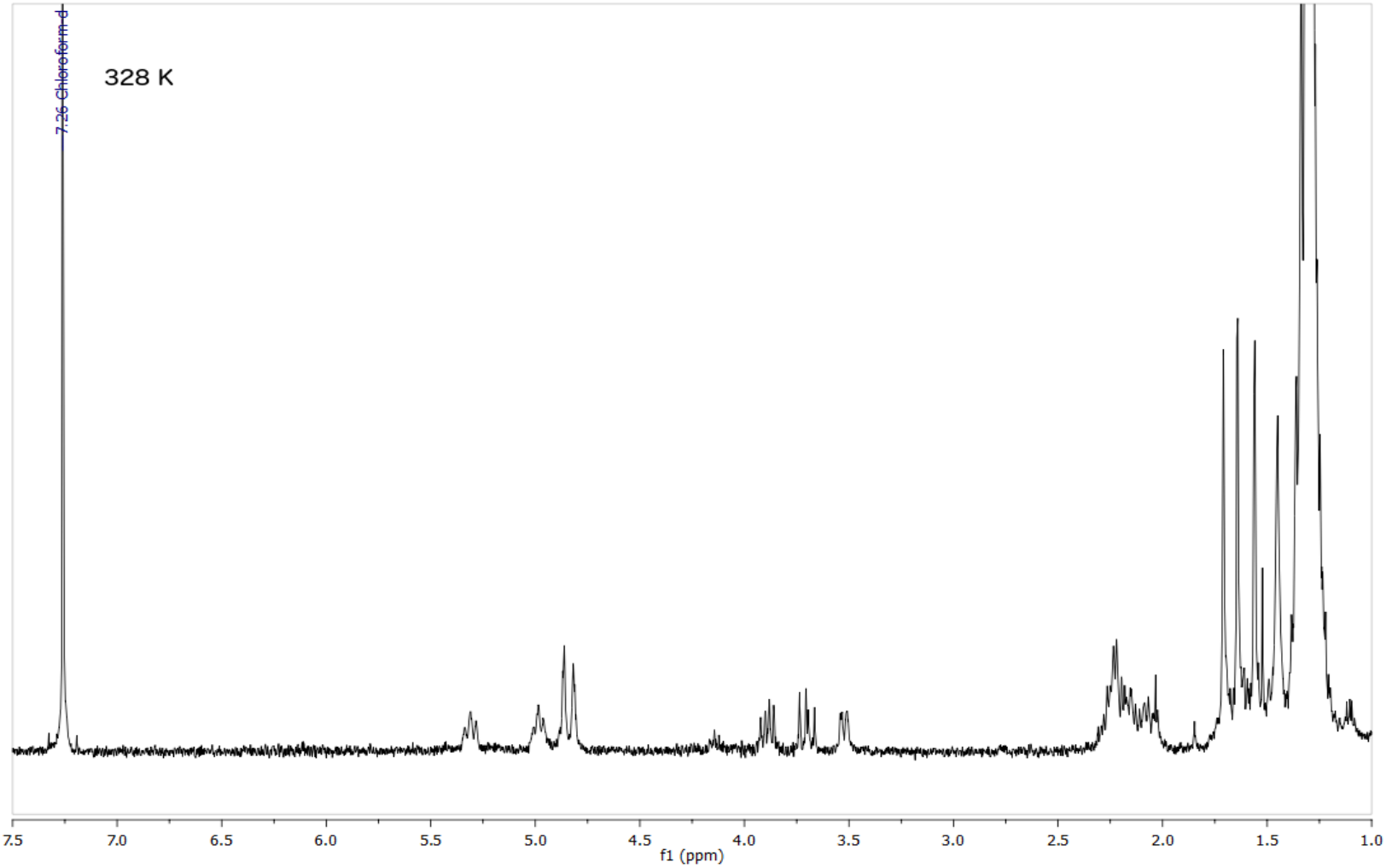


Figure 8.3 ^{1}H-NMR (300 MHz, $CDCl_3$, 328 K) of macrocyclic ether **158**.

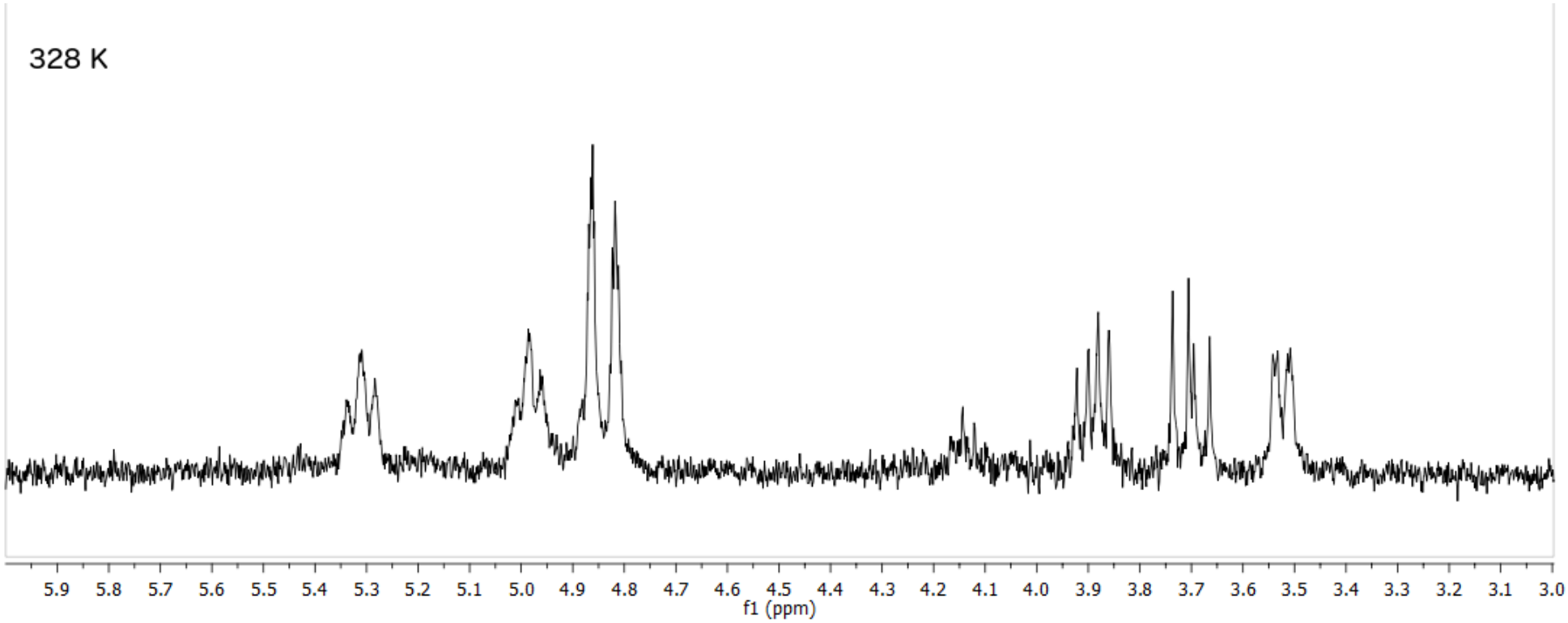


Figure 8.4 ^{1}H-NMR (300 MHz, $CDCl_3$, 328 K) of macrocyclic ether **158**. Expansion of the region from 3.0 to 6.0 ppm.

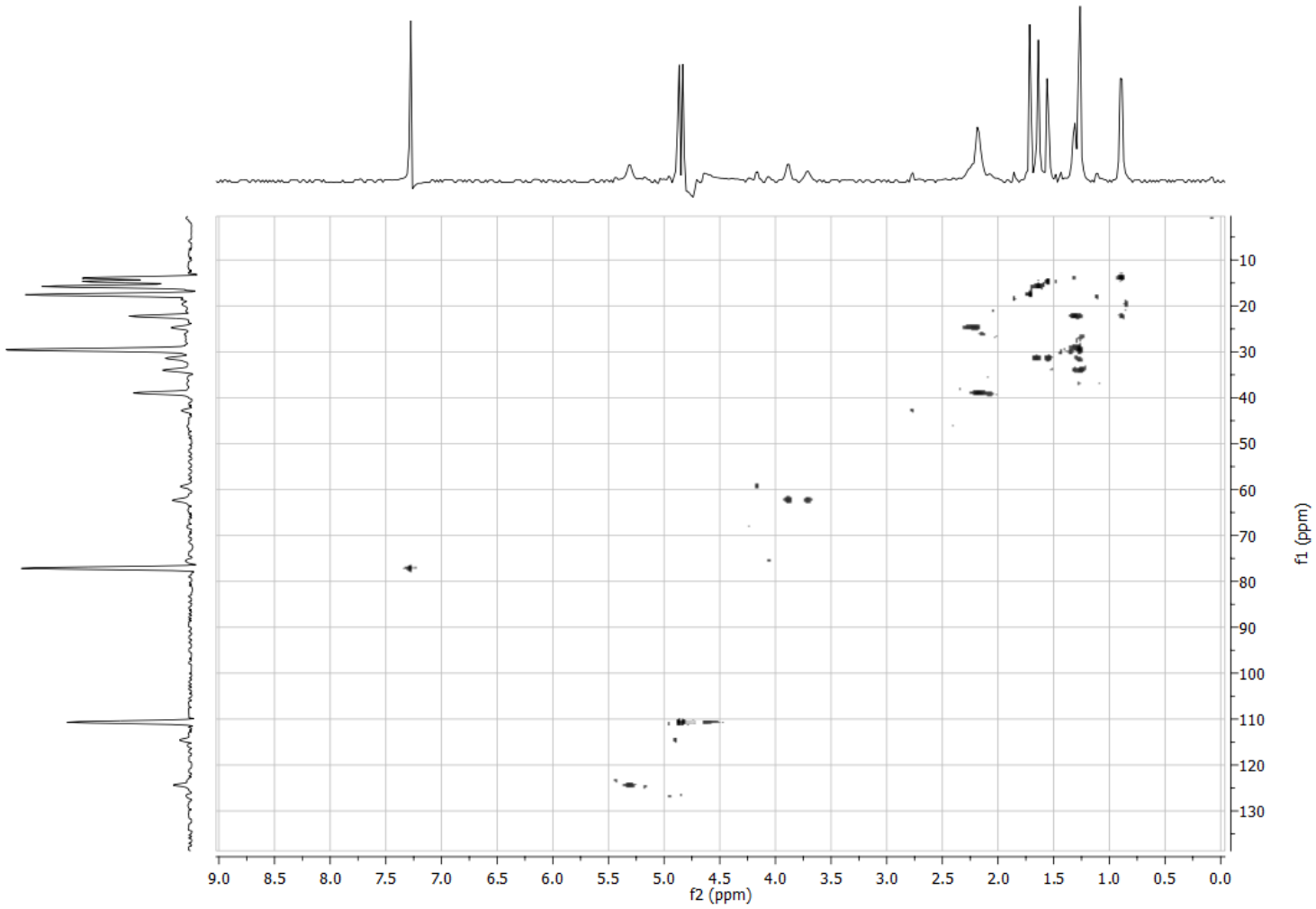


Figure 8.5 ^{1}H-^{13}C HSQC (300 MHz, $CDCl_3$, 298 K) of macrocyclic ether **158**.

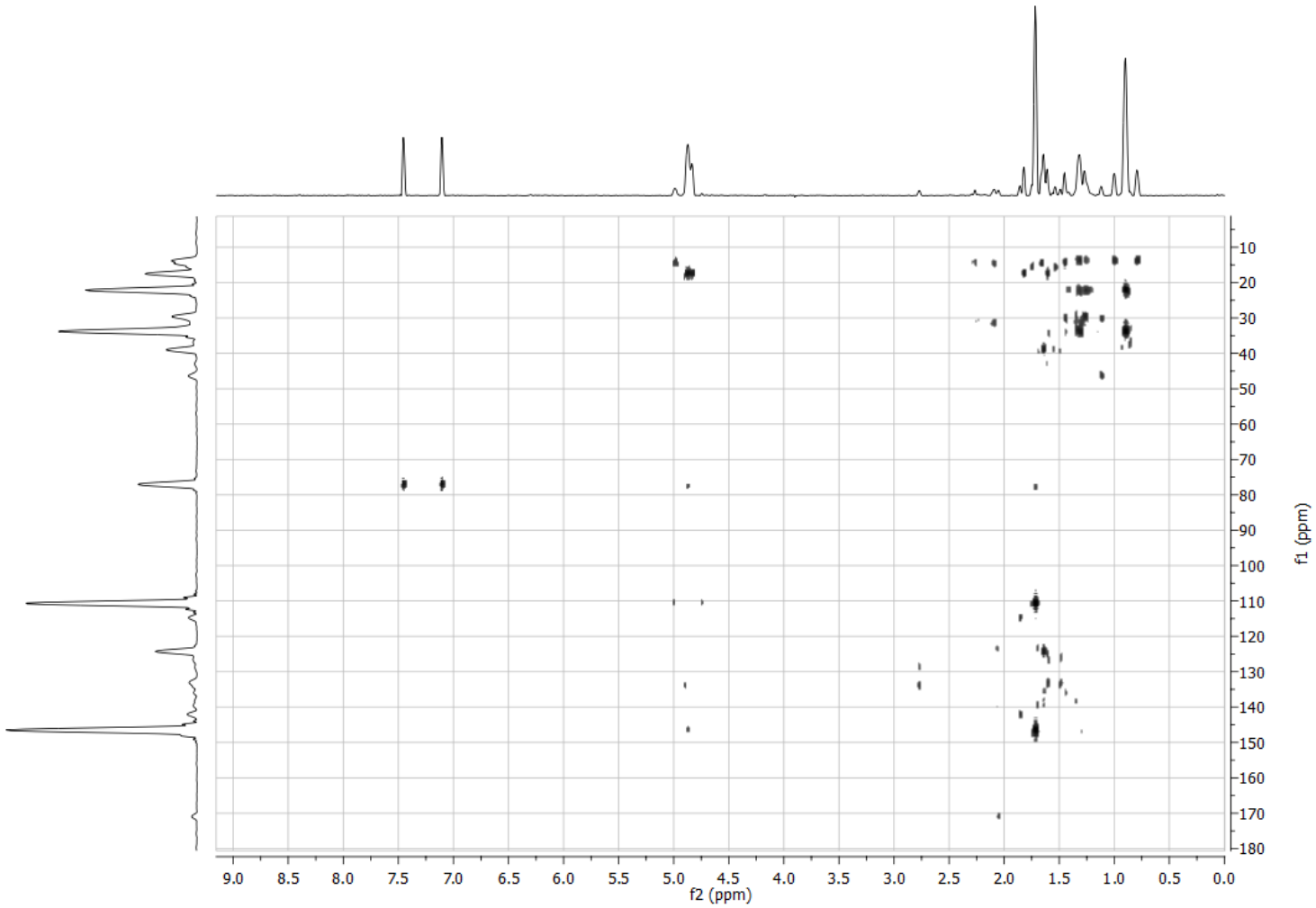


Figure 8.6 ^{1}H-^{13}C HMBC (300 MHz, $CDCl_3$, 298 K) of macrocyclic ether **158**.